

Lecture Notes in Electrical Engineering 517

Qilian Liang

Xin Liu

Zhenyu Na

Wei Wang

Jiasong Mu

Baoju Zhang *Editors*

Communications, Signal Processing, and Systems

Proceedings of the 2018 CSPA
Volume III: Systems

Lecture Notes in Electrical Engineering

Volume 517

Series Editors

Leopoldo Angrisani, Department of Electrical and Information Technologies Engineering, University of Napoli Federico II, Napoli, Italy

Marco Arteaga, Departament de Control y Robótica, Universidad Nacional Autónoma de México, Coyoacán, Mexico

Bijaya Ketan Panigrahi, Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India
Samarjit Chakraborty, Fakultät für Elektrotechnik und Informationstechnik, TU München, München, Germany

Jiming Chen, Zhejiang University, Hangzhou, Zhejiang, China

Shanben Chen, Materials Science & Engineering, Shanghai Jiao Tong University, Shanghai, China

Tan Kay Chen, Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore

Rüdiger Dillmann, Humanoids and Intelligent Systems Lab, Karlsruhe Institute for Technology, Karlsruhe, Baden-Württemberg, Germany

Haibin Duan, Beijing University of Aeronautics and Astronautics, Beijing, China

Gianluigi Ferrari, Università di Parma, Parma, Italy

Manuel Ferre, Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de Madrid, Madrid, Madrid, Spain

Sandra Hirche, Department of Electrical Engineering and Information Science, Technische Universität München, München, Germany

Faryar Jabbari, Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA

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

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

Alaa Khamis, German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt

Torsten Kroeger, Stanford University, Stanford, CA, USA

Qilian Liang, Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX, USA

Ferran Martin, Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain

Tan Cher Ming, College of Engineering, Nanyang Technological University, Singapore, Singapore

Wolfgang Minker, Institute of Information Technology, University of Ulm, Ulm, Germany

Pradeep Misra, Department of Electrical Engineering, Wright State University, Dayton, OH, USA

Sebastian Möller, Quality and Usability Lab, TU Berlin, Berlin, Germany

Subhas Mukhopadhyay, School of Engineering & Advanced Technology, Massey University, Palmerston North, Manawatu-Wanganui, New Zealand

Cun-Zheng Ning, Electrical Engineering, Arizona State University, Tempe, AZ, USA

Toyooki Nishida, Graduate School of Informatics, Kyoto University, Kyoto, Kyoto, Japan

Federica Pascucci, Dipartimento di Ingegneria, Università degli Studi "Roma Tre", Rome, Italy

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

Gan Woon Seng, School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore, Singapore

Joachim Speidel, Institute of Telecommunications, Universität Stuttgart, Stuttgart, Baden-Württemberg, Germany

Germano Veiga, Campus da FEUP, INESC Porto, Porto, Portugal

Haitao Wu, Academy of Opto-electronics, Chinese Academy of Sciences, Beijing, China

Junjie James Zhang, Charlotte, NC, USA

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Associate Editor (jasmine.dou@springer.com)

India

Swati Meherishi, Executive Editor (swati.meherishi@springer.com)

Aninda Bose, Senior Editor (aninda.bose@springer.com)

Japan

Takeyuki Yonezawa, Editorial Director (takeyuki.yonezawa@springer.com)

South Korea

Smith (Ahram) Chae, Editor (smith.chae@springer.com)

Southeast Asia

Ramesh Nath Premnath, Editor (ramesh.premnath@springer.com)

USA, Canada:

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries:

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

Christoph Baumann, Executive Editor (christoph.baumann@springer.com)

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

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

Qilian Liang · Xin Liu ·
Zhenyu Na · Wei Wang ·
Jiasong Mu · Baoju Zhang
Editors

Communications, Signal Processing, and Systems

Proceedings of the 2018 CSPS Volume III:
Systems

 Springer

Editors

Qilian Liang
Department of Electrical Engineering
University of Texas at Arlington
Arlington, TX, USA

Zhenyu Na
School of Information Science
and Technology
Dalian Maritime University
Dalian, China

Jiasong Mu
College of Electronic
and Communication Engineering
Tianjin Normal University
Tianjin, China

Xin Liu
School of Information
and Communication Engineering
Dalian University of Technology
Dalian, China

Wei Wang
College of Electronic
and Communication Engineering
Tianjin Normal University
Tianjin, China

Baoju Zhang
College of Electronic
and Communication Engineering
Tianjin Normal University
Tianjin, China

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-13-6507-2 ISBN 978-981-13-6508-9 (eBook)
<https://doi.org/10.1007/978-981-13-6508-9>

Library of Congress Control Number: 2019930976

© Springer Nature Singapore Pte Ltd. 2020

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

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

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

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

Contents

Artificial Intelligence and Deep Learning

Based on Deep Learning CSI Recovery for Uplink Massive Device Dynamic Internet of Thing	3
Yue Xiu, Wenyuan Wang, Yongliang Shen, and Zhongpei Zhang	
DEEP: Detection of Environmental Pollution Using Cooperative Neural Network	10
Yang Zhang	
Imbalanced Data Classification Method Based on Ensemble Learning	18
Yu Xiang and Yongping Xie	
Bayesian Method-Based Learning Automata for Two-Player Stochastic Games with Incomplete Information	25
Hua Ding, Chong Di, and Li Shenghong	
A Double Competitive Strategy-Based Learning Automata Algorithm	34
Chong Di, Mingda Guo, Jinchao Huang, and Shenghong Li	
A Learning Automata-Based Compression Scheme for Convolutional Neural Network	42
Shuai Feng, Haonan Guo, Jichao Yang, Zhengwu Xu, and Shenghong Li	
Improved Efficient Dictionary Learning with Cross-Label and Group Regularization	50
Tian Zhou, Sujuan Yang, Jian Xiong, Jie Yang, and Guan Gui	
Multi-task Cascaded Convolutional Neural Networks for Real-Time Dynamic Face Recognition Method	59
Bin Jiang, Qiang Ren, Fei Dai, Jian Xiong, Jie Yang, and Guan Gui	

Based on the Predicted Blocking Virtual Machine Load Balancing Scheduling Strategy	67
Youhui Jiang	
Research on Fault Diagnosis Based on Artificial Neural Network	73
Rui Liu	
Virtual Samples for Cloud Classification via Supervised Learning	80
Shuang Liu, Mei Li, Zhong Zhang, Mingzhu Shi, and Xiaozhong Cao	
Deep Learning for Optical Character Recognition and Its Application to VAT Invoice Recognition	87
Yu Wang, Guan Gui, Nan Zhao, Yue Yin, Hao Huang, Yunyi Li, Jie Wang, Jie Yang, and Haijun Zhang	
Deep Learning Based Detection Method for SDN Malicious Applications	96
Chi Yaping, Yu Yuzhou, and Yang Jianxi	
Survey of Big Data Application Technology on Multimedia Data of Public Security	105
Huibo Li, Yinan Jiang, Yunxiang Yang, Jing Guo, Xiaocheng Hu, Ke Guo, Bo Zhang, and Jing Cheng	
Pedestrian Detection Based on Deep Neural Network in Video Surveillance	113
Bo Zhang, Ke Guo, Yunxiang Yang, Jing Guo, Xueying Zhang, Xiaocheng Hu, Yinan Jiang, and Xinhai Zhang	
Robust Model for Chinese License Plate Character Recognition Using Deep Learning Techniques	121
Amr Abdussalam, Songlin Sun, Meixia Fu, Yasir Ullah, and Safwan Ali	
A Multi-label Scene Categorization Model Based on Deep Convolutional Neural Network	128
Gaofeng Zhao, Wang Luo, Yang Cui, Qiang Fan, Qiwei Peng, Zhen Kong, Liang Zhu, and Tai Zhang	
Discriminative Structured Dictionary Learning for Face Recognition	136
Ying Zhu	
An Improving Data Stream Classification Algorithm Based on BP Neural Network	145
Baoju Zhang, Guilin Wang, and Lei Xue	
Facial Fatigue Detection Based on Machine Learning	154
Dewei Zheng, Shaohua Cui, and Chenglin Zhao	

Genetic Algorithm-Based Beamforming Using Power Pattern Function 159
 Shuoguang Wang, Shiyong Li, and Houjun Sun

Short-Text Sentiment Analysis Based on Windowed Word Vector 168
 Dongmei Zhao, Yingli Shen, Yabo Shen, Yong Ma, Yun Jin, Shidang Li, and Mingliang Gu

FFnet: Residual Block-Based Convolutional Neural Network for Crowd Counting 175
 Fei Lei, Qinyu Zhang, Peng Zhao, Dongqiang Chen, Xiu Chen, and Xiao Han

Deep Learning-Based V2V Channel Estimations Using VNETs 184
 Qi Song, Tian Lan, Xuanxuan Tian, and Tingting Zhang

Research on Optimization of Evacuation Path Based on Fuzzy Control and Intelligent Agent Technology—With Open Block Emergency Evacuation as Example 193
 Fu Erkang, Zhang Lingfei, Li Xinyun, and Zhang Yujia

Feature Engineering of Click-through-rate Prediction for Advertising 204
 Jie Ren, Jian Zhang, and Jing Liang

Application of Fuzzy C-Means Algorithm in Complex Background Image Segmentation of Forensic Science 212
 Zhuang Chen, ChunYu Li, ZhanQing Jiang, and Yongqiang Zhao

Steady-State Performance Analysis of Quaternion-Valued Least Mean Square Adaptive Algorithm 218
 Sen Li, Fengzhi Liu, Bin Lin, Rongxi He, and Xiaomei Zhu

Research on Concept-Drifting Data Stream Based on Fuzzy Integral Ensemble Classifier System 225
 Baoju Zhang, Yidi Chen, and Lei Xue

An Improved Speech Synthesis Algorithm with Post filter Parameters Based on Deep Neural Network 233
 Shunjie Dong, Chunyang Li, and Hong Zhang

Research on Machine Translation Model Based on Neural Network 244
 Zhuoran Han and Shenghong Li

Multi-node Repair Based on GA_PSO with Fractional Regenerating Code Combined with Prior Replication 252
 Niannian Wang, Ye Wang, Jia Yu, and Siyun Chen

A Super-Resolution Reconstruction Algorithm Based on Learning Improvement	261
Han Gao, Xinwei Li, and Aiping Jiang	
Construction of the Intelligent Small-Sized Plant	270
Yadi He, Yuanyuan Wang, Pengli Zhu, Shuangshuang Zhang, Feichao Zhao, and Lidong Zhang	
Integrating Heterogeneous Datasets by Using Multimodal Deep Learning	279
Fariba Khoshghalbavash and Jean X. Gao	
A Multi-view Deep Learning Approach for Detecting Threats on 3D Human Body	286
Zhicong Yan, Shuai Feng, Fangqi Li, Zhengwu Xu, and Shenghong Li	
An Incremental Scheme with Weight Pruning to Train Deep Neural Network	295
Haonan Guo, Zhicong Yan, Jichao Yang, and Shenghong Li	
A Deep Learning Method of Moving Target Classification in Clutter Background	303
Ningyuan Su, Xiaolong Chen, Xiaoqian Mou, Lin Zhang, and Jian Guan	
Optical System	
Three-Dimensional Laser Scanning for the Bridge Deformation of Shanghai Maglev Train	315
Yanwen Wu, Lei Zhang, V. Badenko, and R. D. Garg	
Access Control System Based on Visible Light Communication	330
Xinpeng Xue, Jinpeng Wang, Ying Yu, and Nianyu Zou	
Portable 3D Laser Scanner for Volume Measurement of Coal Pile	340
Wang Zhang, Deshan Yang, Ying Li, and Wenhai Xu	
Development of a New Type of Long Base Stress Sensor for Hull Structure	348
Wei Wang, Libo Qiao, and Yuliang Li	
Research of Single-Phase Photovoltaic Grid-Connected System Based on MATLAB	354
Hao Yang	
Analysis of Output Characteristics of Photovoltaic Arrays Under Shaded Conditions	363
Hao Yang	

A Novel Fiber Polarization Splitter Based on Orthogonal Dual Core 378
 Yu Hou

Theoretical Research on the Strain Characteristics of the Selective-Filling Birefringent Photonic Crystal Fiber 384
 Jin Zuo, Tingting Han, and Jingping Yang

Design and Analysis of a Novel Metamaterial Structure Realized in Low-Frequency Band 392
 Xiu Zhang, Honghao Sun, and Xin Zhang

Evanescence-Mode Waveguide Filter with Transmission Zeros Created by Shorted Waveguide Shunted in Coupling Region 399
 Liu He, Xingjian Zhong, Zhendong Fan, Quan Zhang, and Wei Zhang

Time-Aware Routing and Spectrum Assignment Assisted by 3D-Spectrum Auxiliary Graph in Elastic Optical Networks 405
 Li Zhang, Cunqian Yu, and Rongxi He

Routing and Spectrum Assignment for Software-Defined Elastic Optical Networks 413
 Ziwei Lin, Rongxi He, and Tongtong Liu

Numerical Implementation of a Wideband Chaotic Light Based Ring-and-Spur Long-Reach Passive Optical Network: Architectures and Real-Time Secure Communications 422
 Xinyu Dou, Hongxi Yin, and Bin Wu

A Theoretical Method for Constructing the Boundary of the Color Domain of the System 429
 Sile Liu, Yan Li, Yali Huang, and Dan Zhen

Research and Evaluation of Color Gamut Extension Algorithm Based on Image 436
 Yali Huang, Yan Li, Sile Liu, and Dan Zhen

A Novel Hinge Structure of Fiber Bragg Grating Acceleration Sensor 445
 Yuliang Li, Wei Wang, and Libo Qiao

Circuit System Design

Study of Giant Magnetostrictive Thin Film Pressure Sensor Based on Villari Effect 459
 Liyuan Dong, Shaopeng Yu, Tingting Han, Bowen Wang, and Xinxin Cui

Research on ATML-Oriented Test Application Development Platform 468
 Jin Luo, Hua Yang, and Jing Huang

A School Violence Detection Algorithm Based on a Single MEMS Sensor	474
Jifu Shi, Liang Ye, Hany Ferdinando, Tapio Seppänen, and Esko Alasaarela	
Multichannel High-Speed Data Caching System on FPGA for RAID Storage	482
Haixin Wang, Xue Bai, and Qiongzhi Wu	
Research on Multiple Switched Flat-Top Beam Smart Antenna	492
Wei Liang, Xiuzhen Luan, and Kejun Tan	
Design of Portable Power Supply System	499
Xinqiang Zhang, Jiaqi Li, Ya Tu, Changyun Ge, and Xiujie Zhao	
Design of Direct Current Motor Servo Control System Based on SOPC	505
Tu Ya, Jialu Du, and Peng Cui	
Design and Implementation of a Numerical Control High-Precision Low-Temperature Drift Constant Current Source	513
Xiwei Guo, Chaochao Yang, and Deliang Liu	
Compact Hybrid-Integrated Circular Polarized Double-Ring Antenna for Satellite Application	523
Sihao Chen, Dongliang Fei, and Lianxing He	
Influence of Center of Mass Movement on Steering Characteristics of Front-Wheel Steering Vehicles and Four-Wheel Steering Vehicles	529
Shuaijun Yang, Yinshan Wang, and Shaoyun Lu	
Reflectors for Multiple Applications Based on Flux Compensation Method	538
Wang Guangzhen, Hou Yu, and Li Jia	
The Design of Low-Power and High-Precision Electronic Scale Based on Single Chip	549
Dun Liu and Fang Qu	
The Main Damage Characteristics of Semiconductor Devices Under High-Power Microwave	558
Kaibai Chen	
A Configuration-Based Automatic Test System for Armored Vehicle Information Acquisition Devices	564
Linghui Zhang, Ruina Zhao, Lei Guo, Shao Li, and Weizheng An	

Efficient Sensitivity Analysis of Dynamic Neuro-space Mapping for Transistor Modeling 575
 Lin Zhu, Jian Zhao, and Wenyan Liu

Open-Loop Carrier Synchronization Design and Its FPGA Implementation for Short Burst Communication at Low SNR 584
 Wen Che and Jinhui Fang

Design and Simulation of a Sector-Shaped Microstrip Antenna Fed by a T-Shaped Probe 592
 Weiyang Mao and Lizhong Song

Design of a Wideband Receiving Antenna for High-Frequency Ground Wave Radar 599
 Hongbo Li, Yang Song, and Changjun Yu

Optimal Design of S Band with Graphene Frequency Multiplier 606
 Huili Chen, Yong Fang, Xiaoling Zhong, Qingyan Song, Xueshi Hou, and Haoxuan Sheng

Intelligent Street Lamp Management System 612
 Shimin Wang and Yongjie Yang

Control System Design of Reactor Robot for Object Salvaging Underwater 621
 Xiaochen Huang, Lingyu Sun, Xiaojun Zhang, Manhong Li, and Minglu Zhang

Application of EWT and PSO-SVM in Fault Diagnosis of HV Circuit Breakers 628
 Bing Li, Mingliang Liu, Zijian Guo, and Yamin Ji

A Dynamic Programming Track-Before-Detect Algorithm with Adaptive State Transition Set 638
 Hao Xing, Jidong Suo, and Xiaoming Liu

New Processing Method Based on Intelligently Manufacturing Blade with Multiple Space and Compound Angles 647
 Yuanyuan Wang, Yadi He, Pengli Zhu, Lidong Zhang, and Feichao Zhao

Design of Three-Phase AC-Voltage Regulator Energy-Saving System by SOPC 656
 Tu Ya, Jialu Du, and Peng Cui

A Flexible Broadband Single RF Architecture Based on Time-Modulated Array 665
 Xiaofeng Ling and Nan Wang

An Open Integrated Electronic System Software Architecture Design for Launch Vehicle	673
Feng Zhang, Guo-wei Yao, Qian Wang, and Yun Xia	
Read and Write Performance Research and Optimization for eMMC Device Driver	683
Yanlin Chen, Songyan Liu, Yifei Niu, Huan Liu, and Xiaowen Wang	
Power Analysis Method Based on DC Bus Voltage Waveform	687
Yicheng Wang, Zhiyong Huang, Xiaolong Luo, and Yinguo Huang	
Automatic Rapid Electrokinetic Pattern System	699
Yanwei Wang and Jiaqi Zhen	
Wireless System	
A Secondary Surveillance Radar Data Analysis Technique Based on Geometrical Method	707
Jing Gao, Jie Zou, and Ning Guo	
Spectral Efficiency Analysis of Hybrid Precoding in Millimeter Wave MIMO Systems	716
Jing Li, Dianwu Yue, and Meng Wang	
An FPGA-Based Balanced and High-Efficiency Two-Dimensional Data Access Technology for Real-Time Spaceborne SAR	724
Tianyuan Sun, Bingyi Li, Xiaoning Liu, and Yizhuang Xie	
Research on the Realization of Excitation Signal SPWM of Radar Synchronous Motor Based on DSP	733
Liang Zhao and ShiHui Zheng	
Sensitivity of Radar Variables to Signal-to-Noise Ratio in Dual-Polarization Weather Radar	740
Xiao-yi Wang and De-bin Su	
Analysis of Characteristics of Atmospheric Structure Constant of Refractive Index Based on Wind Profiler Radar in Precipitation	745
Yating Li, Debin Su, and Xingang Fan	
Modeling and Simulation of Auto Parts Production Line Based on Petri Net	753
Shuqi Jin, Shaohua Cui, and Chenglin Zhao	
A Radar Electromagnetic Environment Sensing Method Based on Cyclic Spectral Algorithm	759
Jurong Hu, Yu Zhang, Xujie Li, Xiaoyong Ni, and Evans Baidoo	

Performance Analysis of Mid-Far Infrared Wave in Satellite-Ground Link 769
 Meng Jing, Li Shuai, Lin Qingqing, and Liu Shuai

Seawater Antenna for High-Frequency Surface Wave Radar 776
 Linwei Wang, Changjun Yu, and Haorong Wang

An Underwater Sensor Networks Based Cooperative Positioning System for Falling Water Containers 784
 Manyu Xu and Ying Wang

Internet of Things

Wisdom Farm Internet of Things Software Design and Selection Program 791
 Chaoyan Man, Li Guo, Yang Gao, and Yu Zhang

Architecture Design of Modern Marine IoT Cloud Server Software 799
 Ling Yu, Yang Gao, Chaoyan Man, and Yu Zhang

Three-Dimensional Structure Measurement and Optimization Method of Indoor Scene Based on Single Image 804
 Ronghe Wang, Xinhai Zhang, Bo Zhang, Jianning Bi, and Xiaolei Guo

The Intelligent Supervision System of Farm Based on “Internet + BDS + GIS” 813
 Wei Fu, X. R. Dong, Weiyi Shuai, Mingqi Yang, and Jun Wang

A Platoon-Based Vehicle-to-Vehicle Connectivity Enhancing Scheme Under Bidirectional Highway Scene 821
 Kun Xu

An Intelligent Water Regimen Monitoring System 829
 Yaping Fan, Heng Dong, Ying Jiang, Jinqiu Pan, Shangang Fan, and Guan Gui

Attitude Stabilization Control Method for Quadrotor UAV Based on ADRC 836
 Sen Yang, Leiping Xi, Guanghong Gong, and Hairui Dong

Fuzzy Adaptive PID Control for Translational Flight of a Tail-Sitter UAV 845
 Leiping Xi, Dizhou Zhang, and Sen Yang

Renewable Energy-Aware IoT Data Aggregation for Fog Computing 852
 Yusong Fu, Dapeng Li, Feng Tian, and Yongan Guo

Renewable Energy Powered IoT Data Traffic Aggregation for Edge Computing 861
 Cunchao Peng, Dapeng Li, Feng Tian, and Yongan Guo

Design and Implementation of Sensory Data Collection and Storage Based on Hadoop Platform 870
 Zhen Bai, Shaohua Cui, and Chenglin Zhao

Design of a Data Acquisition and Transmission System for Smart Factory Based on NB-IoT 875
 Ruijian Zhang, Shaohua Cui, and Chenglin Zhao

Design of Laboratory Monitoring Management System Based on Internet of Things 881
 Nannan Chong

Online Taxi-Hailing Platform Using Blockchain Technology 891
 Chuang Ma, Dou Hu, Xinyi Wang, and Wenlong Liu

The Intangible Culture Heritage “New Ecology” Under Emotional Demand of Interactive Experience 898
 Dongna Cai, Yuning Li, Zhi Li, and Yongjian Huai

Research on Optimization Model of Storage Capacity Based on the Consortium Blockchain 911
 Xiaotian Wei, Jiahua Chen, and Zhihui Li

Blockchain Storage Analysis and Optimization of Bitcoin Miner Node 922
 Junying Gao, Bo Li, and Zhihui Li

Improved Ant Colony Optimization Algorithm for Optimized Nodes Deployment of HAP-Based Marine Monitoring Sensor Networks 933
 Jianli Duan, Yuxiang Liu, Bin Lin, Yuan Jiang, Fen Hou, and Wantong Li

A Resource Allocation Algorithm for D2D Multicast in Internet of Vehicles 942
 Wei Wu, Muchen Yu, Xuanli Wu, and Guoan Bi

Services Ranking Based Random Access Scheme for Machine-Type Communication 950
 Lu Dai, Yunjian Jia, Zhengchuan Chen, Liang Liang, and Guojun Li

A Novel Energy Harvesting Scheme in Interference Networks with UAVs 962
 Bingcai Chen, Manrou Yang, and Yu Chen

Low-Power Wide Area Networks: Changes for Smart Grid 967
 Liang Wan, Yirui Huang, Weihua Li, Yu Zhang, and Zhijian Zhang

A Low Energy Consumption Ocean Environment Information Collection System Designing 975
 Qiuming Zhao, Hongjuan Yang, and Bo Li

System Security

Cyberspace Security Evaluation Technology on the Condition of Attack and Defense Confrontation 983
 Zhang Bo, Wen Tao, Lei Jing, Yang Yunxiang, and Guo Jing

Interoperability Performance Analysis and Assessment of B2a/L5/E5a 990
 Weiyi Shuai, X. R. Dong, Di Yan, Jun Wang, and Wei Fu

Physical Layer Security of a Buffer-Aided Relay Selection for Underlay Cognitive Radio Network 999
 Jian Jia, Ting Jiang, Wei Guo, and Xiaoying Qiu

Generative Adversarial Network-Based Credit Card Fraud Detection 1007
 Xiaobo Xie, Jian Xiong, Liguo Lu, Guan Gui, Jie Yang, Shangan Fan, and Haibo Li

A Novel Improved System Based on CVSS 1015
 Wen Tao, Zhang Yuqing, Zhang Bo, Lei Jing, Yang Yunxiang, and Guo Jing

Efficient Traffic Coordination Strategies at Intersections Using Multiple Collision Sets 1022
 Yangan Mo, Mengqi Wang, Tingting Zhang, and Hongguang Xu

Key-Controlled PEG-LDPC Algorithm Design and Its Application in Secure Communication 1031
 Zhiping Shi, Shujun Zhang, Fengcheng Lyu, Fan Bu, Hongxia Sun, and Qian Zhang

User Relationship Privacy Protection on Trajectory Data 1038
 Zi Yang, Mingda Yang, and Bo Ning

Research on Modulation Algorithm Based on Physical Layer Encryption 1046
 Xiang Li and Yueyong Zhang

Threat Modeling for Cyber Range: An Ontology-Based Approach 1055
 Lei Gong and Yu Tian

An Analysis of a New Detection Method for Spear Phishing Attack 1063
 Yaping Chi, Zhiting Ling, Xuejing Ba, and Shuhao Li

Memory Confidentiality and Integrity Protection Technology 1073
Hongjin Wang, Huimin Meng, Nianmin Yao, and Yishun Cheng

Privacy Preservation of Semi-structured Data Based on XML 1081
Cheng Shi, Mingda Yang, and Bo Ning

A Research on Detection Algorithm of Vehicle Illegal U-Turn 1089
Wenrui Wang, Yiran Zhou, Qiurun Cai, and Yiying Zhou

A Dredge Traffic Algorithm for Maintaining Network Stability 1100
Yanan Zhao, Fusheng Dai, and Jun Shi

Attacks Detection Method Based on Free Space Quantum Secure Direct Communication 1109
Jinlong Liu, Zhutian Yang, and Zhilu Wu

Research on Model and Application of Elevator Safety Remaining Service Life 1117
Qi Li, Zhenfeng Shi, and Shang Sun

Safety Evaluation Model of Intelligent Elevator Cloud Management Platform 1126
Shang Sun, Zhenfeng Shi, and Qi Li

A Chatbot Design Method Using Combined Model for Business Promotion 1133
Jie Zhang, Hao Huang, and Guan Gui

Design of Monitoring and Warning System for Dangerous Gases in Oil Tank 1141
Yuelan Ji, Yongjie Yang, and Zhongxing Huo

Design of Standard Cell for Anti-radiation 1153
Bei Cao, Pengfei Wu, and Danyang Qin

Design and Implementation of Intelligent Crib-Based on Android 1168
Junjiao Zhang, Yongjie Yang, and Zhongxing Huo

Big Data-Based Attack Scenario Reconstruction Architecture in Smart Grid 1178
Liang Guo, Qianqian Jin, Ying Liu, Yuanyi Xia, and Han Hu

Present Situation Analysis and Future Development of the Command and Control Console 1188
Wenli Jin and Fang Bai

Research on Comprehensive Defense Technology of the Emergency Command Vehicle 1194
Fang Bai, Yangyang Hu, and Ruigang Zhao

An Introduction of Cognitive Electronic Warfare System	1202
Huaji Zhou	
The Application of Geomagnetic Models in Flight Path Planning	1211
Zhiyuan Hang and Zhifang Wang	
Author Index	1215

Artificial Intelligence and Deep Learning



Based on Deep Learning CSI Recovery for Uplink Massive Device Dynamic Internet of Thing

Yue Xiu^(✉), Wenyuan Wang, Yongliang Shen, and Zhongpei Zhang

National Key Laboratory of Science and Technology on Communications,
University of Electronic Science and Technology of China, Chengdu, China
xiuyue12345678@163.com

Abstract. This paper investigates uplink massive MIMO communication scenarios in dynamic Internet of things (IoT) networks. In this paper, dynamic IoT mainly consists of the Internet of vehicles (IoV) and the original IoT network. Because the speed of vehicle is very fast, the number of users is constantly changing in the IoT network, which leads the structure of the IoT network to change. We mainly consider how to obtain the channel state information (CSI) of active users. Due to active users and inactive users, the system model is considered a sparse structure. This structure inspired us to give an algorithm suitable for the sparse structure and obtain more accurate channel state information of dynamic IoT networks, though these numerical results, under the premise of guaranteeing performance, can greatly reduce the complexity of the algorithm.

Keywords: Deep learning (DL) · Dynamic Internet of things (DIoT) · Adaptive compressive sensing · Massive MIMO · Sparse structure

1 Introduction

In the past few years due to the increase of network users, the rapid growth of mobile vehicles at the same time, the pressure of Internet of things (IoT) network is bigger, especially when the Internet of vehicles (IoV) is becoming more and more common. To solve these problems, more powerful techniques have been discussed. With the advent of 5G networks, the solution of such problems has become a hot topic again. The upcoming 5G network can increase capacity by 1000 times, by 2020 [1]. 5G massive MIMO can greatly improve the physical layer capacity and spectrum efficiency [2]. At the same time, with the application of large-scale antenna arrays, the massive MIMO system can also utilize the advantages of space division multiplexing, which can increase the spectrum efficiency by 10 times and increase the energy efficiency by 100 times [3]. It is necessary for massive MIMO system to obtain channel state information (CSI) at the base station. In TDD massive MIMO systems, channel reciprocity can be used to obtain channel information [1]. However, in the Internet of things (IoT), the overhead of channel estimation is still difficult to bear because the system needs to provide service to a large number of users in a limited channel coherence time.

In recent years, with the development of artificial intelligence, the deep learning algorithm has been the most powerful classification algorithm. Deep learning is a popular method of machine learning, which has made great progress in the field of traditional machine learning. IoT networks and smart city deployments are generating a large amount of sensor data that needs to be analyzed. The application of deep learning in these fields has been an important research topic. In [4], a method of data recognition for large data obtained by IoT networks is given by deep learning. In [5], a long–short memory network is improved and the deep learning algorithm obtains a large amount of data collected by the IoT network, and we found that the deep learning network is mainly used for data hiding detection error in collecting data. Because the IoT network is a big data network, to obtain useful data from large amounts of data has become an important technique [6]. Previous work [1, 2, 4, 7] has shown that the deep learning approach to solving sparse signal problems has the potential to offer significant improvements, in both accuracy and complexity over traditional algorithms like orthogonal matching pursuit (OMP) and sparse Bayesian learning (SBL). Deep learning has gained great popularity due to its widespread success on many problems, and we consider the application of deep learning to solve the sparse recovery problem. What is more, we combine the detection algorithms mentioned in document [6, 8] with the related AMP family algorithms. We obtain a completely new adaptive compress sensing algorithm.

Our work is mainly focused on several aspects. First, we consider the design of a deep learning network algorithm for dynamic IoT network by using AMP networks. The users in the dynamic IoT network are divided into active users and inactive users. And a network structure of probability learning algorithm for cascade graphs is given. The structure is inspired by ensemble method of machine learning. The probability of active users is set to 1, the probability of inactive users will be set to 0, and we found that effect of SNR and the length of the training signal cascade structure are low. Second, it is compared with the traditional approximate message propagation algorithm.

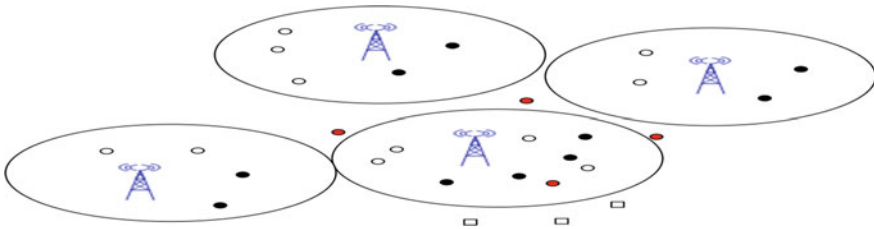


Fig. 1. System model of the massive device communication network

2 System Model

Consider the uplink of multiple-cell network consisting of many BS and user, and a single cell network consists of N users. Because we consider the users of single cell include dynamic users and output users, the number of users of single cell is time-

varying and the number of users is a function on the time t , which can be rewritten as $N(t)$. Denoted by the user set $N = \{1, \dots, N(t)\}$, it is assumed that the BS is equipped with M antennas and the number of antennas N is giant, while each user is equipped with one antenna. The complex uplink channel vector from user i to the BS is denoted by $\mathbf{h}_i \in \mathbb{C}^{1 \times N}$. This paper adopts a block-fading model, in which all the channels follow independent quasi-static flat fading within a block of coherence time, where \mathbf{h}_i 's remain constant but vary independently from block to block (Fig. 1).

The sporadic nature of user traffic is modeled as follows. Because of the dynamic users, the number of users' single cell are varying with time. What is more, we assume that in the t times, the users of single cell are synchronized and each user decides in each coherence block whether or not to access the channel with probability π in an i.i.d. manner. Thus, within each coherence block, only a subset of the users is active. In each particular block, we define the user activity indicator for user i as follows:

$$\hat{\mathbf{h}}_i \begin{cases} 1 \\ 0 \end{cases} \quad (1)$$

so that $P(\alpha_i) = \pi$, $P(\alpha_i) = 1 - \pi$, $\forall i$. Further, we define the set of active users within a coherence block as $\Omega = \{i : \alpha_i = 1, i = 1, \dots, N(t)\}$. We denote the number of active users as K , i.e., $K = |\Omega|$. The overall channel input-output relationship is modeled as:

$$\mathbf{y} = \sum_i \mathbf{h}_i \alpha_i x_i + \mathbf{z} = \sum_{k \in \Omega} \mathbf{h}_k x_k + \mathbf{z} \quad (2)$$

where \mathbf{x}_i , \mathbf{y} , respectively, are the user transmitted signal, the channel output at the BS, and the additive white Gaussian noise (AWGN) vector distributed as $CN(0, \sigma^2 \mathbf{I})$. For simplicity, this paper assumes no power adaptation so that all the devices transmit at a constant power ρ , i.e., $(\mathbb{E}[x_i])^2 = \rho$. Further, we do not account for intercell interference. The objective of the BS is to detect the number of users where the users are active; after finishing these works, the BS estimates the channel information. The channel estimation is important to decode their transmitted messages within each coherence block and compute the QoS.

Consider the second phase of massive device transmission in which each user sends its pilot sequence synchronously through the channel. Define ρ^{pilot} as the identical transmit power of the active users in the second transmission phase. The transmit signal of user n can be expressed as $\alpha_i \sqrt{\xi} \mathbf{a}_i$, where $\xi = L \rho^{pilot}$ denotes the total transmit energy of each active user in the first phase. The received signal at the BS is then

$$\mathbf{Y} = \sqrt{\xi} \sum_{i \in N} \alpha_i \mathbf{a}_i \mathbf{h}_i^T + \mathbf{Z} \quad (3)$$

where $\mathbf{Y} \in \mathbb{C}^{L \times M}$ is the matrix of received signals across M antennas over L symbols, and $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_M]$ with $\mathbf{z}_m \sim CN(0, \sigma^2 \mathbf{I})$ is the independent AWGN at the BS. Now define $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_{N(t)}]$. Let $x_i = \alpha_i \mathbf{h}_i$ and define $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_{N(t)}]^T$. Then, the training phase can be modeled as the following matrix equation:

$$\mathbf{Y} = \sqrt{\xi} \mathbf{A} \mathbf{X} + \mathbf{Z} \quad (4)$$

where the rows of the matrix \mathbf{X} follow a Gaussian distribution.

The goal for the BS in the second phase is to detect the user activities and to estimate the user channels by recovering \mathbf{X} based on the observation \mathbf{Y} . As \mathbf{X} is row sparse, i.e., many \mathbf{x}_n 's are zero, such a reconstruction problem is a compressed sensing problem. Further, as the sparsity pattern is sensed at multiple antennas, this is an MMV compressed sensing setup.

Among many powerful compressed sensing techniques, this paper adopts a low-complexity AMP algorithm to recover the row-sparse matrix \mathbf{X} . Before proceeding to evaluate its performance for user activity detection and channel estimation, we first briefly review the vector version of the AMP algorithm in the next subsection.

3 Amp Algorithm Based on Probability Learning

In the first stage, according to the received notification signal, the base station (BS) obtains the number of users. After the number of users is known in the cell, the active and inactive users are classified according to the received signal and pilot signal. Next, we start introducing a classification algorithm based on deep learning. In deep learning [7], $\mathbf{Y}(:, d)$ denotes the d th column vector of the received signal \mathbf{Y} . The training data $(\mathbf{Y}(:, d); \mathbf{X})_{d=1}^M$ composed of (feature, label) pairs are used to train the parameters of a deep neural network with the goal of accurately predicting the label \mathbf{X} of a test feature $\mathbf{Y}(:, d)$. The deep learning network accepts $\mathbf{Y}(:, d)$ and subjects it to many layers of processing, where each layer consists of a linear transformation followed by a nonlinearity (Fig. 2).

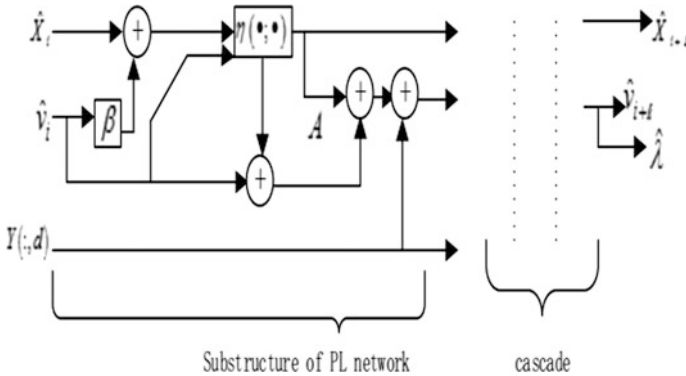


Fig. 2. System model of the massive device communication network

The neural network architecture proposed in [2] is proposed. We can use the AMP network to build the neural network. We rewrite AMP algorithm expression as

$$\hat{\mathbf{x}}_{t+1} = \eta(\mathbf{S}\hat{\mathbf{x}}_t + \mathbf{B}\mathbf{y}; \pi), \mathbf{B} \triangleq \beta\mathbf{A}^H, \mathbf{S} \triangleq \mathbf{I}_N - \mathbf{B}\mathbf{A} \quad (5)$$

AMP algorithm usually uses the EM algorithm to classify the system. However, the information obtained through EM often has vulnerabilities, and the classification accuracy is lower under the condition of low SNR, so we calculate π by cascading AMP networks, which can not only obtain accurate channel information but also improve the estimation accuracy. We introduce the design of neural network based on AMP neural network. Our PL network can learn both the vector π from the training data $(\mathbf{Y}(:, d); \mathbf{x})_{d=1}^M$

Here, π denotes the set of learnable parameters and $\mathbf{X}_T(\mathbf{Y}(:, d); \mathbf{A})$ the output of the T -layer network with input $\mathbf{Y}(:, d)$. Because the learning parameter π is a probability value, the approach was coined ‘‘probability learning network.’’

We propose to construct a neural network from AMP with tunable parameters π learned from training data. The hope is that a ‘‘Probability Learning’’ (PL) will require multilayers. The AMP typically requires iterations in each layer. Figure 3 shows one layer of the PL network. We can combine the one layer of PL network. By processing, we can get the classified information. Note that PL is built on a generalization of AMP network (5). An important question is whether this generalization preserves the independent Gaussian nature of the input error.

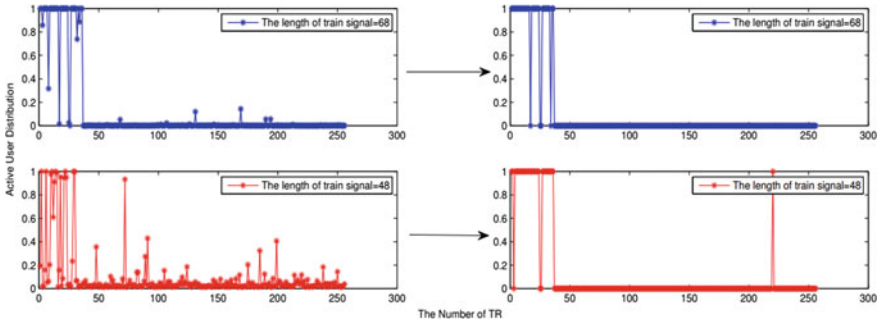


Fig. 3. Detection performance of PL network algorithm

Because the system is a sparse structure, the system element is an approximation of the Bernoulli–Gaussian distribution; in the stage of PL, we have known the location information of the elements, so it is equivalent to the processing of a Gauss variable problem. According to the characteristics of the AMP algorithm, the AMP algorithm is very effective in solving such data with Gauss characteristics. The general state evolution of the AMP algorithm as in [5] applies to any arbitrary channel estimator $\eta_{t,n}(\bullet)$. With the MMSE estimator, the state evolution can be considerably simplified.

4 Simulation Results

In this section, we provide numerical examples to verify the results of this paper. The setup is as follows. There are $100 < N(t) < 200$ devices in the cell with $0 < K < 256$ active devices at a time. Let d_n denote the distance between user n and the BS, $\forall n$. It is assumed that d_n 's are randomly distributed in the regime $[0.5, 1 \text{ km}]$. The path loss model of the coherence time of the wireless channel is 1 MHz and 1 ms, respectively, and thus in each coherence block $T = 1000$ symbols can be transmitted. The transmit power for each user at both the first and second transmission phases is $\rho^{pilot} = 23 \text{ dBm}$. The power spectral density of the AWGN at the BS is assumed to be -169 dBm/Hz . Moreover, the numerical results here are obtained by using Monte Carlo method simulation.

First, we examine the performance of user activity detection achieved by the proposed PL network detector. Given any M , Fig. 4 shows $N(t) = 256$ and $K = 36$, and the PL classified detector at the BS determines the active users from all users. What is more, we show that when the length of train signal $P = 68$ and $P = 64$, the PL detector can accurately get the position information of active users from all users. In Fig. 4, when the received signal and train signal are processed by one-level PL network, the error of detecting is very obvious. When they are processed by multilevel PL network, the error of detecting is very small. Figure 4 shows that the position of active users mainly concentrates upon the 1–36. Figure 4a shows the connection of NMSE and SNR, and this diagram shows the relationship between the PLAMP algorithm and the SNR, and in the diagram, we give the relation between the AMP algorithm and the SNR. On the way, it is easy to see that the performance of PLAMP is better than AMP, and the convergence speed and accuracy of PLAMP are higher in the same number of iterations and pilots.

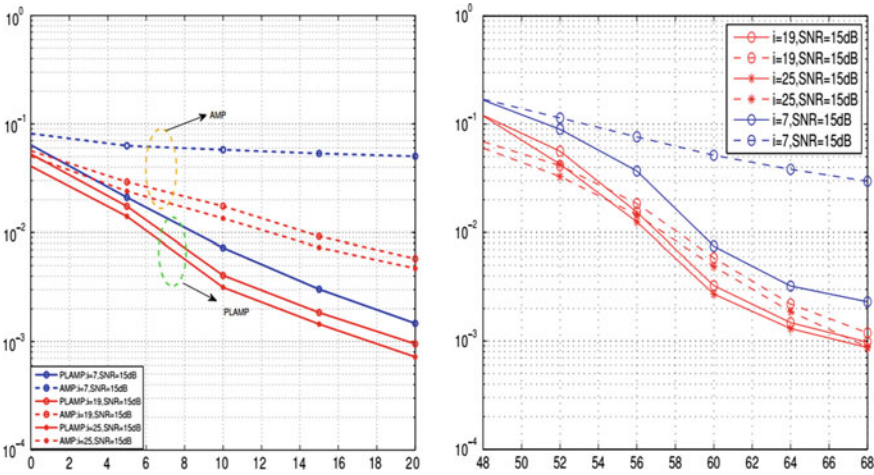


Fig. 4. The NMSE of PLAMP vs AMP with the variation of SNR in (a) and the NMSE of PLAMP vs AMP with the variation of the length of pilot in (b)

Figure 4b shows the connection of NMSE and the length of train signal. In the same number of iterations and in the case of SNR, it is easy to see that the PLAMP algorithm estimates performance is better than traditional AMP algorithm. Moreover, compared to the traditional AMP algorithm, our proposed PLAMP converges faster.

5 Numerical Results

We addressed a multiple measurement vector (MMV) model in practical scenarios. We showed that existing algorithms have poor performance and they have limited ability in practice. To solve this problem, we proposed a completely new deep learning network based on AMP network framework, which can easily get the channel matrix and the performance is better. Based on this framework, we derived new algorithms, namely, PLAMP. Extensive experiments have shown that the proposed algorithms have superior performance to many state-of-the-art algorithms. Theoretical analysis also has shown that the proposed algorithms have given the process of algorithm. In the next work, we will continue discussing the theoretical lower bounds of the algorithm.

References

1. Boccardi, F., Heath, R.W., Lozano, A., et al.: Five disruptive technology directions for 5G. *IEEE Commun. Mag.* **52**(2), 74–80 (2014)
2. Wang, P., Li, Y., Song, L., et al.: Multi-gigabit millimeter wave wireless communications for 5G: from fixed access to cellular networks. *IEEE Commun. Mag.* **53**(1), 168–178 (2015)
3. Niu, Y., Li, Y., Jin, D., et al.: A survey of millimeter wave communications (mmWave) for 5G: opportunities and challenges. *Wireless Netw.* **21**(8), 2657–2676 (2015)
4. Rusek, F., Hu, S.: Sequential channel estimation in the presence of random phase noise in NB-IoT systems. [arXiv:1706.04350](https://arxiv.org/abs/1706.04350) (2017)
5. Hu, S., Berg, A., Li, X., et al.: Improving the performance of OTDOA based positioning in NB-IoT system. [arXiv:1704.05350](https://arxiv.org/abs/1704.05350) (2017)
6. Zhang, Z., Wang, X., Zhang, Y., et al.: Grant-free rateless multiple access: a novel massive access scheme for internet of things. *IEEE Commun. Lett.* **20**(10), 2019–2022 (2016)
7. Alkhateeb, A., El Ayach, O., Leus, G., et al.: Channel estimation and hybrid precoding for millimeter wave cellular systems. *IEEE J. Sel. Top. Signal Process.* **8**(5), 831–846 (2014)
8. Liu, C., Li, M., Collings, I.B., et al.: Design and analysis of transmit beamforming for millimeter wave base station discovery. *IEEE Trans. Wireless Commun.* **16**(2), 797–811 (2017)



DEEP: Detection of Environmental Pollution Using Cooperative Neural Network

Yang Zhang^(✉)

Division of Wireless Optimization, China Mobile Communications Group Co., Ltd.,
Beijing, China
zhangyang@chinamobile.com, 121623036@qq.com

Abstract. High-accuracy detection of environmental pollution (DEEP) schemes to measure a variety of pollutants arouses great interests in industry and research communities. This paper proposes a novel DEEP approach to improve the detection precision by using machine learning theory in which an RBF network for detection is optimized by genetic algorithm. Specifically, this cooperative scheme employs more appropriate relationship in the networks, which can accelerate the convergence of the algorithm and also can enhance the precision. Simulation results demonstrate that the proposed method outperforms conventional schemes in terms of environmental pollution detection accuracy, as well as monitoring different pollutants.

Keywords: Detection of environmental pollution (DEEP) · Radial basis function network · Genetic algorithm

1 Introduction

The explosive growth of industry and society leads to terrible environmental pollution. In particular, air pollution and water pollution have reduced the quality of our lives. To be specific, these pollution problems bring great harm to human health, ecosystem, and the climate, and hence cause an important and complex problem [1]. For the purpose of minimizing adverse effects on human health and the environment caused by the air quality, both the Ambient Air Quality Directive of the Council and the European Parliament [2] and Air Quality Guidelines [3] of WHO announce limits for ambient concentrations of air pollutants.

In order to govern the environmental pollution, it is essential for us to make accurate and complementary environment pollution detection. Some scholars have proposed different methods to detect the environmental pollution. In [4], a way for environmental perception was developed based on the characteristics of the visible light communication channel. Also, the authors proposed a scheme to detect the pollution outflow from Mexico City with the aids of CALIPSO

lidar measurements [5,6] presented a strategy to monitor atmospheric pollution in Brno, Czech Republic through time series. In order to overcome the temporal instability and the lack of reproducibility of semiconductor gas sensors, a calibration and normalization procedure was well investigated [7]. However, these methods fail to detect different kinds of pollutants, and the accuracy is too low to meet our requirements in practice.

Recently, the concepts of artificial intelligence (AI) and machine learning have attracted great interests in academic and industrial fields. Paper [8] explored a new way called random forest which is based on the thought of Bagging. In [9], the author proposed an ARIMA algorithm to make accurate predictions. Furthermore, a novel framework named deep belief network (DBN), which is derived by Hinton, has improved the performance of the conventional neural network [10]. These methods are dedicated to make accurate environmental pollution.

Motivated by the aforementioned considerations, a complementary study of the environmental detection is conducted. Different from previous work, taking the advantages of the machine learning, the performance of conventional neural network is improved and it is employed to make environmental detection. To the best of our knowledge, with the aids of sensors, we first apply the machine learning theory to detect the environmental pollution. Moreover, simulation results have demonstrated our proposed scheme that can make accurate environmental pollution detection.

The rest of this paper is organized as follows. In Sect. 2, we propose a cooperative neural network model. Then, in Sect. 3, algorithms for accuracy environmental pollution detection are proposed. Numerical results for evaluating the performance of the proposed schemes are provided in Sect. 4, which is followed by conclusions in Sect. 5.

2 System Model

Currently, the high-precision environmental detection is realized with the aids of high-precision sensors that cost. Besides a large scale of detection, multiple sensors are supposed to be deployed, which would unquestionably cost more money. Therefore, in this paper, we incorporate with visible light communication (VLC) channel. To be specific, bit error rate (BER), data rate, and time delay in the communication links would change when passing through the experimental environment corrupted by the pollutant. Furthermore, those three indices vary strongly according to the change of pollutant concentration in the experimental environment.

Now that we obtain a set of data like BER, data rate, and time delay based on our environmental pollutant monitoring system, and we would concentrate on how to specifically make predication about air quality. It is of great significance to employ those three features as inputs to make a multivariable linear regression model over pollutant concentration. Multivariable linear regression is a widely used statistical method that would consider largely every data point in the data set, which has been verified as a sharpen weapon in building a mathematical model.

However, this method, to a large extent, keeps a comparatively high error when the following situations happen:

- Quite a few of noise points are recorded, which could not be well exploited by this regression algorithm.
- When the scale of training data is not large, it could not deal with new data that seem not associated with the training data.
- The output obtained from this model is of low accuracy during the simulation, which is not acceptable when it comes to boosting the performance of our system.

In order to enhance the precision of our observation system as well as the adaptive ability in multiple situations, in this paper, we employ a cooperative machine learning method, which consists of radial basis function network (RBF network) and genetic algorithm (GA). The RBF network serves as a predictor in our system, combined with GA, an optimization algorithm to search for the optimal value of weights and center vectors in RBF network.

3 Algorithm for Environment Pollution Detection

3.1 RBF Network

As is shown in Fig. 1, let us consider an RBF network with three neural layers, including an input layer, a hidden layer, and an output layer. There are two main parameters in RBF network, W for weights of the units in hidden layer contributing to the final result in the output layer, and c for center vectors that are included as units in the hidden layer. In this algorithm, we consider the training processing as a supervised learning, which means the major work of training processing is to adjust the parameters in the model. To run up this RBF network, the input data X is supposed to be represented as a vector with N -dimensions as N features of input data. In order to enhance the goodness of fit, activation functions are introduced between input layer and hidden layers in RBF network to provide nonlinear approximation which is of high accuracy. Particularly, the activation function of RBF network is radial basis function, embodying the distance between the input and the center vectors. Furthermore, the Euclid distance is employed to define the distance. Here, we introduce Gaussian function with zero mean and variance for 1 as the radial basis function. And the output layer gets the result of a linear combination of outputs of the units in the hidden layer with weights W .

The least sum squared error (SSE) is introduced to be the object function in this network. In every iteration, the network would be trained to minimize the SSE by automatically adjusting W and c of the network.

RBF network harbors strong ability in nonlinear predicting in reflecting relationships between several variables, and it is now widely applied in machine learning. To be summarized, this RBF network boasts these following strengths:

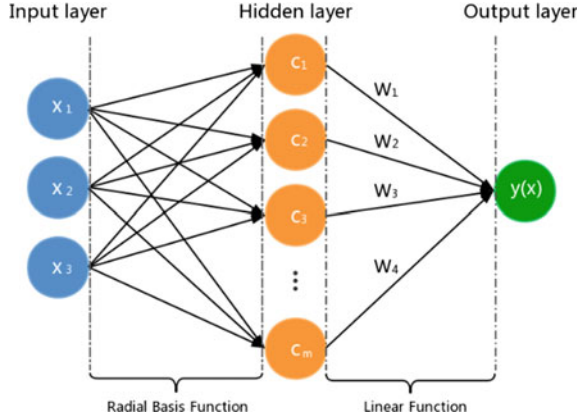


Fig. 1. The diagram of RBF network

- The activation function of hidden layer is usually set to be with localized receptive field, which means the neurons would only answer the input data that within a diminutive and specific space.
- The concept of distance between input data and center vectors brings about the feasibility to avoid random partition generated by the hyperplanes.

Basically, BER, data rate, and time delay characterize the pollutant concentration in our system. Consequently, we set those three aforementioned variables as a three-dimensional vector as input data, and pollutant concentration is regarded as the target data. With the aids of the RBF network, the high precision could be realized in principle. However, the weight and center vectors of every neuron are still hard to be optimized, for the initial step to choose the center vectors is of randomness, which would bring complicated influence to the network that could not be described concretely. Honestly, the values of W could be certain by linear method called linear least square method. Yet, nonlinear method for optimization is supposed to be cited to deal with optimization for other indices such as center vectors. Assuming the linear method is deployed, the answer we get would probably converge to local optimum. Accordingly, to estimate the value of other parameters, in our algorithm we introduce genetic algorithm (GA) to do this work.

3.2 Genetic Algorithm

GA is a widespread algorithm used to seek for optimal solution of sophisticated problems, which is promoted by the process of natural selection.

To complete this algorithm, we would propose a characteristic function to compute the sufficiency of every individual in a generation. The chromosomes are then ranked on basis of their specific sufficiency, the result of which decides the probability of every individual with their chromosomes. Set the whole number of individuals to be noted as N , the rank of chromosomes is denoted as $m(i)$, and

the probability of chromosome is assumed as $P(i)$; then the relationship between $m(i)$ and $P(i)$ could be computed as the ratio of $N - m(i)$ and N .

Then among a generation, the individuals would probably mate, and their offspring would get a genetic variation. It is obvious that the chromosome of offspring is the hybridization of that of parents. The partition of the chromosomes is conventionally randomly selected; here, in this paper, we employ grid search to seek for a relatively optimized partition for the crossover process. Note that this component exactly plays a decisive role in the value of output, and leading to the generations of combination of chromosomes would have impacted the descendants or be influenced by the ancestors. Additionally, to avoid falling into local optimum as well as enlarge the search space, a mechanism of mutation is treated as an introduction in our algorithm. Given a small probability that a piece of chromosomes would get reversal, changing the polarity is encoded. For instance, the encoded piece of chromosomes would transform from "010" to "101" if a mutation is happened. This method, to be frank, has a goodish performance when the sum of data is inadequate or the encoded value of chromosomes fail to be even enough.

If an individual harbors a high sufficiency, it would have a high probability to mate and pass its grate gene to the descendants. Accordingly, the sufficiency of the next generation would change. The result is that the optimized genes with very high sufficiency would eventually survive after hundreds of thousands of generations.

On the basis of such a process of survival of the fittest, we could implement this algorithm to find the best weight and center vectors for our RBF network. The first procedure is to encode the parameters to chromosomes.

To sum up, algorithm for training our RBF network can be listed as followed:

- (1) Because the basic RBF network could hardly converge to global optimization, we adopt GA to handle it. Noting that every parameter is a piece of chromosome (within S), the sufficiency is the sum of the Euclid distance between the input and the output. Besides, ε is imported to represent the precision of the algorithm to control the precision of sufficiency. Therefore, we are able to control the precision of our environmental pollutant monitoring system by making adjustment about the precision of RBF network.
- (2) Calculate the length of chromosome. Set the number of generations and then encode every individual binarily.
- (3) Implement crossover and mutation among the encoded sequences.
- (4) Compute the sufficiency of every individual respectively. According to the sufficiency, make selections among the individuals. In other words, formulate $P(i)$ for every individual and choose mating individuals by Roulette.
- (5) Repeat (3)–(4) until the number of iterations meets the set number. Then acquire the weights and center vectors of the RBF network in terms of the optimum output from the genetic algorithm, and set the number of neurons in the hidden layer.
- (6) Most importantly, set the precision learned from the algorithm to be ζ . If $\zeta > \varepsilon$, continue to learn and get a precision ζ again until $\zeta < \varepsilon$.

4 Numerical Result and Analysis

In order to obtain good performance, our proposed cooperated RBF network must be well trained. In this paper, we designed a training experiment to train the network as well as a test experiment to evaluate its capability. Without loss of generality, our experiments are repeated for 25 times in the same conditions and their average values are formulated in our analysis.

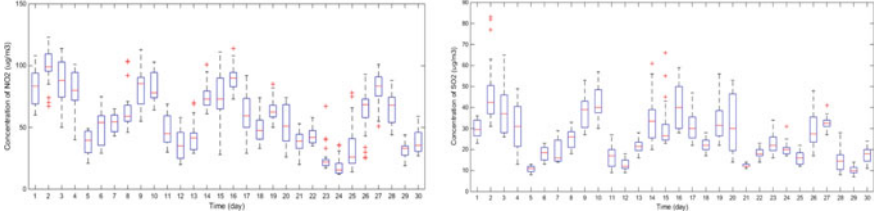


Fig. 2. The box plot of daily concentration in Nanjing, China

Here are the procedures of training process.

- (A) Set the current pollutant concentration as h in experiment environment, using NO_2 as experimental pollutant, which must be in $[0, 600]$ (ug/m^3).
- (B) In every turn, the input module sends a sequence of bit data for every value of h , which would be increased with the same pace at the beginning of a single turn.
- (C) Send input data like, BER, data rate, and time delay, combined with pollutant concentration as the output value to the RBF network and repeat the training procedure (1)–(7) in part in result of well-trained networks.
- (D) Change NO_2 to be other pollutants like SO_2 , O_3 , then we are enabled to get specific network for those pollutants.

In our experiment, we recorded the input data by stepping up the pollutant concentration of NO_2 and SO_2 for $1\text{ ug}/m^3$ per minute from 1 to $600\text{ ug}/m^3$. The collecting process spanned 600 minutes, and we recorded the data every minute, which means we got a set of 600 pairs of input and output data and cause the concentration that could hardly over $600\text{ ug}/m^3$. Then, we regard the three features as input data, according to concentration as target data into our RBF network to train it.

After that, testing experiment is introduced to validate our model. To make it more scientific, we gained the concentration of NO_2 and SO_2 from the website of “<http://pm25.in/nanjing>”, hourly from January 1, 2016 and January 30, 2016 in Nanjing. This website provides six main pollutants in criteria like NO_2 , SO_2 , etc. Nanjing is a huge city in China with a population of 827 million, where industries of all fields are well developed. Hence, the impact of air quality would play an important role in citizens’ daily lives as well as their well-being. There exist several stations in Nanjing, which are distributed decentralized in the city. The statistical representation of the data is shown in Fig. 2. In the box-plot figure, the

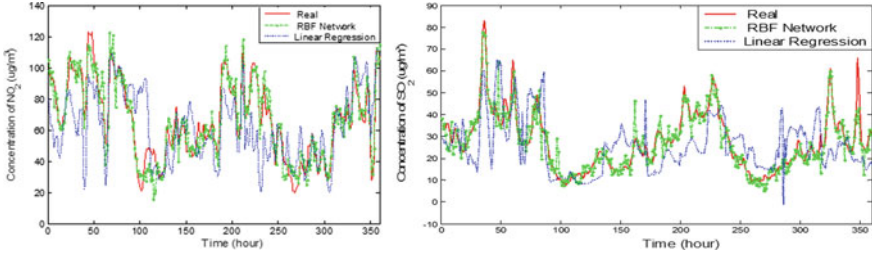


Fig. 3. Comparison of real concentration and predicted one

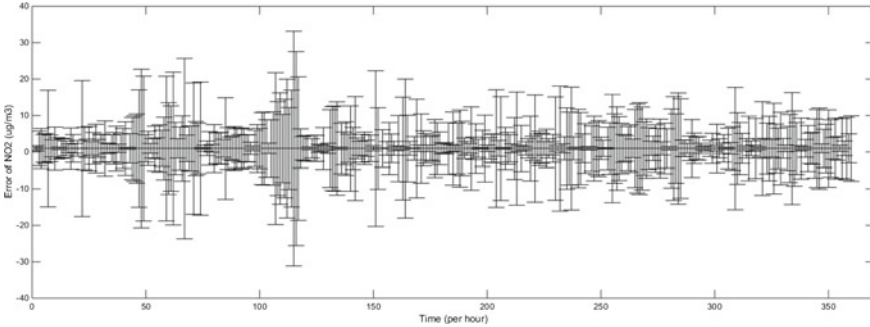


Fig. 4. Error of RBF network on NO_2

daily concentration of stations is comparatively and centrally distributed. With the aim of getting a universal result, we employed the global concentration data in Nanjing city as our experimental data. In terms of that, the concentrations of NO_2 and SO_2 were simulated in our laboratory, and meanwhile collected the BER, data rate, and time delay as input data. Simultaneously, we both employed our RBF network and multivariable linear regression to make a comparison. The result is clearly shown in Figs. 3 and 4.

In Fig. 3, take the hourly concentration of NO_2 generated from RBF network and linear regression as example. As for the former one, in a month, the concentration underwent several considerable fluctuations, the top of which reached around 120 ug/m^3 and the lowest of which reached nearly 20 ug/m^3 . It could be concluded that the concentration of NO_2 in realistic environment proves to be tough to make predication. Moreover, it could be found that the output of the RBF network was really acceptable for both the tendency of variation and the value of every hour pollutant concentration. Then, when it comes to linear regression, the performance got worse. It could not be tough to understand that the line regression is comparatively weak in predicting if the input is not so close to the training one. The result is similar when the pollutant is SO_2 . Needless to say, it would be better to use RBF network to gain higher precision in pollutant concentration detection.

Specifically, Fig. 4 indicates the overall error of concentration of RBF network when the pollutant is NO_2 . It could be observed from Fig. 4 that in most cases

the error would be lower than 10 ug/m^3 , manifesting the correctness of estimates of concentration. Yet, the biggest error reached to 33.87 ug/m^3 , which it cannot meet the requirement. To have some insight into what exactly contributed to the error, the conclusion is that when the pollutant concentration meets with a huge change, the inertia of RBF network drives the predicted concentration to deviate.

5 Conclusion

This paper proposed a high-precision method for pollution concentration detection. To validate the effectiveness of our detection system, simulation results of the pollutant concentration of NO_2 in our experimental environment with the realistic data in Nanjing are presented. Having analyzed the result of experiment, it could be concluded that our detection system is highly capable to work out agreeable results. It is also worthy to be indicated that the error of our system negatively influences the detection for the whole duration of experiment, which affords us an opportunity to explore our algorithm in the future.

References

1. Holeovský, J., Šampulová, M., Michálek, J.: Semiparametric outlier detection in nonstationary times series: case study for atmospheric pollution in Brno, Czech Republic. *Atmos. Pollut. Res.* (2017)
2. Parliament, E., Union, C.O.: Directive 2008/50/ec of the European Parliament and of the Council of 21 may 2008 on ambient air quality and cleaner air for Europe. *Off. J. Eur. Communities*, 1–43 (2008)
3. Saldiva, P.H.N., Künzli, N., Lippmann, M.: Air quality guidelines: global update 2005. Particulate matter, ozone, nitrogen dioxide and sulfur dioxide. *Indian J. Med. Res.*, 492–493 (2006)
4. Huang, H., Wang, J., Zhou, X., et al.: Cooperative VLC systems for data transmission and environment perception. In: 2017 9th IEEE International Conference on Communication Software and Networks (ICCSN), pp. 624–629. Guangzhou (2017)
5. Kar, J., et al.: Detection of pollution outflow from Mexico City using CALIPSO lidar measurements. *Remote. Sens. Environ.* **169**, 205–211 (2015)
6. Hrdliková, Z., Michálek, J., Kolář, M., Veselý, V.: Identification of factors affecting air pollution by dust aerosol PM in Brno City, Czech Republic. *Atmos. Environ.*, 8661–8673 (2008)
7. Nichani, S.: System or method for identifying contents of a semi-opaque envelope. US patent, US5970166 (1999)
8. Buschjäger, S., Morik, K.: Decision tree and random forest implementations for fast filtering of sensor data. *IEEE Trans. Circuits Syst. I: Regul. Pap.* **65**(1), 209–222 (2018)
9. Chen, B.-S., Lee, B.-K., Peng, S.-C.: Maximum likelihood parameter estimation of F-ARIMA processes using the genetic algorithm in the frequency domain. *IEEE Trans. Signal Process.* **50**(9), 2208–2220 (2002)
10. Hinton, G.E., Osindero, S., Teh, Y.W.: A fast learning algorithm for deep belief nets. *Neural Comput.* **18**(7), 1527–1554 (2006)



Imbalanced Data Classification Method Based on Ensemble Learning

Yu Xiang^(✉) and Yongping Xie

School of Information and Communication Engineering, Dalian University
of Technology, Dalian, Liaoning, China
dutxiangyu@126.com

Abstract. Imbalanced data classification is one of the problems that emerged when classifier learning algorithms used in the worlds of business and industry. This paper proposes the methodology to improve the performance of imbalanced data classification. We balance data sets by using synthetic minority oversampling technique (SMOTE); noise generated by new data sets is eliminated by Tomek links (T-Links), support vector machine (SVM), k-nearest neighbor (KNN), and logistic regression (LR) which are selected as the base classifiers to improve classification by using stacked generalization, and the final result is generated by weighted voting. In the experiments, six UCI datasets are tested, and the experimental results show that the method is highly representative and can effectively improve the classification ability.

Keywords: Imbalanced data · SMOTE–Tomek links · Ensemble learning · Stacked generalization · Weighted voting

1 Introduction

In practical applications, imbalanced data exists widely such as credit card fraud [1, 2] and software defect detection [3]. Classification methods are mostly based on balanced data sets. When the methods are used to solve the imbalanced classification problems, their classification accuracy often declines [4]. There are mainly two ways to solve imbalanced classification problem [5]. For imbalanced data sets, we can reduce the degree of sample imbalance by changing the proportion of training set samples. In classification tasks, we can improve classification algorithm to adapt to imbalanced classification. SMOTE algorithm [6] is the commonly used oversampling method. However, it does not take into account the characteristics of the neighboring samples, which easily lead to noise samples and affect classification boundaries. Evaluation criteria suitable for imbalanced data classification are harmonic mean F-measure and geometric mean G-mean [7] based on the confusion matrix.

This paper uses SMOTE + T-Links method [8] to resample the original dataset. In the process of constructing classification model, we use one kind of ensemble learning method stacked generalization to improve the classification performance, the method first-layer algorithm combines SVM, KNN, and LR as sub-classifier, and second-layer algorithm uses weighted voting method. This procedure further enhanced the efficiency and performance of classification.

2 Background

SMOTE produces the same number of artificial samples for each minority sample to achieve data sets balance; however, the new generated samples increase the overlap of different classes. T-Links as a data cleaning method reduces classes overlap.

Ensemble learning methods use multiple learning algorithms to improve predictive performance and generalization ability. We use staked generalization method to combine different learning algorithms, in order to achieve higher prediction accuracy.

SVM analyze data is used for classification and regression analysis; its basic model is the largest linear classifier in the feature space. The model can minimize the empirical error and maximize the geometric margins.

KNN is a supervised learning algorithm, which classifies new instance based on the majority of the k-nearest neighbor category (distance in the feature space).

LR is used to estimate the probability of a binary response based on one or more predictor (or independent) variables (features).

Weighted voting is a method in which not all voters have the same amount of influence over the outcome. In this paper, the weighted voting method determines the results of each sub-classifier.

3 Experimental Design

3.1 Imbalance Data Preprocessing

In this experiment, the first step is to build balanced data sets through the combination of oversampling and undersampling two techniques; the principle is as follows:

SMOTE method adds minority samples to achieve data sets balance. First, a sample E in minority class is selected and k-nearest neighbors are found. Then, a sample E' is randomly selected from the k-nearest neighbors, and finally, a new minority sample is constructed according to the following formula:

$$E_{new} = E + rand(0, 1) \times (E' - E) \quad (1)$$

$rand(0, 1)$ means a value selects from $(0, 1)$.

SMOTE algorithm progress is as follows:

1. Each sample x of minority class calculates the distance to all the samples of minority class except itself using Euclidean distance, and then gets its k-nearest neighbors.
2. Set ratio N according to the sample imbalance ratio; each minority sample E randomly selects N samples from k-nearest neighbor, named as E_i ($i = 1, 2, 3, \dots, N$).
3. E and E_i ($i = 1, 2, 3, \dots, N$) repeatedly build a new sample with the formula.

SMOTE can reduce the degree of data over-fitting; however, it will introduce new noise and affect the classification boundary. We reduce the kind of noise by using T-Links method.

T-Links is an undersampling method, given two samples E_i and E_j belong to different classes. $d(E_i, E_j)$ is the distance between two samples, when there is no such sample E_l satisfies

$$d(E_i, E_l) < d(E_i, E_j) \text{ or } d(E_j, E_l) < d(E_i, E_j) \quad (2)$$

Then, sample pair (E_i, E_j) is considered noise named Tomek lines. SMOTE method has well-balanced class distribution; all Tomek lines are stripped out by T-links.

3.2 Ensemble Learning

After obtaining balanced data, we train classification model by using stacked generalization method [9]. The sub-classifiers choose LR, SVM, and KNN as the first-layer algorithm and weighted voting method [10] as the second-layer algorithm; stacked generalization method process is described as follows:

1. Input: initial data sets $D = \{x_i, y_i\}_{i=1}^p$
2. Data sets are divided into training sets and test sets $D \rightarrow D_{tr}, D_{ts}$
3. The first-layer algorithm L_1, L_2, L_3
4. The second-layer algorithm L (weighted voting)
5. Step 1: basic classifier training
6. For $t = 1$ to 3 do
7. Based on D_{tr} and L_t learning prediction function h_t
8. End for
9. Step 2: based on test sets D_{ts} to make prediction
10. For $i = 1$ to n do
11. $x'_i = \{h_1(x_i), h_2(x_i), h_3(x_i)\}$ generate new data sets $D_h = \{x'_i, y_i\}$
12. End for
13. Step 3: learn to get a meta-classifier
14. $h' = L(D_h)$
15. Output: $H(x) = h'(h_1(x), h_2(x), h_3(x))$.

The output of each classifier is probability. The final classification result is obtained by using the weighted voting, and the formula is as follows:

$$result = \max \left(\sum_{n=1}^m w_n \times P_{\min(n)}, \sum_{n=1}^m w_n \times P_{\max(n)} \right) \quad (3)$$

w_n is the classifier weight, $P_{(\min)n}$ is minority class probability for classifier n , $P_{(\max)n}$ is majority class probability for classifier n , and result is output. Figure 1 illustrates the framework of how the method works.

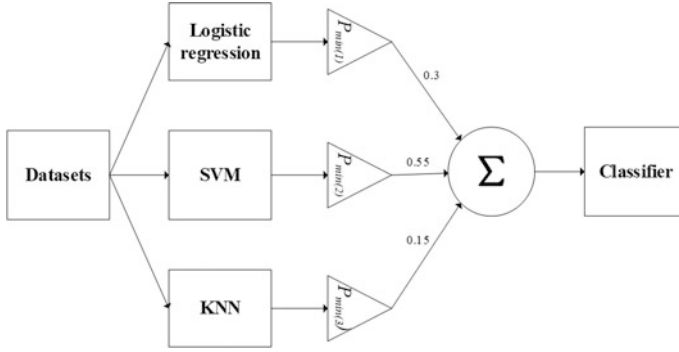


Fig. 1. Framework of classification

4 Experimental Result

4.1 Evaluation Index Selection

Classical classifiers use accuracy as performance criteria, but in imbalanced data sets this index may be biased toward majority class. Researchers proposed F-measure, G-mean based on confusion matrix (as shown in Table 1). TP (true positive) and FN (false negative) indicate correct classification of minority class and incorrect classification of majority class, TN (true negative) and FP (false positive) indicate correct classification of majority class and incorrect classification of minority class.

Table 1. Confusion matrix for two-class problem

Class	Predicted positive	Predicted negative
Actual positive	TP	FN
Actual negative	FP	TN

F-measure is used to evaluate minority class performance:

$$F - measure = \frac{(1 + \beta^2) \times Recall \times precision}{\beta^2 \times Recall + Precision} \quad (4)$$

Recall and precision are defined as follows:

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

β is used to adjust weight between recall and precision, in this paper, β is 1.

G-mean is used to measure classification performance:

$$G - mean = \sqrt{TPR \times FPR} \quad (7)$$

TPR and FPR are defined as follows:

$$TPR = Precision = TP/(TP + FP) \quad (8)$$

$$FPR = TN/(TN + FN) \quad (9)$$

4.2 Information of Data Sets

Six groups UCI data sets are chosen to test method, which belong to different categories and application backgrounds. When data sets have two or more categories, we take the least number of class sets as minority class and merge other classes as majority class. Data sets are shown in Table 2, and ratio represents imbalanced degree, which is the ratio of the majority to the minority. Comparing the experimental with results of RBBag, OvBag, uNBBag0.5, and KAcBag algorithms coming from literature [11], 10-fold cross-validation method is used to verify classification results, and average result is the final result. Python language is used to implement the method.

Table 2. Information of data sets

Data sets	Sample number	Attributes	Minority class	Ratio
Ionosphere	351	34	b	1.79
Balance scale	625	4	B	11.76
Pima	768	8	1	1.87
Vehicle	846	18	van	3.25
New thyroid	215	5	2	5.14
Abalone	4177	8	0-4:16-29	11.47

4.3 Experimental Results

As shown in Tables 3 and 4, two indexes are better than any other algorithms compared to ionosphere, balance scale, and vehicle data sets. However, indexes about pima are lower than RBBag, NBBag0.5, and KAcBag, in which there are undersampling algorithms. Because imbalanced degree is not high in data sets, undersampling classification is more accurate.

Table 3. Minority class F-measure comparison of different sampling algorithms

Data sets	RBBag	OvBag	uNBBag0.5	KAcBag	my
Ionosphere	89.00 (2)	88.84 (3)	88.79 (4)	87.92 (5)	92.56 (1)
Balance scale	15.80 (3)	0.51 (5)	15.86 (3)	19.47 (2)	47.19 (1)
Pima	68.54 (2)	66.24 (5)	67.93 (4)	68.70 (1)	68.08 (3)
Vehicle	89.44 (4)	90.38 (2)	89.79 (3)	87.92 (5)	95.39 (1)
New thyroid	91.70 (5)	92.03 (3)	91.82 (4)	93.49 (2)	97.78 (1)
Abalone	39.34 (3)	42.34 (2)	38.37 (4)	6.21 (5)	49.82 (1)

Table 4. Minority class G-mean comparison of different sampling algorithms

Data sets	RBBag	OvBag	uNBBag0.5	KAcBag	my
Ionosphere	90.67 (3)	90.47 (4)	90.71 (2)	90.44 (5)	95.17 (1)
Balance scale	54.23 (3)	1.40 (5)	32.76 (4)	54.91 (2)	62.11 (1)
Pima	75.64 (1)	73.54 (5)	74.80 (3)	75.12 (2)	73.69 (4)
Vehicle	95.44 (3)	94.61 (4)	95.53 (2)	94.46 (5)	96.39 (1)
New thyroid	96.58 (3)	95.36 (5)	96.49 (4)	98.59 (1)	97.89 (2)
Abalone	79.32 (3)	61.95 (5)	79.59 (2)	79.71 (1)	66.31 (4)

5 Conclusion

This paper presents an imbalanced data classification method based on ensemble learning. First, we build the balanced data based on SMOTE + T-Links method. Then we use stacked generalization model, in which first-layer algorithm combines with three basic classification models (SVM, KNN, and LR), and second-layer algorithm uses weighted voting method to obtain the final result. Finally, experiments on UCI data sets show that the proposed method can effectively utilize the distribution of data and improve the classification performance.

References

1. Wei, W., Li, J., Cao, L., Ou, Y., Chen, J.: Effective detection of sophisticated online banking fraud on extremely imbalanced data. *World Wide Web-Internet Web Inf. Syst.* **16**(4), 449–475 (2013)
2. Wang, G., Hao, J., Ma, J., Jiang, H.: A comparative assessment of ensemble learning for credit scoring. *Expert Syst. Appl.* **38**(1), 223–230 (2011)
3. Yi, P.E.N.G., Gang, K.O.U., Guoxun, W.A.N.G., Wenshuai, W.U., Yong, S.H.I.: Ensemble of software defect predictors: an ahp-based evaluation method. *Int. J. Inf. Technol. Decisi. Mak.* **10**(01), 187–206 (2011)
4. Chawla, N.V.: *Data Mining for Imbalanced Datasets: An Overview*. Data Mining and Knowledge Discovery Handbook. Springer, US (2005)
5. He, H., Garcia, E.A.: Learning from imbalanced data. *IEEE Trans. Knowl. Data Eng.* **21**(9), 1263–1284 (2009)

6. Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: Smote: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**(1), 321–357 (2002)
7. Zhi-Fei, Y.E., Wen, Y.M., Bao-Liang, L.U.: A survey of imbalanced pattern classification problems. *Caai Trans. Intell. Syst.* (2009)
8. Batista, G.E.A.P.A., Prati, R.C., Monard, M.C.: A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explor. Newsl.* **6**(1), 20–29 (2004)
9. Graczyk, M., Lasota, T., Trawiński, B., Trawiński, K.: Comparison of bagging, boosting and stacking ensembles applied to real estate appraisal, vol. 5991, pp. 340–350 (2010)
10. Rojarath, A., Songpan, W., Pong-Inwong, C.: Improved ensemble learning for classification techniques based on majority voting. In: *IEEE International Conference on Software Engineering and Service Science*, pp. 107–110. IEEE (2017)
11. Bingyan, Xiong, Guoying, Wang, Weibin, Deng: Under-sampling method based on sample weight for imbalance data. *J. Comput. Res. Dev.* **53**(11), 2613–2622 (2016)



Bayesian Method-Based Learning Automata for Two-Player Stochastic Games with Incomplete Information

Hua Ding¹(✉), Chong Di², and Li Shenghong²

¹ School of Biomedical Engineering, Dalian University of Technology, No. 2 Linggong Road, Ganjingzi District, Dalian 116023, Liaoning, China

dinghua@mail.dlut.edu.cn

² School of Cyber Security, Shanghai Jiao Tong University, Shanghai 200240, China
dichong95@sjtu.edu.cn, shli@sjtu.edu.cn

Abstract. In the field of artificial intelligence, learning automaton (LA) is a self-adaptive decision-maker which plays an important role in reinforcement learning (RL). Games of learning automata are stochastic games with incomplete information that have received frequent usage. Traditional learning automata schemes using in games are parameter-based schemes which exist a tunable parameter (stepsize) changing with different environments. In this paper, we proposed Bayesian method-based parameter-free learning automata (BPFLA) for two-player stochastic games with incomplete information. The parameter-free property indicates that a set of parameters in the scheme can be universally applicable for all configurations of games. Besides, simulation results demonstrate that BPFLA has much faster convergence rate than traditional schemes using games of learning automata with equal or higher accuracy.

Keywords: Games of learning automata · Learning automata · Reinforcement learning · Bayesian inference

1 Introduction

As a simple self-adaptive decision-maker, learning automaton (LA) plays an important role in reinforcement learning (RL) [1]. Through interacting with the random environment, LA can explore the optimal action or maximize the reward using pure or mixed strategies [2]. Games with complete information indicate that all players are perfectly informed of all the other players and their payoffs for all possible action profiles [3], while games with complete information refer that players are uncertain of the payoffs to other players. Players in games of learning automata are independent and unaware of other players. Thus, they are games with incomplete information and have received frequent usage [4].

© Springer Nature Singapore Pte Ltd. 2020

Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 517,

https://doi.org/10.1007/978-981-13-6508-9_4

Zero-sum games and games with identical payoff are two major kinds of games of learning automata which were first proposed in [5]. There are a tremendous amount of study targeting on both games. Fixed-structure automata [2], variable structure automata (L_{RI}) [6], and estimator-based automata (DPL) [7] are all used to explore the best strategies for each player in games. However, all schemes are parameter-based schemes. That is, there is a tunable parameter (stepsize) in schemes and to achieve best performance, the stepsize is tuned by interacting with environment which is unrealistic in practical applications [8]. Besides, the convergence rate of the schemes is very slow and cannot meet the requirements in practice.

In this paper, we proposed Bayesian method-based parameter-free learning automata (BPFLA) for two-player stochastic games with incomplete information. The parameter-free property indicates that a set of parameters in the scheme can be universally applicable for all configurations of games, rather than no configurable parameters are involved in the schemes [9]. The scheme achieves much higher convergence rate and makes parameter-free possibly using Bayesian inference methods. Our work presents a set of contributions and we summarize in the following:

1. We proposed Bayesian method-based parameter-free learning automata for two-player identical payoff as well as zero-sum games. The advantages of BPFLA embody in two aspects: (1) The scheme can converge with much higher rate than all the other schemes. (2) The parameter-free property ensures that the applications of BPFLA in learning automata games do not need manual parameter tuning.
2. To the best of our knowledge, this is the first time that Bayesian method-based LA has been applied to the games of learning automata.
3. A comprehensive comparison with both variable structure automata (L_{RI}) and estimator-based automata (DPL) is given to establish the superiority of proposed BPFLA scheme.

This paper is organized as follows. In Sect. 2, we introduce the common structure of LA and describe the semantics of two-player identical payoff games and zero-sum games. The proposed BPFLA scheme is discussed detailedly in Sect. 3. Experimental results are shown in Sect. 4 to indicate the advantages of BPFLA over the previous schemes in various game situations. Finally, Sect. 5 concludes this paper.

2 La and Game Setup

2.1 A Learning Automaton and an Environment

A learning automaton can be represented by a quintuple $\langle \Phi, \mathbb{A}, \mathbb{B}, \mathbb{F}, \mathbb{H} \rangle$, while the environment \mathbb{E} is represented by a triple $\langle \mathbb{A}, \mathbb{B}, \mathbb{C} \rangle$.

- Φ is a set of internal states.

- \mathbb{A} is a set of actions. That is, the outputs of the automaton or the inputs of the environment.
- \mathbb{B} is a set of responses. That is, the inputs of the automaton or the outputs of the environment.
- \mathbb{C} is a set of reward probabilities of actions in \mathbb{A} in environment \mathbb{E} .
- \mathbb{F} is a transition function that maps current state and input into next state.
- \mathbb{H} is an output function that maps current state and input into current output.

At each iteration t , the automaton selects an action a_t according to the output function \mathbb{H} to interact with the environment \mathbb{E} . Then, the environment \mathbb{E} responses a feedback \mathfrak{F}_t according to the reward probability c_t of action a_t to the automaton where $\mathfrak{F}_t = 1$ represented a reward and $\mathfrak{F} = 0$ represents a penalty.

2.2 Game Setup

In this subsection, the statement of the problem as well as the setup of two-player stochastic games with incomplete information will be introduced.

Identical Payoff Game In a two-player identical payoff game, both players (automata) participating in the game get the same payoff (feedback) at each iteration of the game [2]. Assume each player P (automaton) has r possible strategies (actions). At each iteration in the game, each player chooses a strategy, while they have no knowledge about another player and its choice. Then both players get a common payoff from the environment according to their strategies. The game is stationary stochastic and the payoff \mathfrak{F} is a random variable taking values 0 or 1, with standing for a penalty or a reward.

Define a game matrix $\mathbb{D} = [d_{i_1 i_2}]$ of dimension $r_1 \times r_2$, where

$$d_{i_1 i_2} = E[\mathfrak{F} | P_1 \text{ choose action } a_{i_1}, P_2 \text{ choose action } a_{i_2}] \quad (1)$$

Suppose that there is a set of strategies m_p^e for each player P_p ; it satisfies that

$$d_{m_1^e m_2^e} = \max(\mathbb{D}) \quad (2)$$

Then, the strategy m_p^e is called as the Nash equilibrium strategy for player P_p . The objective of identical payoff game is to make each automaton converge to the Nash equilibrium strategy.

Zero-Sum Game Similarly, define a two-player zero-sum game matrix $\mathbb{D} = [d_{i_1 i_2}]$ of dimension $r_1 \times r_2$, where

$$d_{i_1 i_2}^1 = E[\mathfrak{F} | P_1 \text{ choose action } a_{i_1}, P_2 \text{ choose action } a_{i_2}] \quad (3)$$

$$d_{i_1 i_2}^2 = 1 - d_{i_1 i_2}^1 \quad (4)$$

Considering a two-player zero-sum game, the element $d_{i_m i_n}^1$ satisfies that

$$d_{i_k i_n}^1 < d_{i_m i_n}^1 < d_{i_m i_l}, \forall k \neq m, l \neq n \quad (5)$$

Then, $d_{i_m i_n}^1$ is called as the saddle point of \mathbb{D} . That is, i_m is the Nash equilibrium strategy for player P_1 while i_n is the Nash equilibrium strategy for player P_2 . The purpose of zero-sum game is to learn the equilibrium strategies for both automata.

3 BLA Algorithms

Consider a game with P-model environment $\mathbb{E} \sim \langle \mathbb{A}, \mathbb{B}, \mathbb{C} \rangle$ that we have no prior knowledge about the random environment. In Bayesian statistics, the posterior distribution of each action's reward probability is a beta-distribution, denoted as $E = \{e_1, e_2, \dots, e_r\}$, where $e_i = \text{Beta}(\alpha_i, \beta_i)$, the parameters α_i and β_i record the rewarded times and penalized times of action a_i . The probability density function (PDF) and cumulative distribution function (CDF) of the beta-distribution are as follows:

$$f(x; \alpha_i, \beta_i) = \frac{1}{B(\alpha_i, \beta_i)} x^{\alpha_i-1} (1-x)^{\beta_i-1} \quad (6)$$

$$F(x; \alpha_i, \beta_i) = \int_0^x f(x; \alpha_i, \beta_i) dx \quad (7)$$

where $B(\alpha_i, \beta_i)$ is the beta function, and $B(x, \alpha_i, \beta_i)$ is the incomplete beta function. Meanwhile, the Maximum Likelihood Estimates (MLE) of all possible actions are denoted as $\bar{E} = \{\bar{e}_1, \bar{e}_2, \dots, \bar{e}_r\}$, where $\bar{e}_i = \frac{Z_i}{S_i}$, Z_i is the number of times that action a_i has been rewarded and S_i is the number of times that action a_i has been selected. That is,

$$\alpha_i = Z_i = S_i \cdot \bar{e}_i, \beta_i = S_i - Z_i = S_i \cdot (1 - \bar{e}_i) \quad (8)$$

$$\bar{e}_i = \frac{Z_i}{S_i} = \frac{\alpha_i}{\alpha_i + \beta_i} \quad (9)$$

Thus, we have

$$e_i = \text{Beta}(\alpha_i, \beta_i) = \text{Beta}(Z_i, S_i - Z_i) = \text{Beta}(S_i \cdot \bar{e}_i, S_i \cdot (1 - \bar{e}_i)) \quad (10)$$

According to the law of large numbers, we have

$$\bar{e}_i = \mu_i = \frac{Z_i}{S_i} \rightarrow c_i, \quad \text{as } S_i \rightarrow \infty \quad (11)$$

In line with the mechanism of beta-distribution, the Bayesian estimate of action a_i is also gradually become reliable as the number of times S_i that action a_i has been selected grows. According to [10], we have

$$\lim_{S_i \rightarrow \infty} \int_{|x-c_i| \leq \epsilon} f(x; \alpha_i, \beta_i) dx = 1, \quad \forall \epsilon \in (0, 1) \quad (12)$$

Then, for each player P_p , at each iteration in a game, assume a_{p_m} and a_{p_n} are the actions with the highest and submaximal maximum likelihood estimate $\bar{e}_{p_m} = \max\{\bar{E}_p\}$, $\bar{e}_{p_n} = \text{submax}\{\bar{E}_p\}$. The definition of probability vector Pr_p is

$$Pr_{p_m} = 1 - \int_0^{\bar{e}_{p_n}} f(x; \alpha_{p_m}, \beta_{p_m}) dx = 1 - F(\bar{e}_{p_n}; \alpha_{p_m}, \beta_{p_m}) \quad (13)$$

$$Pr_{p_i} = \frac{1 - Pr_{p_m}}{r_p - 1}, \forall a_{p_i} \in \mathbb{A}_i \quad (14)$$

The automaton converges whenever both $Pr_{p_m} > \tau$, where τ is a predefined threshold. The processes of BPFLA applied in identical payoff and zero-sum games are summarized in Algorithms 1 and 2.

Algorithm 1 BPFLA in Identical Payoff Game

Require: $\tau = 0.999$: a convergence threshold

```

1: Initial For each player  $P_p$ ,  $\alpha_{p_i} = 2, \beta_{p_i} = 1$  for  $i = 1, 2, 3, \dots, r_p$ 
2: repeat
3:   for  $i = 1$  to  $r_1$  do
4:     for  $j = 1$  to  $r_2$  do
5:       Player  $P_1$  chooses action  $a_{1_i}$  and player  $P_2$  chooses action  $a_{2_j}$ .
6:       Interact with the random environment.
7:       Receive a feedback from environment and update the parameters:
           
$$\begin{cases} \alpha_{1_i} = \alpha_{1_i} + 1, \alpha_{2_j} = \alpha_{2_j} + 1 & \text{if feedback} = 1 \text{ (reward)} \\ \beta_{1_i} = \beta_{1_i} + 1, \beta_{2_j} = \beta_{2_j} + 1 & \text{if feedback} = 0 \text{ (penalty)} \end{cases}$$

8:     end for
9:   end for
10:  Update the probability vector  $Pr_p$  by (13) and (14).
11: until  $Pr_m \geq \tau$ 
    
```

4 Simulation Results

To verify the effectiveness of proposed BPFLA scheme, simulations among BPFLA, L_{RI} and DPL in different environments are provided in this section.

The definitions of the metrics used in this section are described as follows:

- The *accuracy* is defined as the ratio between the number of correct convergence and the number of all experiments.
- The *iteration* is the average number of required iterations for a correct convergence.

Since L_{RI} and DPL are parameter-based schemes, the “best” stepsize parameter α in each game matrix is determined as the biggest value with the accuracy higher than 99.5% which corresponds to the fastest convergence rate.

4.1 Two-Player Identical Payoff Game

The simulations of identical payoff games were generated in the following two 2×2 game matrices [2, 7].

$$\mathbb{D}_1 = \begin{bmatrix} 0.4 & 0.9 \\ 0.2 & 0.6 \end{bmatrix} \quad \mathbb{D}_2 = \begin{bmatrix} 0.8 & 0.6 \\ 0.3 & 0.1 \end{bmatrix}$$

It is clear that game matrix \mathbb{D}_1 has the mode (1, 2) and \mathbb{D}_2 has the mode (1, 1). Figures 1 and 2 show the changes of probability vectors of each action in typical simulations in game matrices \mathbb{D}_1 and \mathbb{D}_2 . The accuracy and iteration of all schemes in game matrix \mathbb{D}_1 and \mathbb{D}_2 are summarized in Table 1.

Table 1 indicates that BPFLA converges much faster than L_{RI} and DPL with higher accuracy in both game matrices. Compared with L_{RI} , BPFLA dramatically improves the convergence rate by 63.2% in \mathbb{D}_1 and 63.3% in \mathbb{D}_2 , while outperforms DPL by 69.5% in \mathbb{D}_1 and 73.3% in \mathbb{D}_2 .

Algorithm 2 BPFLA in Zero-sum Game

Require: $\tau = 0.999$: a convergence threshold

1: **Intial** For each player P_p , $\alpha_{p_i} = 2, \beta_{p_i} = 1$ for $i = 1, 2, 3, \dots, r_p$

2: **repeat**

3: **for** $i = 1$ to r_1 **do**

4: **for** $j = 1$ to r_2 **do**

5: Player P_1 chooses action a_{1_i} and player P_2 chooses action a_{2_j} .

6: Interact with the random environment.

7: Receive a feedback from environment and update the parameters:

$$\begin{cases} \alpha_{1_i} = \alpha_{1_i} + 1, \beta_{2_j} = \beta_{2_j} + 1 & \text{if feedback} = 1 \\ \beta_{1_i} = \beta_{1_i} + 1, \alpha_{2_j} = \alpha_{2_j} + 1 & \text{if feedback} = 0 \end{cases}$$

8: **end for**

9: **end for**

10: Update the probability vector Pr_p by (13) and (14).

11: **until** $Pr_m \geq \tau$

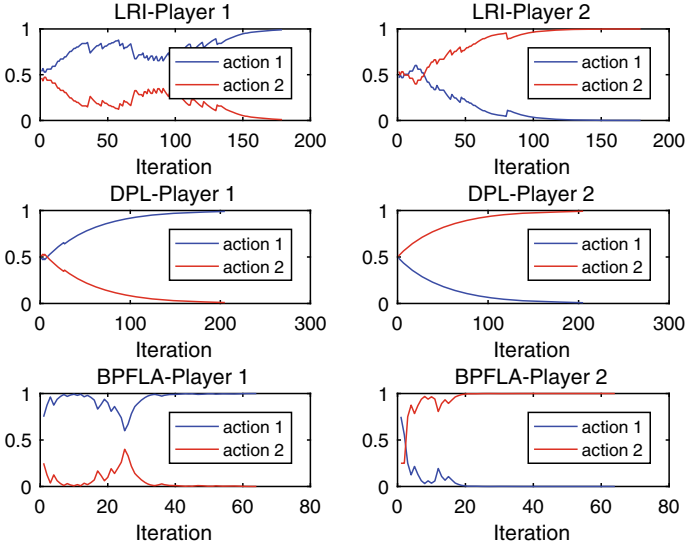


Fig. 1. Action probabilities Pr_{p_i} for L_{RI} , DPL, BPFLA in identical payoff game \mathbb{D}_1

Table 1. Performance comparison of L_{RI} , DPL, and BPFLA in identical payoff games

Game matrix	L_{RI}			DPL			BPFLA	
	α	Accuracy	Iteration	α	Accuracy	Iteration	Accuracy	Iteration
\mathbb{D}_1	0.07	0.998	174	0.02	0.997	210	0.999	64
\mathbb{D}_2	0.05	0.999	326	0.01	0.996	412	0.999	67

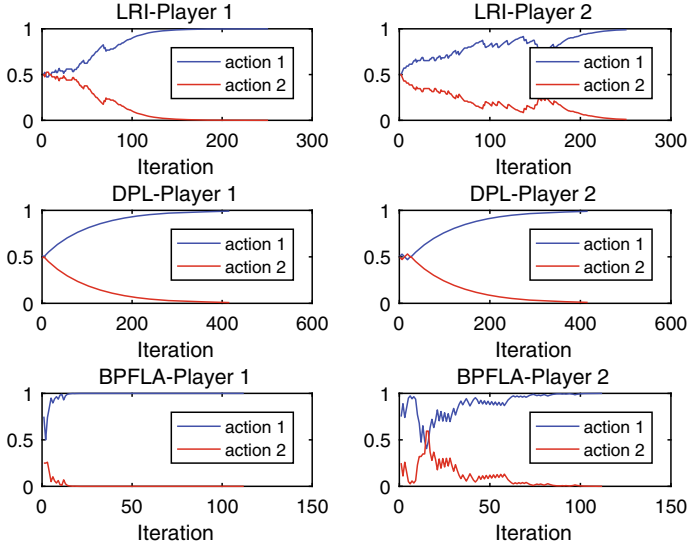


Fig. 2. Action probabilities Pr_{p_i} for L_{RI} , DPL, BPFLA in identical payoff game \mathbb{D}_2

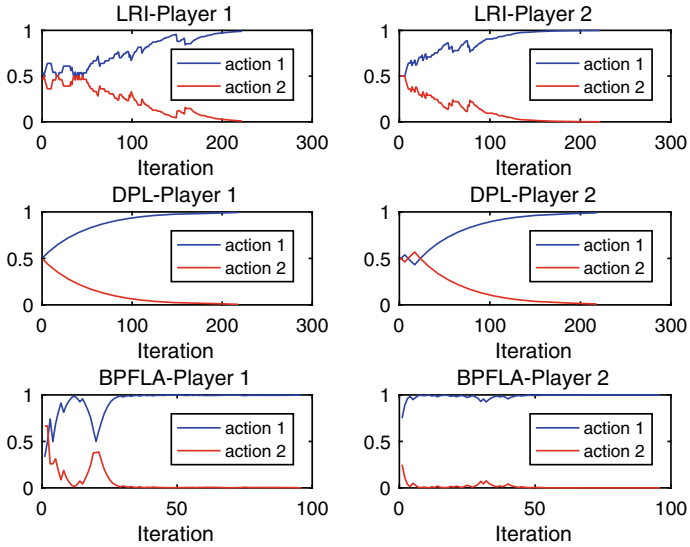


Fig. 3. Action probabilities Pr_{p_i} for L_{RI} , DPL, BPFLA in zero-sum game \mathbb{D}_1

4.2 Two-Player Zero-Sum Payoff Game

Similarly, the zero-sum games were generated in game matrices \mathbb{D}_3 and \mathbb{D}_4 [7], where \mathbb{D}_3 has the saddle point (1, 1) and \mathbb{D}_4 has the saddle point (1, 2).

$$\mathbb{D}_3 = \begin{bmatrix} 0.6 & 0.8 \\ 0.3 & 0.7 \end{bmatrix} \quad \mathbb{D}_4 = \begin{bmatrix} 0.9 & 0.8 \\ 0.5 & 0.3 \end{bmatrix}$$

The changes of probability vectors of each action in typical simulations in game matrices \mathbb{D}_3 and \mathbb{D}_4 are demonstrated in Figs. 3 and 4.

In addition, Table 2 demonstrates that BPFLA has faster convergence rate than L_{RI} and DPL with equal or higher accuracy. The convergence rate of L_{RI} and DPL are improved by BPFLA with 55.6%, 54.8% in \mathbb{D}_3 and with 32.6, 52.1% in \mathbb{D}_4 .

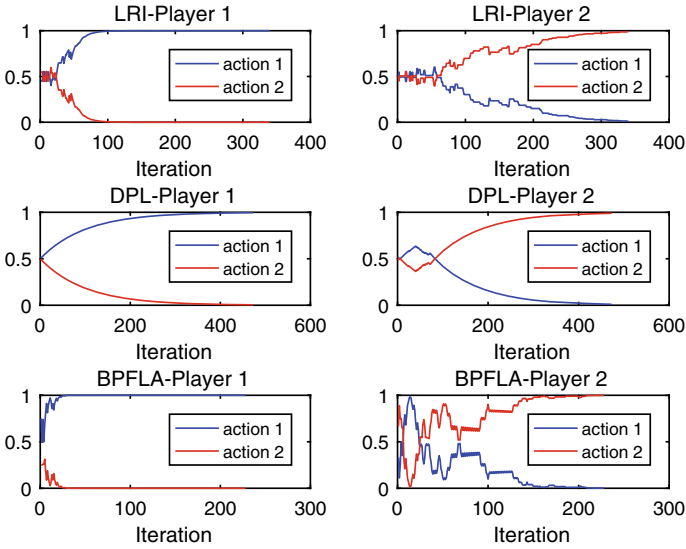


Fig. 4. Action probabilities Pr_{p_i} for L_{RI} , DPL, BPFLA in zero-sum game \mathbb{D}_2

Table 2. Performance comparison of L_{RI} , DPL, and BPFLA in zero-sum games

Game matrix	L_{RI}			DPL			BPFLA	
	α	Accuracy	Iteration	α	Accuracy	Iteration	Accuracy	Iteration
\mathbb{D}_3	0.08	0.998	223	0.02	0.998	219	0.999	99
\mathbb{D}_4	0.1	0.998	340	0.01	0.998	478	0.999	229

5 Conclusion

In this paper, focusing on two-player stochastic games with incomplete information, we proposed a Bayesian method-based parameter-free learning automaton. Compared with previous LA schemes using identical payoff and zero-sum games, the proposed BPFLA has parameter-free property which indicates that BPFLA can be applied in games without the process of parameter tuning. Besides, simulation results demonstrate that BPFLA has much faster convergence rate than L_{RI} and DPL with equal or higher accuracy in both identical payoff and zero-sum games.

Nevertheless, the BPFLA scheme is far from perfect, and there are still some limitations. The exploration strategy in multi-action games and how the scheme can be extended to multi-player games are the two major challenges.

References

1. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An introduction, vol. 1, No. 1. Cambridge, MIT press (1998)
2. Narendra, K.S., Thathachar, M.A.: Learning Automata: An Introduction. Courier Corporation (2012)
3. Thomas, L.C.: Games, Theory And Applications. Courier Corporation (2012)
4. Wang, H., et al.: Reinforcement learning for constrained energy trading games with incomplete information. *IEEE Trans. Cybern.* **47**(10), 3404–3416 (2017)
5. Fu, K.S., Li, T.J.: Formulation of learning automata and automata games. *Inf. Sci.* **1**(3), 237–256 (1969)
6. Lakshminarayanan, S., Narendra, K.S.: Learning algorithms for two-person zero-sum stochastic games with incomplete information: a unified approach. *SIAM J. Control. Optim.* **20**(4), 541–552 (1982)
7. Tilak, O., Ryan, M., Mukhopadhyay, S.: Decentralized indirect methods for learning automata games. *IEEE Trans. Syst. Man Cybern. Part B (Cybernetics)* **41**(5), 1213–1223 (2011)
8. Ge, H., et al.: A parameter-free gradient Bayesian two-action learning automaton scheme. In: Proceedings of the 2015 International Conference on Communications, Signal Processing, and Systems. Springer, Berlin (2016)
9. Ge, H.: A parameter-free learning automaton scheme (2017). [arXiv:1711.10111](https://arxiv.org/abs/1711.10111)
10. Gupta, A.K., Nadarajah, S. (eds.): Handbook of Beta Distribution and its Applications. CRC Press, Boca Raton (2004)



A Double Competitive Strategy-Based Learning Automata Algorithm

Chong Di¹, Mingda Guo², Jinchao Huang¹, and Shenghong Li¹(✉)

¹ School of Cyber Space Security, Shanghai Jiao Tong University, 800 Dong Chuan Road, Shanghai 200240, China

{dichong95,hjc2015,shli}@sjtu.edu.cn

² School of Mechanical Design manufacture and Automation Major, Taiyuan University of Technology, Ying Ze Road, Taiyuan 030000, China

2528288925@qq.com

Abstract. Learning automaton is considered as one of the most potent tools in reinforcement learning. The family of estimator algorithms is proposed to improve the convergence rate of learning automaton and has made significant achievements. However, the estimators perform poorly on estimating actions' reward probabilities in the initial stage of the learning process. In this situation, a lot of rewards would be assigned to nonoptimal actions. Thus, numerous extra iterations are required to compensate for these wrong rewards. To further improve the speed of convergence, we propose a new P-model absorbing learning automaton using a double competitive strategy to update the action probability vector. The proposed scheme overcomes the drawbacks of the existing action probability vector updating strategy. And, extensive experimental results in benchmark environments demonstrate that the proposed learning automata perform more effectively than the most classic learning automaton SE_{RI} and the current fastest learning automaton $DGCPA^*$.

Keywords: Learning automata · Stationary environments · Estimator algorithms · Reinforcement learning

1 Introduction

Learning automaton (LA) is a decision maker that can update its strategy to learn the optimal action through interacting with the random environment [1]. As a powerful tool in the adaptive learning system, LA has a wide range of applications [2–4].

In general, the convergence rate is one of the vital considerations of learning algorithms. Therefore, Thathachar et al. designed a new class of learning automata, called estimator algorithms [5, 6]. The estimator algorithms have a faster convergence rate than all the previous ones and SE_{RI} is the most estimator algorithm.

Due to the superiority of estimator algorithms, there are many novel estimators [4, 7, 8] are proposed in recent years. In 2015, Ge et al. [7] proposed a deterministic estimator-based LA (Discretized Generalized Confidence Pursuit Algorithm, *DGCPA*) and extended the algorithm to a stochastic estimator-based scheme, of which the estimate of each action is the upper bound of a confidence interval. The improved stochastic estimator-based *DGCPA** is the current fastest LA.

Although the family of estimator-based learning automata has achieved significant improvements, there are still some drawbacks. Because of the fundamental defect, the estimates cannot always be strictly unmistakable. Especially in the initial stage of the learning process, the estimator may perform poorly on estimating the reward probabilities of each action. Because the update scheme of probability vector is discrete and linear, redundant iterations are needed to compensate for the wrong updates of probability vector whenever the nonoptimal action is rewarded. Thus, many extra iterations are required to converge to the optimal action.

In this paper, to overcome the drawbacks of existing estimator algorithms and further speed up convergence, a novel double competitive strategy-based nonlinear method is introduced to update the action probability vector. The first competitive strategy of *DCA* is the standard scheme which has been used in most LA. It is only the action which has the current highest estimate of reward probability that will be rewarded. And, the second strategy is nonlinear that whenever the optimal action which has the highest estimate changes, the probability of new optimal action gets a huge increase while the original optimal action will be severely punished. Accordingly, the wrong rewards could be corrected instantly. Thus, the *DCA* can converge rapidly without extra iterations and outperforms the current fastest LA *DGCPA**.

The fundamental contributions of this paper are summarized as follows.

- To improve the action probability vector updating strategy, we proposed a double competitive strategy-based learning automata algorithm.
- The proposed *DCA* is compared with the most classic LA *SE_{RI}* and the fastest LA *DGCPA** in various stationary P-model random environments. The results indicate that the proposed *DCA* is more effective than both two schemes.

This paper is organized as follows. In Sect. 2, we introduce the general idea of LA and the estimator algorithms. The *DCA* scheme is presented in Sect. 3, and extensive simulation results are presented to show the superiority of the proposed *DCA* scheme over the most classic LA *SE_{RI}* and the fastest LA *DGCPA** in Sect. 4. We conclude the paper in the last section.

2 Learning Automata and Estimator Algorithms

2.1 Learning Automaton

A LA is defined by a quintuple $\langle A, B, Q, F(\cdot, \cdot), G(\cdot) \rangle$, where

- $A = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$ is the set of outputs or actions, and α_t is the action chosen by the automata at any time instant t .
- $B = \{\beta_1, \beta_2, \dots, \beta_m\}$ is the set of inputs to the automata, and β_t is the input at any time instant t . The set could be finite or infinite. In this paper, we consider the case that $B = \{0, 1\}$, where $\beta = 0$ represents the events that the LA has been penalized, and $\beta = 1$ represents the events that the LA has been rewarded.
- $Q = \{q_1, q_2, \dots, q_s\}$ is the set of finite states, and q_t is the state of the automata at time instant t .
- $F(\cdot, \cdot) : Q \times B \rightarrow Q$ is a mapping in terms of the state and input at any time instant t , such that, $q(t+1) = F(q(t), \beta(t))$.
- $G(\cdot)$ is the output function which determines the output of the automata depending on the state q_t , such that, $\alpha(t) = G(q(t))$.

2.2 Estimator Algorithms

The estimator algorithms keep estimating the reward probabilities of actions, and then the estimates are utilized to update the probability vector. According to the definition of estimators, the estimator algorithms could be divided into two classes: deterministic estimator algorithms and stochastic estimator algorithms.

The deterministic estimator $D'(t) = [d'_1(t), \dots, d'_r(t)]$ can be computed using the following formula [9, 10]:

$$d'_i(t) = \frac{W_i(t)}{Z_i(t)}, \forall i = 1, 2, \dots, r \quad (1)$$

where $W_i(t)$ is the number of times the action α_i has been rewarded until the current time t and $Z_i(t)$ is the number of times the action α_i has been selected until the current time t .

The implementation of stochastic estimator $U(t)$ [11] is to impose a random perturbation to the deterministic estimate, such that

$$u_i(t) = d'_i(t) + R_i^t \quad (2)$$

where $u_i(t)$ is the stochastic estimate of action α_i 's reward probability at time t . R_i^t is a random number which is uniformly distributed in an interval. The width of the interval depends on a design parameter γ and $Z_i(t)$. That is, $R_i(t) \in (-\frac{\gamma}{Z_i(t)}, \frac{\gamma}{Z_i(t)})$.

3 Double Competitive Algorithm

The proposed *DCA* is a learning automaton which updates the action probability vector with a double competitive strategy. The first competitive strategy of *DCA* is that only the action which has the highest current stochastic estimate of the reward probability gets the opportunity to increase its probability. That is, at each iteration, if action a_m has the highest stochastic estimate,

$$\begin{aligned} p_i(t) &= \max\{p_i(t) - \Delta, 0\}, \forall i \neq m \\ p_m(t) &= 1 - \sum_{i \neq m} p_i(t). \end{aligned} \quad (3)$$

where $\Delta = 1/r/n$ is the step size. Moreover, the second strategy is nonlinear that whenever the optimal action which has the highest estimate changes, the probability of new optimal action α_m gets a huge increase while original optimal action α_{m_0} will be severely punished.

$$\begin{aligned} p_m(t) &= p_m(t) + \mu * p_{m_0}(t) \\ p_{m_0}(t) &= (1 - \mu) * p_{m_0}(t) \end{aligned} \quad (4)$$

where $\mu \in (0, 1)$ is an attenuation factor.

With the unique two competitive strategies, the wrong rewards could be corrected instantly. The optimal action would always be changing as the estimator is not reliable enough in the early stages of learning which will cause the probability of each action fluctuates continuously. However, when the estimator is sufficiently reliable, the LA will converge rapidly.

Besides, since the dramatic changes of the probabilities of any possible action during the learning process, the actions whose probabilities used to be relatively small get more opportunities to be selected. Then their deterministic estimates would be further updated. Therefore, during the learning process, the estimate of each nonoptimal action gets more opportunities to be updated. According to the Law of Large Numbers, the precision of the estimators would be higher. So the exploration scheme in *DCA* would be more effective than other estimator based schemes.

The procedure of *DCA* is briefly introduced in Algorithm 1.

Note that, the double competitive strategy is reflected twice in the probability updating procedures. Line 8–line 11 are the implementation of the first competitive strategy, only the action which has the highest current stochastic estimate will be rewarded, and to satisfy $\sum_{i=1}^r p_i(t) = 1$ the probabilities of all the other actions decrease. The second competitive strategy is summarized in line 12–line 15 where whenever the optimal action which has the highest current stochastic estimate changes ($m \neq m_0$), the probability of the original optimal action α_{m_0} is reduced by μ . Moreover, the new optimal action α_m will get an additional reward which equals to the reduced probability of action α_{m_0} .

4 Simulation Results

In the following, the proposed *DCA* scheme is compared with the most classic LA *SE_{RI}* and the current fastest LA *DGCPA**.

Within the context of LA, the speed of convergence is compared by the iterations required for convergence under the five benchmark environments [11]. The actions reward probabilities for each environment are as follows:

- E_1 : $D = \{0.65, 0.50, 0.45, 0.40, 0.35, 0.30, 0.25, 0.20, 0.15, 0.10\}$.
- E_2 : $D = \{0.60, 0.50, 0.45, 0.40, 0.35, 0.30, 0.25, 0.20, 0.15, 0.10\}$.
- E_3 : $D = \{0.55, 0.50, 0.45, 0.40, 0.35, 0.30, 0.25, 0.20, 0.15, 0.10\}$.
- E_4 : $D = \{0.70, 0.50, 0.30, 0.20, 0.40, 0.50, 0.40, 0.30, 0.50, 0.20\}$.
- E_5 : $D = \{0.10, 0.45, 0.84, 0.76, 0.20, 0.40, 0.60, 0.70, 0.50, 0.30\}$.

Algorithm 1 The *DCA* scheme

Require: n, γ resolution parameters, μ attenuation factor

Initialize $\Delta = 1/r/n$ the step size

Initialize m_0 = a random integer within $[1, r]$

Initialize $p_i(t) = 1/r$ for $1 \leq i \leq r$

- 1: Initialization: Select each action for 10 times and initialize $W_i(t), Z_i(t)$ for $1 \leq i \leq r$
 - 2: **repeat**
 - 3: At time t , choose an action $\alpha(t)$, according to the probability distribution $P(t)$.
 - 4: Receive a feedback $\beta(t) \in \{0, 1\}$ from stochastic environment.
 - 5: Set $W_i(t) = W_i(t-1) + \beta(t)$, $Z_i(t) = Z_i(t-1) + 1$.
 - 6: Compute the deterministic estimate $d'_i(t)$ according to (1).
 - 7: Compute stochastic estimates $u_i(t)$ according to (2).
 - 8: **if** $\beta(t) = 1$ (reward) **then**
 - 9: Select the action α_m that has the highest stochastic estimate of reward probability, where $\alpha_m = \max\{u_i(t)\}$.
 - 10: Update the probability vector $P(t)$ according to (3).
 - 11: **end if**
 - 12: **if** $m \neq m_0$ **then**
 - 13: Update the probability vector $P(t)$ according to (4).
 - 14: Update $m_0 = m$
 - 15: **end if**
 - 16: **until** Convergence
-

In all the simulations performed in this section, we have the same setting as [11] that the parameter tuning process and convergence threshold T are all the same. The convergence threshold T is set to 0.999 and the attenuation factor μ for *DCA* is set to 90%.

Before presenting the overall simulation results, a typical ordinary experiment would be executed to show the difference between *DCA* and *SE_{ri}* during the learning process. The curves that represent the probability of the optimal action are presented in Fig. 1. The results indicate that the probability of *DCA* changes

dramatically in the initial stage of the learning process. With the number of iterations increasing, the stochastic estimator becomes gradually reliable. When the estimator is sufficiently reliable, the learning automaton converges rapidly. And during the learning process of the SE_{RI} scheme, once the probability of the optimal action decreases, a lot of extra iterations are needed to compensate for the loss probability.

The accuracies (number of correct convergences/number of experiments) of DCA , SE_{RI} and $DGCPA^*$ in environment E_1 to E_5 when using the best learning parameters are presented in Table 1. The results show that DCA always has better accuracy than the other two algorithms. To ensure that the performance comparisons between DCA , SE_{RI} and $DGCPA^*$ is fair, let us verify the number of iterations required to achieve the same accuracy, a series of experiments have been carried out. The results are shown in Tables 2 and 3 which demonstrate that the DCA scheme converges with a much faster speed than SE_{RI} and $DGCPA^*$.

Compared with the most classic LA model SE_{RI} , the proposed scheme DCA achieves a great improvement in the speed of convergence in all benchmark environments. For example, in environment E_2 , The DCA converges in 633

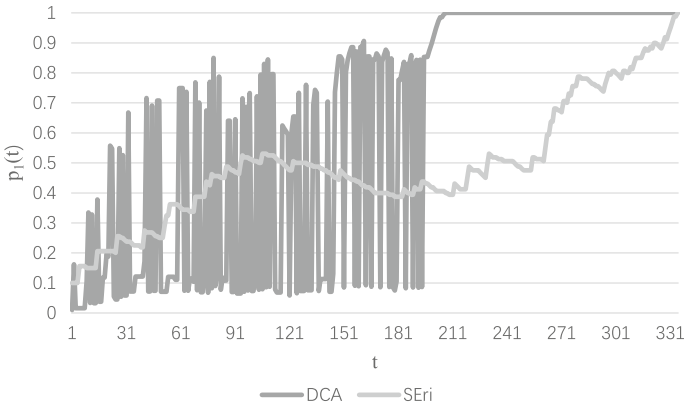


Fig. 1. The curves of optimal actions' probabilities of DCA and SE_{ri} when operating in environment E_1 . For both schemes, the best learning parameters are used and the extra 100 iterations used to initialize the estimator vectors are not included.

Table 1. Accuracy (number of correct convergences/number of experiments) of DCA and SE_{RI} in environment E_1 to E_5 , when using the best learning parameters (250,000 experiments were performed for each scheme in each environment)

	E_1	E_2	E_3	E_4	E_5
DCA	0.998	0.997	0.996	0.999	0.998
SE_{RI}	0.997	0.996	0.995	0.998	0.997
$DGCPA^*$	0.997	0.996	0.995	0.998	0.997

Table 2. Comparison of the average time required for convergence achieving the same accuracy as SE_{RI} shown in Table 1 in environment E_1 to E_5 (250,000 experiments were performed for each scheme in each environment)

Environment	DCA		SE_{RI}		$DGCPA^*$	
	Parameter	Iterations	Parameter	Iterations	Parameter	Iterations
E_1	$n = 10, \gamma = 6$	338	$n = 16, \gamma = 8$	426	$n = 3, \gamma = 5$	351
E_2	$n = 18, \gamma = 8$	633	$n = 32, \gamma = 12$	834	$n = 6, \gamma = 9$	678
E_3	$n = 30, \gamma = 16$	1990	$n = 105, \gamma = 25$	2540	$n = 19, \gamma = 20$	2032
E_4	$n = 9, \gamma = 5$	282	$n = 13, \gamma = 6$	325	$n = 2, \gamma = 4$	298
E_5	$n = 28, \gamma = 7$	582	$n = 33, \gamma = 12$	729	$n = 5, \gamma = 7$	598

Table 3. Comparison of the average time required for convergence achieving the same accuracy as SE_{RI} shown in Table 1 in environment E_1 to E_5 (250,000 experiments were performed for each scheme in each environment)

Environment	DCA		SE_{RI}		$DGCPA^*$	
	Parameter	Time (ms)	Parameter	Time (ms)	Parameter	Time (ms)
E_1	$n = 10, \gamma = 6$	0.162	$n = 16, \gamma = 8$	0.239	$n = 3, \gamma = 5$	3.423
E_2	$n = 18, \gamma = 8$	0.339	$n = 32, \gamma = 12$	0.467	$n = 6, \gamma = 9$	7.417
E_3	$n = 30, \gamma = 16$	1.167	$n = 105, \gamma = 25$	1.420	$n = 19, \gamma = 20$	26.577
E_4	$n = 9, \gamma = 5$	0.126	$n = 13, \gamma = 6$	0.182	$n = 2, \gamma = 4$	2.744
E_5	$n = 28, \gamma = 7$	0.351	$n = 33, \gamma = 12$	0.408	$n = 5, \gamma = 7$	7.252

iterations, while the SE_{RI} requires 834 iterations. Thus, an improvement of 24.10% in comparison with SE_{RI} is obtained.

On the other hand, the current fastest LA model $DGCPA^*$ performs less competitively than the proposed DCA scheme. The superiority of DCA is not only reflected in the fewer number of iterations for convergence but also established in the time efficiency. Because of the complexity of $DGCPA^*$ model when computing the confidence interval, the time required for convergence increases rapidly. Thus, the superiority of the proposed DCA scheme is clear.

5 Conclusions

In this paper, a novel P-model absorbing learning automaton is introduced. With the use of double competitive strategy, the proposed DCA scheme overcomes the drawbacks of existing estimator algorithms. The benefits of the proposed scheme are analyzed, and extensive simulations have been performed in five benchmark environments, and the results indicate that the proposed DCA scheme converges faster and performs more effectively than the most classic LA SE_{RI} and the current fastest LA $DGCPA^*$. The future work will focus on proving the proposed DCA scheme is ϵ -optimal.

Acknowledgement. This research work is funded by the National Key Research and Development Project of China (2016YFB0801003).

References

1. Narendra, K.S., Thathachar, M.A.: Learning automata: an introduction. Courier Corporation (2012)
2. Wang, Y., et al.: Learning automata based cooperative student-team in tutorial-like system. In: International Conference on Intelligent Computing. Springer, Cham (2014)
3. Zhao, Y., et al.: A cellular learning automata based algorithm for detecting community structure in complex networks. *Neurocomputing* **151**, 1216–1226 (2015)
4. Jiang, W.: A new class of optimal learning automata. In: International Conference on Intelligent Computing. Springer, Berlin (2011)
5. Thathachar, M.A.L., Sastry, P.S.: A new approach to the design of reinforcement schemes for learning automata. *IEEE Trans. Syst. Man Cybern.* **1**, 168–175 (1985)
6. Thathachar, M.A., Sastry, P.S.: Estimator algorithms for learning automata (1986)
7. Ge, H., et al.: A novel estimator based learning automata algorithm. *Appl. Intell.* **42**(2), 262–275 (2015)
8. Jiang, W., et al.: A new prospective for learning automata: a machine learning approach. *Neurocomputing* **188**, 319–325 (2016)
9. Sastry, P.S.: Systems of learning automata: Estimator algorithms applications. Dissertation, Ph.D. thesis, Department of Electrical Engineering, Indian Institute of Science, Bangalore, India (1985)
10. Thathachar, M.A.L.: Discretized reward-inaction learning automata. *J. Cybern. Inf. Sci.* **2**, 24–29 (1979)
11. Papadimitriou, G.I., Sklira, M., Pomportsis, A.S.: A new class of ϵ -optimal learning automata. *IEEE Trans. Syst. Man Cybern. Part B (Cybernetics)* **34**(1), 246–254 (2004)



A Learning Automata-Based Compression Scheme for Convolutional Neural Network

Shuai Feng^(✉), Haonan Guo, Jichao Yang, Zhengwu Xu, and Shenghong Li

School of Cyber Security, Shanghai Jiaotong University, Shanghai 200240, China
{feng_shuai,shli}@sjtu.edu.cn

Abstract. The convolutional neural network has been proved to be the state-of-the-art technique in image classification problems. In general, the improved recognition accuracy of the CNN is often accompanied by the increase of structure complexity. However, apart from the accuracy issues, computational resources and operating speed need to be considered on some occasions. Therefore, we propose an efficient compression scheme based on learning automata, which are usually used to choose the optimal action as a reinforcement learning method in this paper. Our proposed method can help the trained CNN to delete insignificant convolution kernels according to the actual requirements. According to the results of experiments, the proposed scheduling method can effectively compress the number of convolutional kernels at the expense of losing weak classification accuracy.

Keywords: Convolutional neural network · Learning automata · Redundant convolution kernels

1 Introduction

In recent years, artificial intelligence (AI) has made a spurt of progress with the help of deep learning algorithm. Therefore, the convolution neural network (CNN), an indispensable member of the depth learning algorithm, also naturally draws the attention of many researchers. In 1998, Yann LeCun proposed the classic LeNet-5 network structure [1] to recognize handwritten numbers and achieved outstanding experimental results. Then, a number of excellent models of CNN subsequently emerged in the subsequent years [2], and the substantial improvements on network structure also have a considerable impact on advancing the state of CNN which constitute the improvements in performance as well as application areas.

With the continuous development of artificial intelligence, CNN has also been used widely in different fields. For example, lots of image retrieval systems tend to take advantage of the features extracted from CNN. In addition, how to apply the neural networks in embedded devices is widely concerned as well. In these situations, operating speed and computational resources must be treated seriously.

In order to address the above challenges, many researchers have already focused on simplifying the convolution neural network with different strategies which turn out to be successful. Reference [3] proposes to prune some weakly connected weights which can be found by LA [4] algorithm during training. Reference [5] removes insignificant connections according to the set threshold and then train the network structure again to achieve the effect of compressing the network.

In this thesis, instead of pruning CNN during the training process, we propose a method that can reduce the redundant convolution kernels as much as possible under our performance goal on the trained CNN. A reinforcement learning method called learning automata [6] is used to attain this objective. When setting a criterion of accuracy, the trained CNN can automatically delete the redundant convolution kernels and select the appropriate convolution kernel combination to confirm the current accurate requirements. The contributions of this paper are as follows:

1. We propose a kernel removing method by combining multiple learning automata to reduce the redundant convolution kernels of the trained CNN for the first time.
2. The dynamic rewards-based scheme is used to improve the performance of learning automata on reducing the number of convolution kernels.
3. We validated our method on both the MINST dataset and the CFAR-10 dataset, and the simulation results show that the number of convolution kernels has been greatly reduced with tiny accuracy loss.

The rest of this paper is organized as follows: in the Sect. 2, we will briefly introduce the related theory of convolutional neural networks and learning automata, Sect. 3 gives a particular account of the proposed method. Numerical experiments are carried out in Sect. 4 to demonstrate the effectiveness of the our approach, We conclude the paper in Sect. 5.

2 Related Background

2.1 Convolution Neural Network

CNN, a mathematical model, can approximate some complex mathematical functions such as classification function, after the parameters are trained well through backpropagation algorithm [7]. Although there are lots of structural models of CNN, all of them generally consist of two parts: feature extraction and classifier. Different from ordinary deep neural networks, the feature extraction part of CNN is made up of convolution layer and pooling layer. A typical CNN is demonstrated in Fig. 1.

Convolution layer is illustrated as c-layer in Fig. 1. Except for the input of the first convolutional layer, the input and output of other convolutional layers is called feature maps. Each feature map consists of a number of neurons arranged in a rectangle and the neurons in the same feature map share the same weights, which is called the convolution kernel. Before we started training CNN, the convolution kernels of c-layer are generally initialized in the form of a

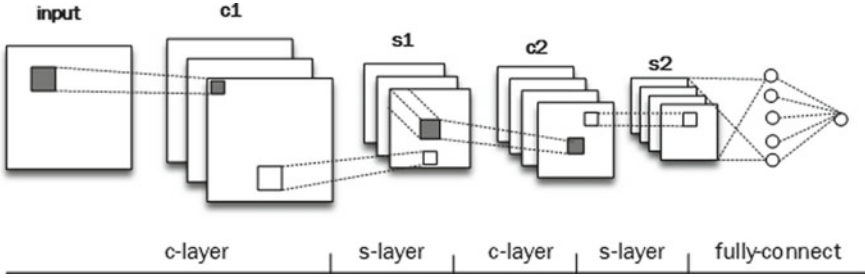


Fig. 1. The structure of CNN

random number matrix, then the convolution kernel learns reasonable weights-based backpropagation algorithm during the training phase. The relationship between the outputs and the inputs of c-layer can be formalized as

$$x^l = f(c^l), c^l = W \cdot m_n^{l-1} + b \quad (1)$$

where x^l denotes a neuron of the output feature map. m_n^{l-1} denotes the partial matrix of n input feature maps. W is the convolution kernel and b is a scalar bias term. $f(\cdot)$ denotes a nonlinearity activation function which is chosen as ReLu function in this paper.

The essence of pooling layer is actually subsampling, which is depicted as s-layer in Fig. 1. Pooling layer, following the convolution layer, can extract the feature from feature map ulteriorly and avoid over-fitting by aggregational statistics of the features in different locations.

2.2 Learning Automata

Learning automata (LA) is a mathematical model that mimics the learning process of the biology. With a small amount of prior knowledge, the automaton achieves a certain goal by constantly responding to a predetermined set of rules or interacting with the environment [8]. As illustrated in Fig. 2, the process of learning consists of two modules: the random environment and the learning automata. To ease out the understanding of the principles applied in our solution approach, we briefly introduce some fundamental concepts of the LA in the following paragraphs.

The LA is defined by a quintuple $\{A, B, Q, T(\cdot, \cdot), G(\cdot)\}$, where:

- $A = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$ represents a finite output set, and α_t is the action chosen by the automaton at any time instant t .
- $B = \{\beta_1, \beta_2, \dots, \beta_m\}$ is the set of inputs and β_t represents the input at any instant t .
- $Q = \{q_1(t), q_2(t), \dots, q_s(t)\}$ is the set of finite states, where $q(t)$ denotes the state of the automaton at any instant t .
- $T(\cdot, \cdot): Q \times B \times A \rightarrow Q$ is the transition function of learning automaton, which determines the state of the automaton at time $t + 1$ depending on the A , B , and the state Q at any time instant t .

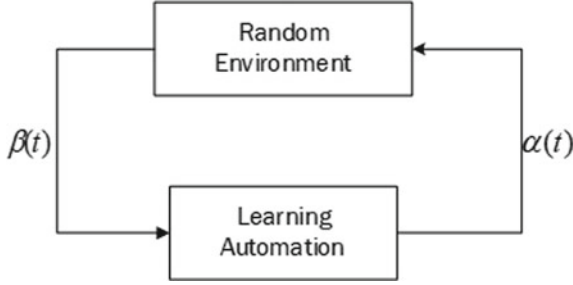


Fig. 2. The automaton-environment feedback loop

- G is a mapping $G: Q \rightarrow A$, and is called the output function. According to the current state, this mapping can help the learning automaton determine the output at any instant t .

The object interacting with the automaton is the environment, and the environment is defined by a quintuple $\{A, B, C\}$. $C = \{c_{ij} = \Pr\{\beta(n) = \beta_j \mid \alpha(n) = \alpha_i\}\}$. If c_{ij} is not a function of time, we call the environment a stationary environment, otherwise the environment is a nonstationary environment.

3 Methods

In order to improve the usability of CNN, we hope to compress the CNN with little or no loss of accuracy. Let K be the set of kernels and k denotes a subset, where $k \subseteq K$. T is the required accuracy threshold, and F represents the network structure of CNN. X is the input of CNN. So the problem can be given below as

$$\min\{k\}, \text{ where } k \subseteq K, F_k(X) \geq T \quad (2)$$

Exhaustive algorithm, evaluating every possible combination of the set kernels, can be used to find the optimal subsets of kernels. It needs $2^n - 1$ iterations and n denotes the number of kernels. However, the complexity of the algorithm cannot be accepted as n increases. To solve the problem, we propose the method based on multiple automatons.

Before presenting the LA-CNN scheme, we would explain that the training set and test set we use to prune kernels are still the same as those of training convolution neural network. In the LA-CNN scheme, each candidate convolution kernel is equivalent to a learning automaton, which has two behaviors: KEEP or REMOVE. For a convolution kernel set, CNN is considered as the interactive environment which gives feedback $B = \{0, 1\}$, where 1 corresponds to reward and 0 corresponds to penalty.

At the beginning of the algorithm, each LA needs to be initialized. T_k and T_r denotes the probability of KEEP and REMOVE respectively. Then, the state Q of each LA can be expressed as $Q = \{T_k, T_r\}$. The behavior vector can be expressed as $A = \{1, 0\}$, where $a = 1$ is the KEEP and $a = 0$ denotes REMOVE,

so LA can choose a behavior based on probability vector Q . We define the $L = \{Q_1, Q_2, \dots, Q_n\}$, where n denotes the number of the kernels. Next, we need to update the structure F of the neural network. We assume that the parameters of the convolution kernel are $W = \{w_1, w_2, \dots, w_n\}$. Bias are $B = \{b_1, b_2, \dots, b_n\}$. $A = \{A_1, A_2, \dots, A_n\}$. S represents other parameters of the CNN, such as the parameters of full link layer. According to the following formula, we can get a new network structure F at time t .

$$F((W^t, B^t, S)) = F((W^{t-1}, B^{t-1}) \cdot A^T, S) \quad (3)$$

In the learning phase, each kernel in networks work as follows: we need to use the train data to get a optimal convolution kernel subset that satisfies the condition. We assume that the threshold is T_a . Once the accuracy of $F(\text{train_dataset})$ exceeds the threshold, we will update the Q . The most important part of the transition function operates as follows: we increase the T_k of the LA whose action is KEEP in the A , but the increment is dynamic and depends on the number of convolution kernels reserved. On the contrary, we take advantage of the fixed steps to increase the $1 - T_k$ of convolution kernels with action REMOVE. The benefit of doing so is to encourage a smaller number of subsets to get higher chances in the next selection. The steps of transition function at each iteration can be formalized as

$$\gamma = \lambda \cdot \frac{\Delta}{\sum_{i=1}^n a_i} \quad (4)$$

where Δ denotes the fixed steps, γ is the dynamic steps, and λ denotes the learning rate. And now, we will present the procedure of LA-CNN scheme in Algorithm 1.

4 Experiments

In order to verify the effectiveness of the proposed algorithm, two different datasets: MINST and Cifar-10, are used, which are widely adopted for training and testing in the field of machine learning. MINST handwritten digital dataset has the handwritten digit images of 0–9, all of which are taken from about 250 individuals. The Cifar-10 dataset contains 32×32 pixel color images in 10 categories and each class are 6000 pictures. This dataset is divided into 5 training sets and a test set, each with 10,000 pictures.

4.1 The Results on MINST

First of all, we trained the five CNNs with the same structure but different numbers of convolution kernels on the MINST dataset. The total number of convolution kernels is 24, 48, 96, 128, 192, respectively. For all the models, we utilize the ReLU activation function and stochastic gradient descent (SGD) optimizer. There are seven layers including two convolution layers in each model which is similar to the LeNet-5 [1]. Before presenting the simulation results, we will show the values of the related parameters. We set the same threshold $T_a = 0.98$ for

Algorithm 1 Learning automata algorithm to prune CNN

input: Given a CNN with n convolution kernels, dataset;
initialize: $T_a, T_k, T_r, \Delta, \lambda$;
repeat
 At time instant t , select an action a according to the probability distribution Q_i
 for every LA;
 update the action set $A = \{A_1, A_2, \dots, A_n\}$;
 According to the set A and new network structure $F((W^t, B^t, S)) = F((W^{t-1}, B^{t-1}) \cdot A^T, S)$;
 $accuracy \leftarrow F(\text{train dataset})$;
 $n = \sum_{i=1}^n a_i$;
 if $accuracy \geq T_a$ **then**
 for subset A **do**
 if $a_i = KEEP$ **then**
 Update the probability vector Q_i according to the following equations:
 $T_k(t) = \max\{T_k(t-1) + \lambda \cdot \Delta/n, 1\}$;
 $T_r(t) = 1 - T_k(t-1)$;
 else
 $T_k(t) = \min\{T_k(t-1) - \Delta, 0\}$;
 $T_r(t) = 1 - T_k(t-1)$;
 end if
 end for
 update the statue set $L = \{Q_1, Q_2, \dots, Q_n\}$;
 end if
until All convolution kernels (LA) converge ;
Remove the convolution kernels with state REMOVE in the CNN;
Use the final CNN model to make the predict the label of test data.

each model and the learning rate $\lambda_1 = \lambda_2 = \lambda_3 = 1$, $\lambda_4 = \lambda_5 = 5$, The step size $\Delta = 0.01$. Table 1 shows that the CNN with the different numbers of convolution kernels at the same threshold, the proportion of the convolution kernels which are removed and the proportion of decline in the accuracy of the test set. From the Table 1, it can be seen that the method proposed in this paper can remove the redundant kernels successfully with about 1% loss of accuracy. In addition, the experiment also shows that simply increasing the number of convolutions have a minor effect on the improvement of accuracy when the number of convolution kernels reaches a certain level.

Table 1. The experiment in MINST

CNN	LA + CNN	Kernels removed (%)	Original acc (%)	Acc drop (%)
24	18	25.0	98.76	0.79
48	24	50.0	98.77	1.23
96	37	61.4	98.86	1.38
128	63	51.9	98.87	1.15
196	65	66.1	98.84	0.97

4.2 The Results on Cifar-10

Different from the above experiment, the CNN trained on Cifar-10 dataset is fixed. Taking into account the complexity and run time of the network, We make a simple CNN model which is similar to the VGGNet structure and the concrete structure is shown in Fig. 3. We set $T_k = 0.9$ and $\Delta = 0.01$ in the following experiments. The different threshold T_a can be set by decreasing in steps of 2% original accuracy. Considering the run time of convergence for LA, we set twenty hundred thousand iterations as a threshold in this experiment.

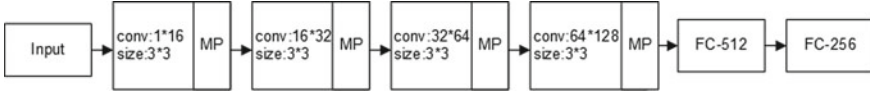


Fig. 3. The structure of CNN. MP represents the pooling layer

Table 2. The experiment in Cifar-10

Original kernels	LA + CNN	Kernels removed (%)	Accuracy (%)
240	240	0	74.44
	205	14.8	73.03
	195	18.8	72.09
	187	22.1	71.12
	174	27.5	70.02

Table 3. The distribution of the redundant convolution kernels

Convolution kernels	Kernel_174	Kernel_187	Kernel_195	Kernel_205
Layer-16	2	3	2	1
Layer-32	1	0	0	0
Layer-64	4	3	2	2
Layer-128	59	47	41	32

Table 2 demonstrates the number of redundant convolution kernels at the loss of different classification accuracy. Compared with the result on the MINST, it can be seen that the Cifar-10 does need more kernels because there is just 14.58% kernels removed at expense of about 1% classification accuracy. But it is possible to get a better result by improving the numbers of iterations. The distribution of the redundant convolution kernels is shown in Table 3. That is obvious that the most of redundant convolution kernel generally appears at the last layer, so the proposed scheme can be also used to verify whether the CNN model designed is reasonable.

5 Conclusion

We propose a method-based learning automata to remove the redundant kernels of CNN in this paper. We adopt the multiple learning automata to help each kernel to decide if it is redundant or not. The experiments have been conducted to show that the proposed scheme can shrink the size of CNN with a little decrease in performance. What is more, it is entirely possible to further improve the effect through adjusting parameters. Further work includes testing the method on different datasets and the diverse models of CNN.

Acknowledgements. This research work is funded by the National Key Research and Development Project of China (2016YFB0801003) and the Sichuan province & university cooperation (Key Program) of science & technology department of Sichuan Province (2018JZ0050).

References

1. Lecun, Y., et al.: Gradient-based learning applied to document recognition. *Proc. IEEE* **86**(11), 2278–2324 (1998)
2. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: *International Conference on Neural Information Processing Systems* Curran Associates Inc. pp. 1097–1105 (2012)
3. Guo, H., et al.: A new learning automata based pruning method to train deep neural networks. *IEEE Internet Things J.* pp. 99, 1–1 (2017)
4. Tsetlin, M.L.: Automaton theory and modeling of biological systems. *Am. Econ. Rev.* 234–244 (1973)
5. Han, S., Mao, H., Dally, W.J.: Deep compression: compressing deep neural networks with pruning, trained quantization and Huffman coding. *Fiber* **56**(4), 3–7 (2015)
6. Thathachar, M., Sastry, P.S.: Varieties of learning automata: an overview. *IEEE Trans. Syst. Man Cybern. (A Publication of the IEEE Systems Man & Cybernetics Society)* **32**(6), 711–722 (2002)
7. Zipser, D., Andersen, R.A.: A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature* **331**(6158), 679–684 (1988)
8. Mostafaei, H., Meybodi, M.R.: Maximizing lifetime of target coverage in wireless sensor networks using learning automata. *Wirel. Pers. Commun.* **71**(2), 1461–1477 (2013)



Improved Efficient Dictionary Learning with Cross-Label and Group Regularization

Tian Zhou, Sujuan Yang, Jian Xiong, Jie Yang^(✉), and Guan Gui

College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
{b14010323, jyang, guiguan}@njupt.edu.cn

Abstract. Existing works have revealed that dictionary learning can effectively preserve the label property when applied to signal reconstruction and face recognition. As time goes, the methods based on dictionary learning become increasingly popular due to their superior accuracy and efficiency. Based on this, an improved dictionary learning model is proposed in this paper to find the balance between the time cost of operating the algorithms and the residuals generated when reconstructing signals with the learnt dictionary sparse codes. Demonstrated by the results of experiments, the proposed method intends to determine a more reasonable sparse dimension, which can not only obtain desired classification results but also decrease much of the redundancy in experimental computation.

Keywords: Cross-label suppression · Computational complexity · Dictionary learning · Face recognition · Compressive sensing

1 Introduction

Much attention has been focused on dictionary learning over the past a few decades and has been applied to many aspects of signal processing, such as signal recovery [1–3], cluster [4], and classification [5]. In essence, dictionary learning is a method to accurately reconstruct a sparse physical signal from an undetermined linear system of equations, where the significant representative matrix is called dictionary and each column of the dictionary is called ‘atom’.

The classic K-SVD algorithm adopts the method of SVD to learn the dictionary one by one column [6]. To obtain lower computational complexity, a structured dictionary model has been utilized in [7]. Supervised learning methods have more advantages for face recognition, and the sparse representation-based classification (SRC) method proves it [5]. In most cases of the supervised learning methods, ℓ_0 -norm [8] or ℓ_1 -norm [9] is considered as a prior item. However, the redundancy of computational time is obviously a major drawback in these works. Thus, ℓ_2 -norm has been adopted in dictionary learning to obtain an iteration solution with closed form, which smooths the iterative manner. What is more, cross-label suppression as well as group regularization

(CLS-GR) is utilized in the dictionary learning [10] in order to achieve a higher accuracy in classification.

In this paper, we propose an improved method which takes both accuracy and efficiency into consideration with cross-label suppression and group regularization. Furthermore, the dimension of sparse representation influences the computational complexity mostly, so that a quantitative relationship is built between sparse dimension and cost time without traversing the situation of all dimensions. Owing to the purpose of saving time, a method of estimating the residuals of signal recovery without operating the algorithms plays a particularly important role of the proposed method.

2 Background

2.1 CLS-GR Dictionary Learning Model

In this section, we will briefly introduce the dictionary learning model—CLS-GR [10]. The image samples of face are linearly combined in the matrix in the form of column vectors, and the information of labels corresponding to the image samples is given, assuming that the number of class is C totally. The dictionary learning model—CLS-GR can be formulated as (1)

$$\begin{aligned} \min_{\mathbf{D}, \mathbf{X}} \sum_{c=1}^C \left(\|\mathbf{Y}^c - \mathbf{D}\mathbf{X}^c\|_F^2 + \beta \|\mathbf{X}^c\|_F^2 \right) \\ + \lambda \sum_{c=1}^C \|\mathbf{P}^c \mathbf{X}^c\|_F^2 + \gamma \text{tr} \left(\mathbf{X}^c \tilde{\mathbf{L}}^c (\mathbf{X}^c)^T \right) \\ \text{s. t. } \|d_k\|_2 = 1, \forall k \end{aligned} \quad (1)$$

This is a general model of dictionary learning with cross-label suppression and group regularization, where $\mathbf{Y} \in R^{m \times p}$ represents the facial image samples and the superscript means belonging to the c -th class, $\mathbf{D} \in R^{m \times n}$ is the learnt dictionary obtained by the training samples, and $\mathbf{X} \in R^{n \times p}$ means the sparse codes corresponding to the image signals. $\mathbf{P}^c \in R^{n \times n}$ is adopted to the method for the purpose of “cross-label suppression,” and \mathbf{P}^c is defined as (2)

$$\mathbf{P}^c(m, n) = \begin{cases} 1 & m = n \ \& \ m \notin (L^c \cup L^0) \\ 0 & \text{elsewhere} \end{cases} \quad (2)$$

where $c = 1, 2, \dots, C$, $\mathbf{P}^c(m, n)$ denotes the (m, n) -th entry of \mathbf{P}^c , and L^c denotes the c -th label. The Laplacian matrix of an N -vertex is defined as (3)

$$\tilde{\mathbf{L}} = \mathbf{M}^{-\frac{1}{2}} \mathbf{L} \mathbf{M}^{-\frac{1}{2}} = \mathbf{I} - \mathbf{M}^{-\frac{1}{2}} \mathbf{W} \mathbf{M}^{-\frac{1}{2}} \quad (3)$$

where \mathbf{W} is the adjacency matrix of graph, \mathbf{M} is a degree matrix with these K maps, and the total variance ($\text{totalVar}(f)$) can be obtained as (4)

$$totalVar = \sum_{k=1}^K \mathbf{f}_k^T \tilde{\mathbf{L}} \mathbf{f}_k = \text{tr} \left(\begin{bmatrix} \mathbf{f}_1^T \\ \mathbf{f}_2^T \\ \dots \\ \mathbf{f}_K^T \end{bmatrix} \tilde{\mathbf{L}} [\mathbf{f}_1, \dots, \mathbf{f}_k] \right) \quad (4)$$

According to (4), the smaller the value of $totalVar(f)$ is, the smoother the map is. In another aspect, the features of corresponding sparse representations with the same label shall be more similar.

2.2 Time Complexity of Different Methods

In computer science, time cost of the algorithms is corresponding to the theories of time complexity. Thus, the analysis of each step in CLS-GR dictionary learning method as well as their corresponding time complexity has been deeply carried out.

The dictionary learning model named CLS-GR utilizes k-means to initialize the label-particular dictionary \mathbf{D}^c one by one, and the time complexity [6] can be expressed as $O(p \cdot T \cdot n \cdot m)$, where p represents the number of samples, n represents the number of clusters as well as the number of atoms in dictionary, m represents the number of features of each sample, and T represents the number of iteration.

In order to initialize the sparse representation— \mathbf{X} , corresponding to all the training samples, and given the initialized dictionary \mathbf{D}_0 , \mathbf{X}_0 can be obtained as (5)

$$\mathbf{X}_0 = (\mathbf{D}_0^T \mathbf{D}_0 + \beta \mathbf{E})^{-1} \mathbf{D}_0^T \mathbf{Y} \quad (5)$$

Its time complexity is $O(n^3)$, and n represents the number of atoms in the dictionary; in another way, it is also the dimension of the sparse codes.

Instead of updating the whole dictionary at one time, the model renews the dictionary atom by atom to fully utilize those that have already been updated. The dictionary is updated by the formula as follows:

$$\tilde{\mathbf{Z}} = \mathbf{Y} - \sum_{k \notin L^c} \mathbf{d}_k \bar{\mathbf{x}}_k - \sum_{k \in L^c, k \neq i} \mathbf{d}_k \bar{\mathbf{x}}_k \quad (6)$$

$$\hat{\mathbf{d}}_i^c = \frac{\tilde{\mathbf{Z}} \cdot \bar{\mathbf{x}}^T}{\|\tilde{\mathbf{Z}} \cdot \bar{\mathbf{x}}^T\|_2} \quad (7)$$

where $\hat{\mathbf{d}}_i^c$ represents the estimated value of i -th column which belongs to the c -th class, $\bar{\mathbf{x}}_k$ denotes estimated value of the k -th row in the whole code matrix \mathbf{X} . Its time complexity can be shown as $O(p \cdot n \cdot m)$, where n represents the number of atoms of the dictionary, m represents the number of features of each sample, p represents the number of samples. A parameter estimation method of least mean square (LMS) is adopted to obtain the sparse representation \mathbf{X}^c of the c -th class, formulated as (8)

$$\widehat{\mathbf{X}}^c = [\mathbf{D}^T \mathbf{D} + \lambda(\mathbf{P}^c)^T (\mathbf{P}^c) + (\beta + \gamma)\mathbf{E}]^{-1} (\mathbf{D}^T \mathbf{Y}^c - \gamma \mathbf{X}^c \bar{\mathbf{L}}^c) \quad (8)$$

As the formula goes, the time complexity is $O(n^3)$ and n represents the number of atoms of the dictionary matrix.

3 An Efficient Dictionary Learning Corresponding to a Reasonable Sparse Dimension

3.1 Cubic Fitting for Sparse Dimension and Time Cost

A cubic fitting method is adopted to estimate the time cost, in which only a few experimental samples with low-dimensional sparse codes is demanded by this method. Make an assumption that the samples are divided into C classes totally, and s can be defined as (9)

$$s = \frac{n}{C} \quad (9)$$

where n represents the dimension of sparse codes. Given a signal with C classes, the time cost can be approximatively written as (10)

$$\sum_{c=1}^C t_c \approx \omega_1 \cdot s^3 + \omega_2 \cdot s^2 + \omega_3 \cdot s^1 + \omega_4 \quad (10)$$

where $\omega_1, \omega_2, \omega_3, \omega_4$ represent coefficients of the cubic function, t_c represents the time of operating the samples of the c -th class.

So that the cubic fitting function connecting s to total running time— \hat{t} can be obtained as (11).

$$\hat{t} \approx \omega_1 \cdot s^3 + \omega_2 \cdot s^2 + \omega_3 \cdot s^1 + \omega_4 \quad (11)$$

It is significant to select a reasonable number of samples which are used to fit the cubic curve. With the aid of it, we can obtain the estimated time cost to substitute the experimental time cost by any value of sparse dimension without traversing all the situations.

3.2 Corresponding Dimension to Residuals Based on SVD

On the premise of the CLS-GR dictionary learning model of dictionary learning, the sparse representations \mathbf{X} 's energy are concentrated on some particular parts of the matrix. So, the residuals can be rewritten as (12)

$$\|\mathbf{Y}^c - \mathbf{D}\mathbf{X}^c\|_F^2 \approx \|\mathbf{Y}^c - \bar{\mathbf{D}}^c \bar{\mathbf{X}}_{L^c}^c\|_F^2 \quad (12)$$

Similarly, separate \mathbf{Y}^c by SVD, so the residuals can be estimated as (13)

$$\begin{aligned} \|\mathbf{Y}^c - \mathbf{D}\mathbf{X}^c\|_F^2 &= \|\mathbf{Y}^c - \sum_{j \notin L^c} \mathbf{d}_j \bar{\mathbf{x}}_j^c - \mathbf{D}^c \mathbf{X}_{L^c}^c\|_F^2 \\ &= \|\mathbf{E}^c - \bar{\mathbf{D}}^c \bar{\mathbf{X}}_{L^c}^c\|_F^2 \\ &\approx \|\mathbf{Y}^c - \sum_{i=1}^s \Lambda(i, i) \mathbf{U}_{c=i} \mathbf{V}_{c=i}^T\|_F^2 \end{aligned} \quad (13)$$

Assume \mathbf{D}^c is a matrix with s columns, $\mathbf{Y}^c = \mathbf{U}\mathbf{A}\mathbf{V}^T$. Based on the classic K-SVD algorithm, we select the first s columns of \mathbf{U} , where $\mathbf{U}_{c=i}$ represents the i -th column in \mathbf{U} . Furthermore, the diagonal matrix composed of the first s singular values multiplying the transpose of the first s columns in \mathbf{V}^T corresponds to $\mathbf{X}_{L^c}^c$. $\bar{\mathbf{X}}_{L^c}^c$ denotes the rows in $\bar{\mathbf{X}}^c$ which belongs to the c -th label. The residuals can be formulated as (14)

$$\|\mathbf{Y}^c - \mathbf{D}\mathbf{X}^c\|_F^2 \approx \left(\sum \Lambda^2(i, i) \right) - \sum_{i=1}^s \Lambda^2(i, i) \quad (14)$$

Based on formula (11) and (14), the residuals of sparse signal reconstruction and the cost time of CLS-GR algorithm can be estimated to determine a reasonable sparse dimension without practical experiments.

3.3 Optimization Object

In order to take both residual and accuracy into consideration, the model has been developed as (15)

$$\begin{cases} \min_s \sum_{c=1}^C \left(\|\mathbf{Y}^c - \mathbf{D}\mathbf{X}^c\|_F^2 + \delta t^c(s) \right) \\ \min_{\mathbf{D}, \mathbf{X}} \sum_{c=1}^C \left(\|\mathbf{Y}^c - \mathbf{D}\mathbf{X}^c\|_F^2 + \beta \sum_{j=1}^{N^c} \|\mathbf{x}_j^c\|_F^2 \right) + \lambda \sum_{c=1}^C \|\mathbf{P}^c \mathbf{X}^c\|_F^2 + \gamma \text{tr}(\mathbf{X}^c \tilde{\mathbf{L}}^c (\mathbf{X}^c)^T) \end{cases} \quad (15)$$

s. t. $\|d_k\|_2 = 1, \forall k$

Among it, $t^c(s)$ represents the cost time of the whole sections, including selecting training samples, initializing the c -th class sparse representations \mathbf{X}^c , and updating of the c -th class dictionary \mathbf{D}^c and sparse representations \mathbf{X}^c , where δ is a parameter to decide the influences of time cost in practical applications. Quantitatively, the most reasonable value of s can be obtained from (16)

$$\begin{aligned}
s &= \arg \min_s \sum_{c=1}^C \|\mathbf{Y}^c - \bar{\mathbf{D}}^c \bar{\mathbf{X}}_{L^c}^c\|_F^2 + \delta \bar{t}^c(s) \\
&= \arg \min_s \frac{1}{C} \sum_{c=1}^C \left(1 - \frac{\sum_{i=1}^s (\boldsymbol{\Lambda}^c(i, i))^2}{\sum (\boldsymbol{\Lambda}^c(i, i))^2} \right) + \delta' (\omega_1 \cdot s^3 + \omega_2 \cdot s^2 + \omega_3 \cdot s^1 + \omega_4)
\end{aligned} \tag{16}$$

The learning methods with the solution of closed form by least square include two steps, obtaining sparse learning of \mathbf{X} by (8), and the dictionary is updated atom by atom by (6) and (7). Repeat the steps above, and the dictionary can be learnt and the sparse representations can be obtained.

3.4 Global Coding Classifier with Reasonable Sparse Dimension

The recognition scheme with a reasonable value of dimension is named global coding classifier with reasonable sparse dimension (RSD-GCC). Given an image sample \mathbf{y} as well as the learnt dictionary \mathbf{D} , due to the label information is unknown, the sparse representation with a general model can be defined as

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 + \beta \|\mathbf{x}\|_2^2 = (\mathbf{D}^T \mathbf{D} + \beta \mathbf{I})^{-1} \mathbf{D}^T \mathbf{y} \tag{17}$$

Due to the structure of the dictionary and the learning method, if the image in form of column vector \mathbf{y} belongs to the c -th class, the large coefficients should be distributed in the particular area of sparse matrix with the c -th label. Considering the analysis, we can define the following metric for face recognition:

$$\text{label}(\mathbf{y}) = \arg \min_c \frac{\|\mathbf{y} - \sum_{k \notin L^c} \mathbf{d}_k \mathbf{x}_{L^c}\|_2^2}{\sum_{k \notin L^c} |\mathbf{x}_{L^c}|} \tag{18}$$

where $\bar{\mathbf{X}}_{L^c}$ represents the rows belonging to the L^c of the sparse code corresponding to the image signal.

4 Simulation Results

In this section, we conduct several computer simulations to compare the proposed combination policy with previous combination rules including both efficiency of our proposed classification and the accuracy of the previous sections. In order to fairly evaluate the computational efficiency, the proposed approach will be operated on the

platform of MATLAB2017b applied in one PC with a 64 bit Windows 10 operating system and equipped with Intel i5-4200H 2.8 GHz CPU and 4 GB memory.

4.1 Face Recognition Based on Extended YaleB Face Dataset

There are 2414 frontal face images with 64 facial images each person belonging to 30 people in Extended YaleB face dataset [11]. On account of comparing the results to the existing methods, the Eigenface feature [12] is compressed in a dimension of 300.

In our experiment, the number of training samples is increased with a step of 5 from 20 to 40, the parameters β, γ, λ are, respectively, set to 8×10^{-3} , 2×10^{-3} , and 2×10^3 . We take both the above two aspects into consideration, including time cost and accuracy of classification, and the performance is shown in Fig. 2a, b.

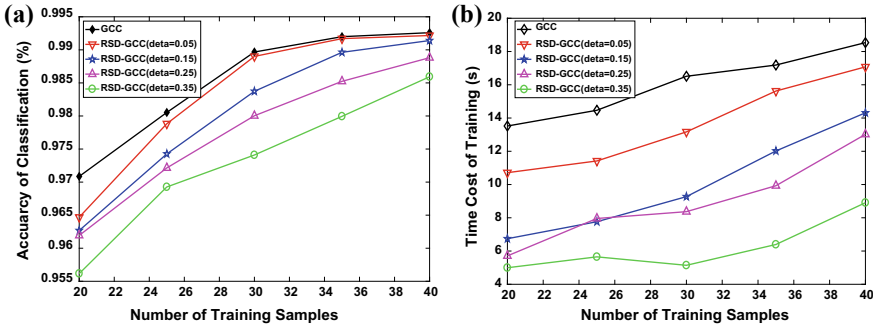


Fig. 1. Accuracies (a) and time cost (b) of face recognition on the extended YaleB dataset

As shown in Fig. 1, on the premise of the loss of the negligible accuracy in face recognition, the time cost of learning can be decreased drastically by our proposed method of determining a reasonable sparse dimension. Compared with GCC [10], on average of less than 1% of accuracy is sacrificed in classification, the training time decreases by more than 20%. In another way, the value of sparse dimension in GCC is not reasonable enough and leads to the redundancy of computation.

4.2 Performance of Cubic Fitting and Estimating Residuals

As shown in Fig. 2, the results of experiments also demonstrate the correctness and strong robustness of our methods of estimation. In the cubic fitting model, k represents the number of fitting samples, and the total number of training sample is 40.

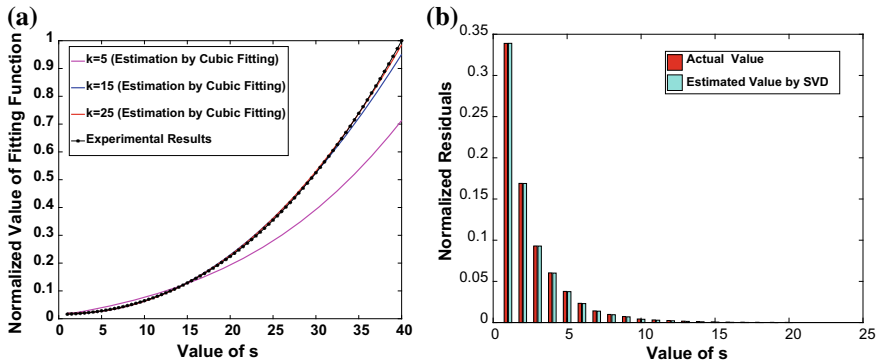


Fig. 2. Effects of cubic fitting method (a) and residual estimation method (b)

Considering the time cost in determining the reasonable dimension, the result of estimation with “ $k = 15$ ” is the most reasonable to obtain a good performance. On the other hand, the residuals estimation obtained by SVD also show accurate estimated values during the process of dictionary learning.

5 Conclusion

In this paper, a cubic fitting model is proposed to substitute the experimental value of cost time in dictionary learning method—CLS-GR. In addition, a method based on SVD is utilized to estimate the residuals in signal reconstruction, and we combine the trade-off between accuracy and computational complexity with the dictionary learning algorithm—CLS-GR. As shown in the experiments, the learning process becomes more reasonable due to the decrease in computation redundancy. Furthermore, the results in the tables compared our method to others, such as SRC [5], LC-KSVD [13], COPAR [14], and the basic GCC [10], and also demonstrate the advantages of our methods in computation reduction.

References

1. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Process.* **54**(11), 4311–4322 (2006)
2. Olshausen, B.A., Field, D.J.: Sparse coding with an overcomplete basis set: a strategy employed by V1? *Vision. Res.* **37**(23), 3311–3325 (1997)
3. Li, Y., et al.: Sparse adaptive iteratively-weighted thresholding algorithm (SAITA) for L_p -regularization using the multiple sub-dictionary representation. *Sens.* **17**(12), 2920–2936 (2017)
4. Wang, F., Lee, N., Sun, J., Hu, J., Ebadollahi, S.: Automatic group sparse coding. In: *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence (AAAI)*. San Francisco, California, USA, 7–11 August 2011

5. Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y.: Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(2), 210–227 (2009)
6. Celebi, M.E., Kingravi, H.A., Vela, P.A.: A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Syst. Appl.* **40**(1), 200–210 (2013)
7. Rubinstein, R., Zibulevsky, M., Elad, M.: Double sparsity: learning sparse dictionaries for sparse signal approximation. *IEEE Trans. Signal Process.* **58**(3), 1553–1564 (2010)
8. Zhang, Q., Li, B.: Discriminative K-SVD for dictionary learning in face recognition. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2691–2698 (2010)
9. Ramirez, I., Sprechmann, P., Sapiro, G.: Classification and clustering via dictionary learning with structured incoherence and shared features. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 3501–3508 (2010)
10. Wang, X., Gu, Y.: Cross-label suppression: a discriminative and fast dictionary learning with group regularization. *IEEE Trans. Image Process.* **26**(8), 3859–3873 (2017)
11. Lee, K.C., Ho, J., Kriegman, D.J.: Acquiring linear subspaces for face recognition under variable lighting. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(5), 684–698 (2005)
12. Turk, M., Pentland, A.: Eigenfaces for recognition. *J. Cogn. Neurosci.* **3**, 209–232 (1991)
13. Jiang, Z., Lin, Z., Davis, L.S.: Label consistent K-SVD: learning a discriminative dictionary for recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(11), 2651–2664 (2013)
14. Wang, D., Kong, S.: A classification-oriented dictionary learning model: Explicitly learning the particularity and commonality across categories. *Pattern Recognit.* **47**(2), 885–898 (2014)



Multi-task Cascaded Convolutional Neural Networks for Real-Time Dynamic Face Recognition Method

Bin Jiang, Qiang Ren, Fei Dai, Jian Xiong, Jie Yang,
and Guan Gui^(✉)

College of Telecommunication and Information Engineering, Nanjing University
of Posts and Telecommunications, Nanjing 21003, China
guiguan@njupt.edu.cn

Abstract. Due to the variety of poses, lighting, and scenes, dynamic face detection and calibration pose a big challenge under unconstrained environment. In this paper, we use the inherent correlation between detection and calibration to enhance their performance in a deep multi-task cascaded convolutional neural network (MTCNN). In addition, we utilize Google's FaceNet framework to learn a mapping from face images to a compact Euclidean space, where distances directly correspond to a measure of face similarity to extract the performance of facial feature algorithms. In the practical application scenario, we set up a multi-camera real-time monitoring system to perform face matching and recognition of collected continuous frames from different angles in real time.

Keywords: Multi-task cascaded convolutional neural networks · Real-time dynamic face recognition · FaceNet · Facial feature algorithms

1 Introduction

Face detection and alignment are indispensable parts of many face-based applications, such as face recognition [1] and facial analysis [2]. Dynamic recognition [3] poses an extreme challenge because of the impact of visual changes on the face, coupled with the effects of posture and lighting.

Recently, convolutional neural network (CNN) [4] was considered as one of the effective methods to achieve good result in visual aspects such as image classification [5] and face recognition [6]. However, these previous face recognition methods often ignore the effect of face angle and lighting. In order to solve these problems, several existing algorithms have been proposed but they are still hard to fully deal with the effect of face angle and lighting [7]. In addition, the performance improvement of the current training sample classifier has certain obstacles [8]. It is necessary to find a classifier that can extract facial features to a greater degree to perform face feature matching in order to achieve a better face recognition effect.

In this paper, we propose a new cascading architecture to improve the dynamic face recognition efficiency. The algorithm consists of three stages: In the first stage, multiple cameras are set up, successive frames are taken from different angles, the candidate

form is retained using MTCNN, and the face part in the video stream is extracted; in the second stage, FaceNet [9] extraction is used. The 128 feature values of the face part are stored in an array. In the third stage, face data sets are classified and matched using KNN [10], and the effect of dynamic face recognition is finally achieved.

2 Algorithm Theory

Multiple cameras are set up from different angles in the scene to collect video streams in real time. Then face frames of the video stream are extracted and sorted on a frame-by-frame basis, and face images that can be accurately used for recognition in consecutive frames are selected, and these filtered consecutive frames are used as MTCNN's inputs, while in the background to detect and correct the face in the picture (Fig. 1).

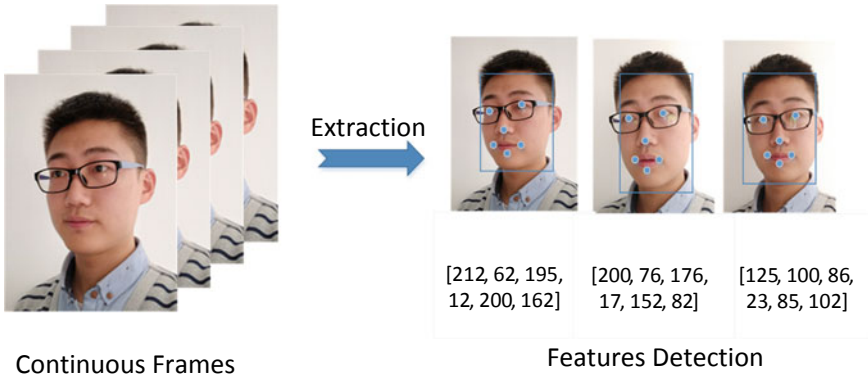


Fig. 1. Continuous images are acquired in real time by camera in real scene, and the feature matrix of keyframe face images is calculated by MTCNN algorithm. The source of the photograph is the photograph taken by the author himself

MTCNN first uses a full convolutional neural network, P-Net, to obtain candidate forms and boundary regression vectors. At the same time, the candidate form is calibrated according to the bounding box. Then, use the NMS method [11] to remove overlapping windows. Then the picture containing the candidate form determined by P-Net is trained in the R-Net network, and the network finally adopts the full connection method for training. Use the bounding box vector to fine-tune the candidate form and then use the NMS to remove the overlapping form. Finally, the network structure is more convolutional than R-Net. The function is the same as that of R-Net. It only shows the key position of the face while removing the overlapping candidate window (Fig. 2).

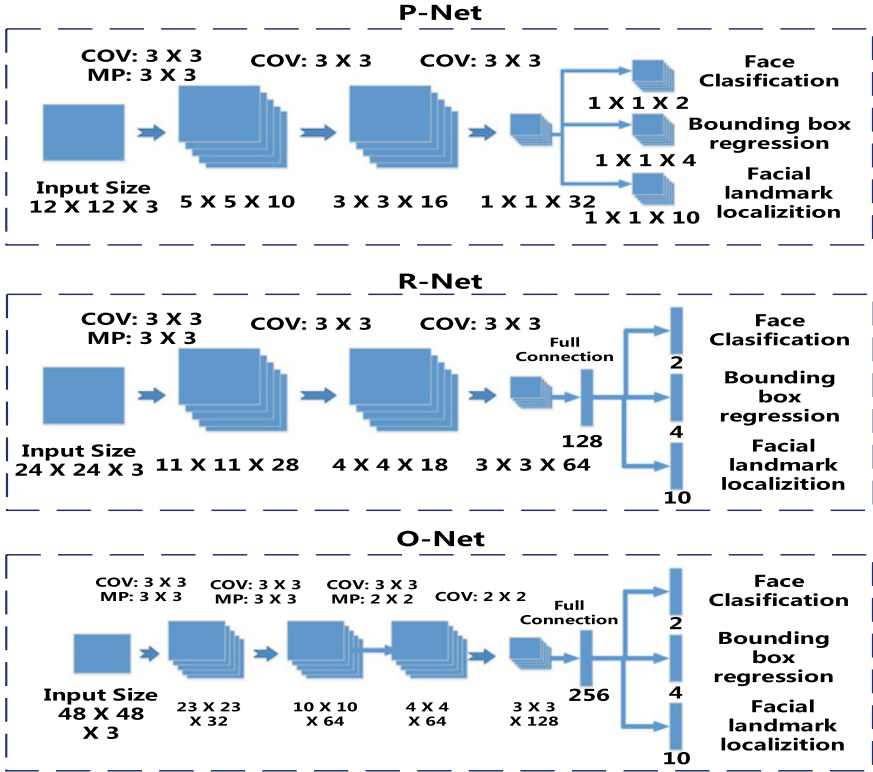


Fig. 2. The proposed MTCNN-based face feature detection method is divided into three steps: (1) proposal net, (2) refine net, and (3) output net

The MTCNN feature descriptor mainly includes three parts: face/non-face classifier, bounding box regression, and landmark location.

$$L_i^{det} = -(y_i^{det} \log(p_i) + (1 - y_i^{det})(1 - \log(p_i))), \text{ where } y_i^{det} \in \{0, 1\} \quad (1)$$

The above formula is a cross-entropy loss function for face classification, where p_i is the probability of the face and y_i^{det} is the real tag of the background.

$$L_i^{box} = \|\hat{y}_i^{box} - y_i^{box}\|_2^2, \text{ where } y_i^{box} \in \mathbb{R}^4 \quad (2)$$

The above equation is the regression loss calculated by the Euclidean distance. Among them, \hat{y} is predicted through the network, and y is the actual real background coordinates and the quad (upper left x, upper left y, long, wide).

$$L_i^{landmark} = \|\hat{y}_i^{landmark} - y_i^{landmark}\|_2^2, \text{ where } y_i^{landmark} \in \mathbb{R}^{10} \quad (3)$$

As with the boundary regression, the Euclidean distance [12] between the predicted landmark position and the actual real landmark is calculated and the distance is minimized. \hat{y} is predicted through the network and y is the actual real landmark coordinates. Since there are a total of five points and two coordinates for each point, y belongs to the ten-tuple. Multiple input source trainings are set as follows:

$$\min \sum_{i=1}^N \sum_{j \in \{det, box, landmark\}} \alpha_j \beta_i^j L_i^j \quad (4)$$

$$\beta_i^j \in \{0, 1\} \quad (5)$$

$$P - Net \quad R - Net \quad (\alpha_{det} = 1, \alpha_{box} = 0.5, \alpha_{landmark} = 0.5) \quad (6)$$

$$O - Net \quad (\alpha_{det} = 1, \alpha_{box} = 0.5, \alpha_{landmark} = 1) \quad (7)$$

The entire training learning process is to minimize the above function, where N is the number of training samples, α_j indicates the importance of the task, β_i^j is the sample tag, and L_i^j is the above loss function.

In the training process, in order to obtain better results, only the gradient of the first 70% of the samples is transmitted backward at a time, so as to ensure that the transmitted numbers are valid. It is similar to latent SVM [13], but it embodies the end-to-end learning of deep learning (Fig. 3).

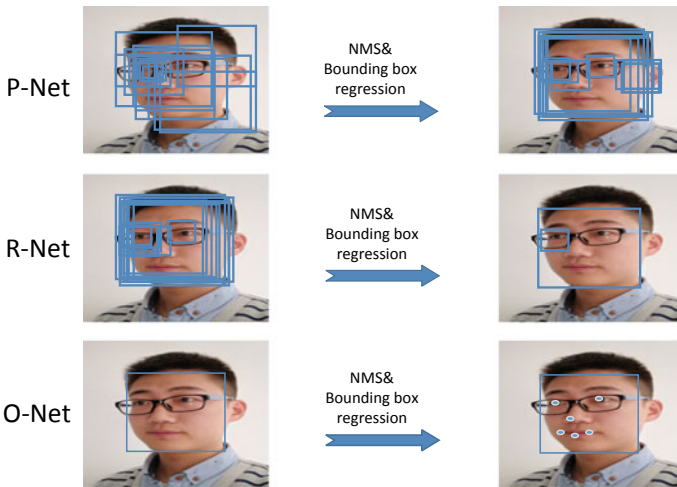


Fig. 3. The flowchart of this method. It can be seen that it is also a three-stage connection

FaceNet learns a mapping from face images to a compact Euclidean space where distances directly correspond to a measure of face similarity. The spatial distance is directly related to the similarity of pictures: different images of the same person have a

small distance in space, and images of different people have a large distance in space. As long as the mapping is determined, the relevant face recognition task becomes simple. FaceNet directly uses the loss function of LMNN (Maximum Boundary Neighbor Classification) [14] based on triplets to train the neural network (Figs. 4 and 5).

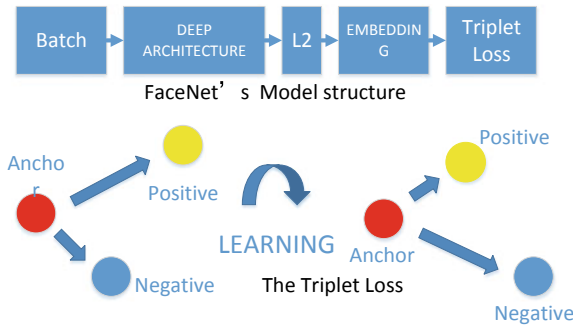


Fig. 4. The network consists of a batch input layer and a deep CNN followed by L2 [15] normalization, which results in the face embedding. This is followed by the triplet loss during training. The triplet loss [16] minimizes the distance between an anchor and a positive, both of which have the same identity, and maximizes the distance between the anchor and a negative of a different identity

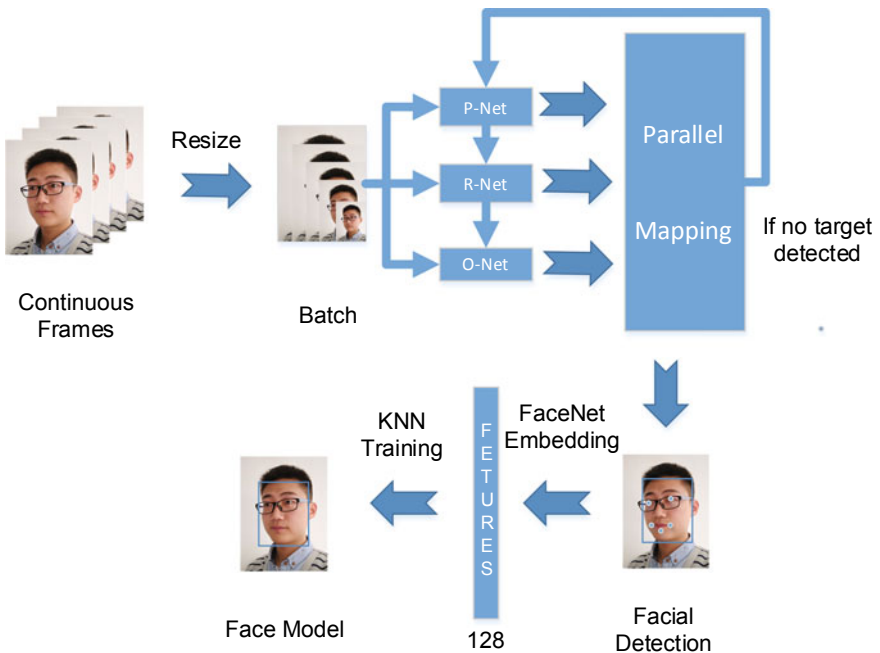


Fig. 5. The flowchart of the entire system

The network directly outputs 128-dimensional vector space.

$$\|x_i^a - x_i^p\|_2^2 + \alpha < \|x_i^a - x_i^n\|_2^2, \forall (x_i^a, x_i^p, x_i^n) \in \mathcal{T} \tag{8}$$

The loss that is being minimized is then $L =$

$$\sum_i^N \left[\|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha \right]_+ \tag{9}$$

where α is the positive/negative boundary. The choice of triplets is very important for the convergence of the model. For x_i^a , we need to select different pictures of the same individual x_i^p making $\operatorname{argmax}_{x_i^p} \|f(x_i^a) - f(x_i^p)\|_2^2$. At the same time, we also need to select different individuals' images x_i^n making $\operatorname{argmin}_{x_i^n} \|f(x_i^a) - f(x_i^n)\|_2^2$. In practical training, it is unrealistic to calculate argmin and argmax across all training samples, and training convergence is difficult due to incorrect label images. So we filter by calculating the subset's argmin and argmax every n steps (Fig. 6).

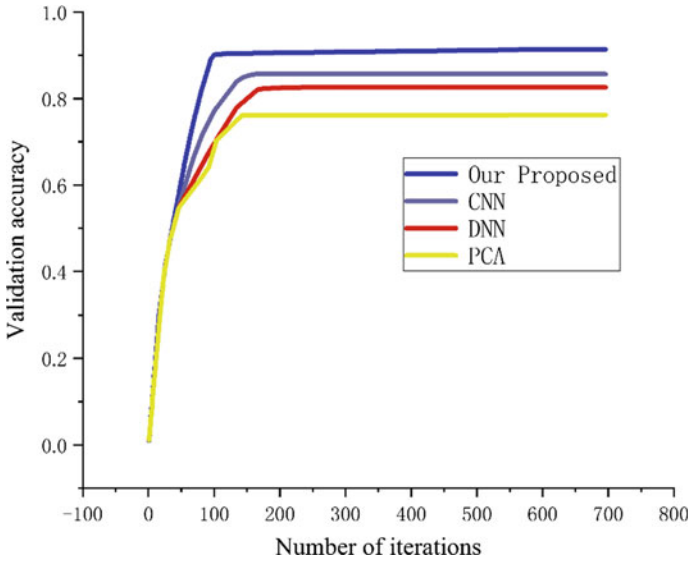


Fig. 6. The relationship between the number of iterations and the accuracy of four different algorithms for real-time face recognition

3 Testing Results

We have set up multiple cameras in the interior to recognize pedestrians from different angles. The camera captures pedestrian photos in real time, detects and corrects faces using the MTCNN model, and then uses these corrected faces as inputs sending to the trained FaceNet model, which make these inputs become a feature matrix with 128 eigenvalues. Finally, KNN training is performed on these feature matrices to achieve face recognition. To train the data set of KNN, we are using CASIA-FaceV5 [17] 4000 Asian face photos, which also includes 100 face photos of our team members. In the training process, the Intersection-over-Union (IoU) ratio of \hat{y} and y is 0.3 nonhuman face, 0.65 face, 0.4 face part, and 0.3 landmark, and training sample ratio is negative sample: positive sample: part sample: landmark = 3:1:1:2. We continue to appear in front of the camera, using the variable n to indicate the number of people identified and the variable m to indicate the number of times the team members were accurately identified. We also used some other algorithms to compare the accuracy of recognition at different angles. The following table is the final test data:

Method	Frame rate	Validation accuracy (%)
Ours	25 FPS	91.34
CNN	16 FPS	85.62
DL	24 FPS	82.57
PCA	27 FPS	76.15

4 Conclusion

On the other hand, the performance improvement of the training sample classifier is hindered, and an online sample for face detection and recognition needs to be designed to train the algorithm to achieve the adaptive training state. This paper proposed an MTCNN method to use the inherent correlation between detection and calibration to enhance their performance. In future, we plan to conduct some experiments further with feature selection, perhaps including more detailed representations of the closed captions. We plan also to investigate temporal-based features to automatically cluster images in adjacent frames. Other areas for further investigation include optimizing automatic selection of the number of clusters (choosing the correct k) by further refining MTCNN and k-means, especially in concert, as well as improving clustering accuracy by further tuning hyperparameters.

References

1. Turk, M.A., Pentland, A.P.: Face recognition using eigenfaces. In: Proceedings of 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Maui, HI, pp. 586–591 (1991)

2. Li, W., Li, M.M.: Research of realtime dynamic face recognition system based on flow compute model storm. In: 2016 International Symposium on Computer, Consumer and Control (IS3C), Xi'an, pp. 1002–1005 (2016)
3. Lawrence, S., Giles, C.L., Tsoi, A.C., Back, A.D.: Face recognition: a convolutional neural-network approach. *IEEE Trans. Neural Netw.* **8**(1), 98–113 (1997)
4. Razavian, A.S., Azizpour, H., Sullivan, J., Carlsson, S.: CNN features off-the-shelf: an astounding baseline for recognition. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops, Columbus, OH, pp. 512–519 (2014)
5. Chen, Y.-N., Han, C.-C., Wang, C.-T., Jeng, B.-S., Fan, K.-C.: The application of a convolution neural network on face and license plate detection. In: 18th International Conference on Pattern Recognition (ICPR'06), Hong Kong, pp. 552–555 (2006)
6. Yu, Z., Liu, F., Liao, R., Wang, Y., Feng, H., Zhu, X.: Improvement of face recognition algorithm based on neural network. In: 2018 10th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), Changsha, China, pp. 229–234 (2018)
7. Lee, K.-C., Ho, J., Kriegman, D.J.: Acquiring linear subspaces for face recognition under variable lighting. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(5), 684–698 (2005)
8. Kumar, N., Berg, A.C., Belhumeur, P.N., Nayar, S.K.: Attribute and simile classifiers for face verification. In: 2009 IEEE 12th International Conference on Computer Vision, Kyoto, pp. 365–372 (2009)
9. Schroff, F., Kalenichenko, D., Philbin, J.: FaceNet: a unified embedding for face recognition and clustering. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, pp. 815–823 (2015)
10. Parveen, P., Thuraisingham, B.: Face recognition using multiple classifiers. In: 2006 18th IEEE International Conference on Tools with Artificial Intelligence (ICTAI '06), Arlington, VA, pp. 179–186 (2006)
11. Hosang, J., Benenson, R., Schiele, B.: Learning non-maximum Suppression. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, pp. 6469–6477 (2017)
12. Sung, K.K., Poggio, T.: Example-based learning for view-based human face detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **20**(1), 39–51 (1998)
13. Lu, S.-X., Wang, X.-Z.: A comparison among four SVM classification methods: LSVM, NLSVM, SSVM and NSVM. In: Proceedings of 2004 International Conference on Machine Learning and Cybernetics (IEEE Cat. No. 04EX826), vol. 7, pp. 4277–4282 (2004)
14. Guillaumin, M., Verbeek, J., Schmid, C.: Is that you? Metric learning approaches for face identification. In: 2009 IEEE 12th International Conference on Computer Vision, Kyoto, pp. 498–505 (2009)
15. Perronnin, F., Liu, Y., Sánchez, J., Poirier, H.: Large-scale image retrieval with compressed Fisher vectors. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, pp. 3384–3391 (2010)
16. Ming, Z., Chazalon, J., Luqman, M.M., Visani, M., Burie, J.C.: Simple triplet loss based on intra/inter-class metric learning for face verification. In: 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), Venice, pp. 1656–1664 (2017)
17. Umer, S., Dhara, B.C., Chanda, B.: Biometric recognition system for challenging faces. In: 2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), Patna, pp. 1–4 (2015)



Based on the Predicted Blocking Virtual Machine Load Balancing Scheduling Strategy

Youhui Jiang^(✉)

College of Computer and Information Engineering, Tianjin Normal University,
Tianjin 300387, China
youhuijiang@mail.tjnu.edu.cn

Abstract. The popularity of the Internet and the rapid development of the massive data promote the expansion of the field of cloud computing applications, which absorb the grid computing, utility computing, and other distributed technology, through the network sharing platform resources to improve the throughput of data processing speed and computing power. As the core of cloud computing virtualization technology, through the construction of multiple virtualization platforms, in order to achieve space expansion, backup, and other operations to achieve the purpose of efficient integration of resources, but in practical applications, with the increase in the size of resources, physical and virtual machine load balance contradiction has become increasingly prominent, if such problems are not a reasonable solution, both a waste of resources, but also interfere with the application performance, thus affecting the user experience. In this paper, we study the process of blocking technology proposed based on load balancing under the conditions of virtual machine scheduling strategy, and through experiments to verify its reliability and efficiency.

Keywords: Cloud computing · Virtual machine scheduling · Congestion prediction · Load balancing

1 Introduction

Load balancing of cloud computing resources mainly exists in two phases of task allocation and application execution. This article mainly analyzes the application execution phase. When the virtual machine monitor detects an abnormal load, it adjusts the resource allocation strategy based on the load information and passes virtually [1]. The machine handles the reconfiguration of resources within the same node or the dynamic migration between different nodes; it makes the resources rebalanced again, which makes the application system environment more stable and improves the data center resource load situation [2].

Virtual machine scheduling strategy is a double-edged sword. The scheduling strategy is scientific and reasonable; it will balance the resource load. On the contrary, it will aggravate the imbalance of resources [3]. This is due to the degree of unbalanced node load relief and the virtual machine that is moved in and out. Released resources are related [4]. For example, if you move out of a virtual machine, you will release the resources of the physical node. If the resources released by the outgoing virtual

machine are few and cannot effectively resolve the load imbalance of the node, you need to migrate the virtual machine again. This will increase the system overhead [5]. Similarly, moving into a virtual machine will increase the resources of the physical node. If you increase the load of a moving node, such migration will be counterproductive and a waste of system resources. Therefore, a scientific and rational scheduling strategy must not only ensure the normal operation of the application, but also maintaining the load balance of the system and balancing the load of the system resources will not only significantly increase resource utilization, but also provide users with a continuous and stable working environment to maximize user satisfaction [6].

2 Load Forecasting Technology

The load forecasting technology predicts the load of the future period of time by using the historical data of the load status of the system in the past [7]. If the load forecasting result is too optimistic, the expected load is lower than the actual, and the idle time is longer, the system delay after the system restarts will be increased, not only affecting the performance but also increasing the energy consumption; if the load forecasting result is too pessimistic, the expected load is higher than actual; although it will not affect the system performance, it does not contribute to energy saving [8].

Load forecasting technology is mainly divided into two types: static prediction and adaptive. A fixed timeout strategy is a simple static prediction technique that customizes a length of time. When the component idle time exceeds the limit, the component transitions to a low-power state. Once a new request arrives, the component is reactivated [9]. The advantage of this strategy is its adaptability, which can be applied to any load working state. Its disadvantage performance will be affected when the component is reactivated.

3 Virtualization Technology

Cloud computing virtualization technology is an energy-saving technology that improves the utilization of server resources [10]. As shown in Fig. 1, on the hardware, multiple virtual machines (VMs) are created through physical machines to allow multiple virtual machines to share physics. Machine resources, in software, introduce the operating system into the abstraction layer between the hardware layer and the virtual hardware layer; independent hardware resources are provided for the virtual machine operating system [11]. The advantages brought by virtualization technology include running multiple operating systems on one physical machine to make full use of server resources, isolating system failures and security threats from the hardware level, and the dynamically allocating system resources such as CPU and memory among virtual machines. Machines are packaged into files that have nothing to do with hardware configuration, and snapshots of virtual machines can be made at any time. Online or offline migration of virtual machines can be performed with a simple file copy [8].

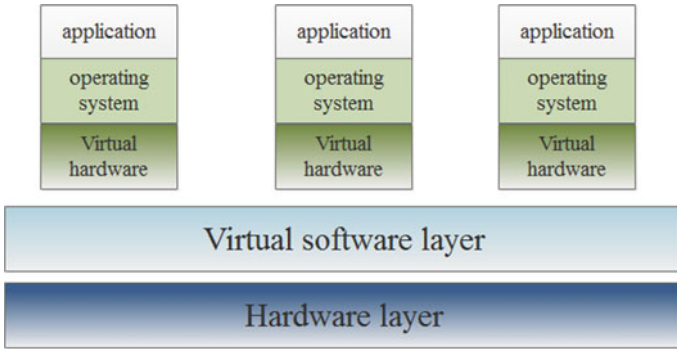


Fig. 1. Virtualization diagram

The virtual machine migration belongs to the data center-level virtualization technology [12]. Users apply for virtual machine resources through the data center scheduling system. The scheduling system allocates resources based on the system load and resource utilization [13]. Figure 2 shows a typical Infrastructure as a Service (IaaS), data center service model; scheduling system as a key part of the data center plays a decisive role in the quality of data center services, and unreasonable allocation of virtual machines will cause additional energy consumption of the data center. The next section will start from the scheduling algorithm and study virtual machine selection scheduling technology [14].

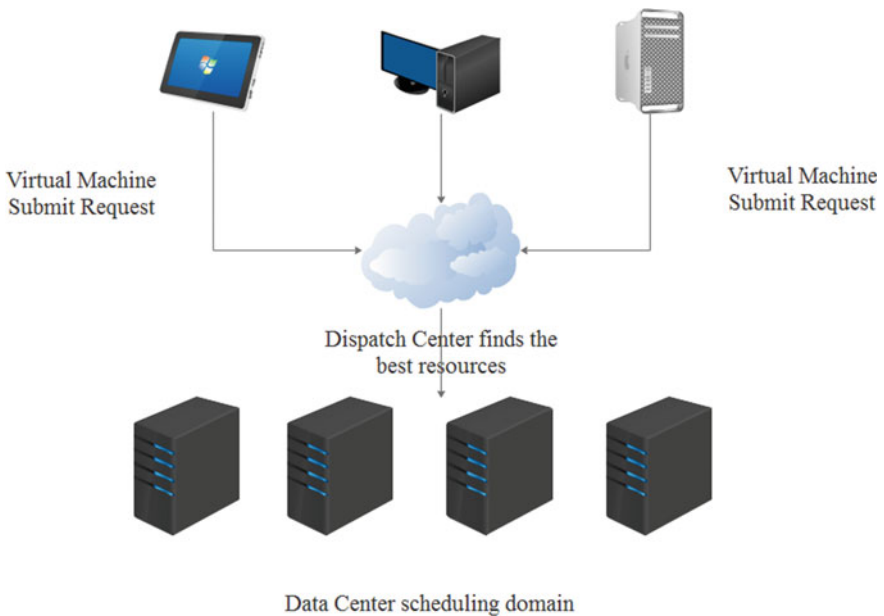


Fig. 2. Data center business process

4 Migration Virtual Machine Selection Algorithm

Process blocking is a passive behavior when waiting for resources. During task scheduling, the operating system will take care of the tasks that are in a blocked state. After blocking the tasks that are waiting for resources in the state, they can change to the ready state. Therefore, we guide the task scheduling by establishing a predictive process blocking model, which in turn affects the virtual machine migration strategy. We first set M equal-length time intervals. The prediction process acquires the resource time T_i ; the actual acquisition resource time is T_j ; each process acquires the same probability and obeys the binomial distribution.

$$\begin{aligned} P_k &= \frac{n_i!}{k!(n_i - k)!} P_{ij}^k (1 - P_{ij})^{n_i - k} \\ &= \binom{n_i}{k} p^k (1 - p)^{n - k} \end{aligned} \quad (1)$$

$k = 1, 2, 3 \dots; i, j = 1, 2, 3 \dots M$, P_k is the probability of acquiring k resources in n_i time T_j ; n_i is the number of resources acquired in T_i ; P_{ij} is the probability of acquiring resources when T_i is needed, and actually acquiring resources when T_i is used. The mathematical expectation and variance of Eq. 1 are expressed as

$$E_{ij}(X) = n_i P_{ij} \quad (2)$$

$$\sigma_{ij}^2(X) = n_i P_{ij} (1 - P_{ij}) \quad (3)$$

The probability distribution T_j of the actual number of acquired resources at P_{Nk} is equal to the cumulative probability distribution of the number of resources acquired at each interval in T_j .

$$\begin{aligned} P_{Nk} &= \frac{n_i!}{k!(n_i - k)!} P^k (1 - P)^{N - k} \\ &= \binom{n_i}{k} p^k (1 - p)^{n - k} \end{aligned} \quad (4)$$

And $k = 1, 2, 3 \dots N; N = \sum_{i=1}^M n_i$. The mathematical expectation and variance of Eq. 4 are expressed as

$$E(X) = \sum_{i=1}^M n_i P_{ij} \quad (5)$$

$$\sigma^2(X) = \sum_{i=1}^M n_i P_{ij} (1 - P_{ij}) \quad i, j = 1, 2, 3 \dots M \quad (6)$$

Assuming that the probability distribution of the actual number of acquired resources in the T_j -period satisfies a normal distribution, its probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (7)$$

In order to further verify the effectiveness of the load balancing strategy, we chose *OpenStack* to use an open source cloud computing platform based on the first experiment. We set up four virtual machines per physical machine, and we simulated the request spans of 150, 300, 600, 1200, and 2400 units in length, respectively. Taking the average, the algorithm strategy for comparison is to sort the requests in descending order of length and allocate virtual machine services (*Firstfit*) accordingly. As shown in Fig. 3, on the server total busy processing time index, the improved policy comparison in this paper compares the virtual machine scheduling policies in descending order of length and has better practicality.

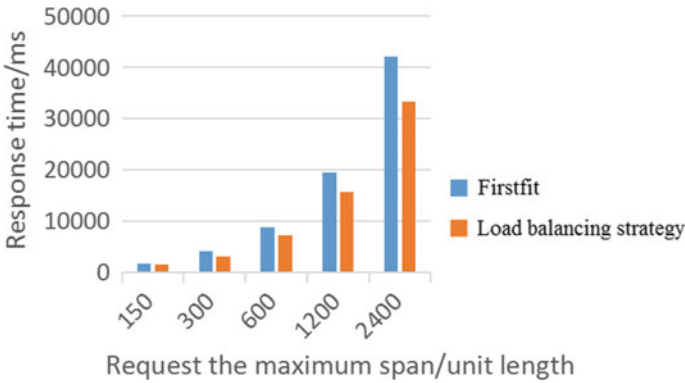


Fig. 3. (Improved) Server processing time comparison

5 Conclusion

This paper describes the background of the cloud computing environment, physical load balancing of the load balancing of physical machines, and virtual machines. Through research and establishment of blocking prediction technology model, a virtual machine scheduling strategy based on load balancing is proposed and their reliability is verified by experiments. The overall optimization of the application system makes the operating environment more stable, and thus effectively improves the data center resource load situation.

References

1. Sun, K., Wu, D.: MPC-based delay-aware fountain codes for live video streaming. In: IEEE International Conference on Communications, ICC (2016)
2. Schmidt, R.O.: Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Propag.* **34**(3), 276–280 (1986)
3. Sun, K., Zhang, H., Wu, D., Zhuang, H.: MPC-based delay-aware fountain codes for real-time video communication. *IEEE Internet Things J. (IoT)* (2016)
4. Tayem, N., Kwon, H.M.: L-shape 2-dimensional arrival angle estimation with propagator method. *IEEE Trans. Antennas Propag.* **53**(5), 1622–1630 (2005)
5. Liu, X., Li, Z.: Information-centric mobile ad hoc networks and content routing: a survey. Article in press (2016)
6. Gruteser, M., Grunwald, D.: Anonymous usage of location based services through spatial and temporal cloaking. In: ACM/USENIX MobiSys (2003)
7. Gu, J.F., Zhu, W.P., Swamy, M.N.S.: Joint 2-D DOA estimation via sparse l-shaped array. *IEEE Trans. Signal Process.* **31**(5), 1171–1182 (2015)
8. Sun, K., Wu, D.: Video rate control strategies for cloud gaming. *J. Vis. Commun. Image Represent.* (2013)
9. Li, Z., Chen, Y.T., Shi, H., Liu, K.: NDN-GSM-R: a novel high-speed railway communication system via named data networking. *EURASIP J. Wirel. Commun. Netw.* (2016)
10. Sun, K., Yan, B., Gharavi, H.: Low complexity content-aware image retar-geting. In: IEEE International Conference on Image Processing (ICIP '2012) (2012)
11. Ghinita, G., Kalnis, P., Skiadopoulos, S.: PRIVE: anonymous location based queries in distributed mobile systems. In: Proceedings of International Conference on World Wide Web (WWW '07), Banff, Alberta, Canada, pp. 1–10 (2007)
12. Li, Z., Liu, K.: MaPIT: an enhanced pending interest table for NDN with mapping bloom filter. *IEEE Commun. Lett.* **18**(11) (2014)
13. Xia, T.Q., Zheng, Y., Wan, Q., Wang, X.G.: 2-D angle of arrival estimation with two parallel uniform linear arrays for coherent signals. *IEEE Radar Conf.* **55**(9), 244–247 (2007)
14. Sun, K., Yan, B.: Efficient P-frame complexity estimation for frame layer rate control of H.264/AVC. In: IEEE International Conference on Image Processing (ICIP '2011) (2011)



Research on Fault Diagnosis Based on Artificial Neural Network

Rui Liu^(✉)

North Automatic Control Technology Institute, Tiyu Road no. 351,
Taiyuan 030006, China
liurui0822@sina.com

Abstract. According to classification framework of classical neural network, contemporary neural network, and soft computing, the basic concepts of the feedforward neural network (MLP, BP, and RBF), the feedback neural network (Hopfield, Boltzmann, Elman), the self-organizing neural network (SOM, ART, CPN), deep and extreme neural network (Deep learning, extreme learning), the novel neural network (SVM, PNN), and soft computing neural networks combined with various methods are introduced, and the research progress and typical applications of the neural network in fault diagnosis are given. The existing problems and future development directions of fault diagnosis are also discussed.

Keyword: ANN · BP · Soft computing · DNN · Fault diagnosis

1 Introduction

Generally speaking, fault diagnosis includes three aspects: fault detection, fault identification, and fault location. Fault detection is to gather state information of the acquisition equipment, and the state information of the equipment can be collected by self-detection (such as BIT [1]) or external equipment (such as ATE) or the acquisition of sensor devices (such as sound, light, and electrical sensors). The fault identification is judged by threshold judgment or parameter matching. Fault location is based on the detection data and failure mode to determine the location and type of the fault. At present, the methods of fault diagnosis mainly include artificial neural network, fuzzy set and fuzzy reasoning, expert system, and evolutionary computation, and contain model- and data-driven method. Artificial neural network has strong nonlinear function approximation ability, adaptive learning ability, fault tolerance ability, and parallel information processing ability. It is an effective way to solve the fault diagnosis problem in nonlinear systems. It has been widely used in rotating system, valve, generator, bearing and manufacturing system of electronic and mechanical equipment [2, 3].

2 Classical Neural Networks

2.1 Feedforward Neural Network

Multilayer perceptron (MLP) is the basic artificial neural network topology. In an MLP network for fault diagnosis, the input layer is usually acquired as a device state signal variable containing the equipment health information and fault information, and the output layer outputs the related fault judgment information of the input variables. The purpose of the MLP network is to minimize the difference between the output value and the expected value, usually by training the hidden layer. The parameters are used to find the optimal solution. MLP network is widely used in fault diagnosis [4].

BP algorithm is the most commonly used algorithm for training MLP network. It belongs to the error multilayer feedforward network. The BP neural network has the advantages of simple network structure, mature algorithm, fast calculation speed, easy simulation implementation, and easy to change. The standard BP algorithm often converges slowly. The weight value correction algorithm of the momentum term is introduced to improve the operation efficiency. Ding et al. [5] improve the convergence of the BP algorithm by increasing the input nodes of the BP neural network. Adedigba et al. [6] introduce the fault tree and event tree method to BP to solve the problem of low efficiency in the complex nonlinear relation. Ma et al. [7] quantized BP neural network to solve the problem of equipment fault diagnosis with qualitative data.

RBF is a novel ANN model. It is a supervised learning multilayer feedforward network. Because of the simple network structure and good convergence of the algorithm, the learning algorithm of RBF network is more complex than BP network, but the training speed is greatly improved. Lei et al. [8] use RBF network to extract the fault features and diagnose the mechanical faults. Compared with MLP network, RBF network has a better performance in recognition accuracy and robustness [9].

2.2 Feedback Neural Network

Hopfield neural network was first proposed by Professor Hopfield of in the United States in 1982. Davey et al. [10] realize fault pattern recognition in dynamic and stable environment based on Hopfield networks. Li et al. [11] introduce wavelet packet analysis and energy calculation as fault feature preprocessing in order to improve the classification efficiency. Compared to the feedforward neural network, the Hopfield neural network considers the lag effect of input and output in time. Feedforward neural networks usually use error correction method, which is slow in computation and slow in convergence. Hopfield network and its associative memory characteristics have good effect in fault diagnosis of incomplete symptoms.

Boltzmann machine is a stochastic recurrent neural network proposed by Hinton in 1985. DRBM can be used as an independent classifier; DDRBM-based classification network can solve multiple classification problems and has better classification performance. The research shows that DRBM has strong ability to extract sample features. Gaussian restricted Boltzmann machine has the ability to provide a closed-form representation of the distribution underlying the training data [12]; He et al. [13] proposed

a novel bearing fault diagnosis method based on Gaussian RBM, and the proposed method can realize the bearing fault diagnosis accurately and effectively.

Elman network can be used to solve the problem of multiple fault identification in nonlinear systems [14]. Combined with rough set and genetic algorithm, the average fault diagnosis accuracy of fault diagnosis system has been greatly improved [15]. Jun and Huang [16] propose an improved particle swarm optimization Elman neural network for aero-engine fault diagnosis.

2.3 Self-organizing Neural Networks

SOM network can be effectively used for fault diagnosis in the case of supervised learning. SOM can be correctly identified with the input signal of the relative decimal order of magnitude, but BP cannot [17]. Because of its competitive learning mechanism, SOM has a significant advantage in dealing with concurrent failures. The SOM network can be used to separate them one by one and return to the corresponding characteristic group [18]. SOM can extract fault features well in offline learning and is suitable for nearest neighbor classifier [19]. SOM is a class of suboptimal model classifier, and its generalization performance is not as good as that of supervised learning model, but it has self-organization and self-adaptive characteristics, and it can better coordinate the requirements of adaptability, stability, and complexity.

ART is widely used in fault pattern recognition. ART can solve the problem of stability and plasticity of neural network learning and can be applied to the fault diagnosis in the manufacturing process [20]. Compared to fault detection and recognition based on rule, ART has the same performance as BP network, but it is easier to implement [21]. Compared with rule-based methods, ART can effectively identify known faults and unknown faults, and only needs little model input. But the ART network depends on the adjustment of the alert parameters. When the alert parameters are high, the ART network is more inclined to produce a new neuron group, but it is easier to make mistakes in the recognition of the unknown fault.

Counter propagation neural network can overcome the shortcomings of BP network which is easy to fall into local minima, slow learning speed, and poor convergence. With advantages of fast training speed and high diagnostic rate, CPN can simulate the fault of analog circuits and can accurately determine the fault mode [22]. Some scholars also set up a fault diagnosis model based on rough sets and colored Petri net and CPN, to improve the generality and reduce the difficulty of diagnosis for complex faults. The combination of spectrum analysis and counter propagation neural network is simple and effective for fault diagnosis and has good robustness and generalization ability.

3 Contemporary Neural Networks

3.1 Deep and Extreme Learning

Deep neural network has been widely used in the fields of pattern recognition and data classification [23]. Deep learning network has a more complex network structure, and the fault feature information is extracted by a small error through a variety of nonlinear

transformations and nonlinear approximations of the unprocessed fault data. Tran takes the deep learning network as the classifier of the input fault data to extract the fault information of the reciprocating compressor. Feng uses unsupervised depth learning network for fault feature mining after training and use supervised learning network for fault classification. This method is suitable for solving the nonlinear matching problems of various health states and corresponding measurement signals [24]. Unsupervised learning of deep neural network also can be used as an automatic encoder for data dimensionality reduction or fault feature extraction [25]. DNN also used in fault classification and diagnosis of the complex electronic information system.

The traditional ELM has the problems of low classification precision and poor network structure stability. Some scholars proposed adaptive chaotic particle swarm optimization algorithm to optimize the ELM parameters. The algorithm enhances the stability of the network and improves the accuracy of data classification. Liu et al. [26] proposed a new method of combining the local mean decomposition (LMD) multi-scale entropy and the ELM. In view of the slow diagnosis speed and difficult parameter selection, Wang et al. [27] introduce ELM to the fault diagnosis of the aeronautical rotating rectifier. Compared with the fault diagnosis method, ELM has better diagnostic efficiency.

3.2 Novel Neural Network

Pulse coincidence neural network model (PCNN) is proposed by Eckhom on the basis of the working principle of the biological vision neural system. On the basis of traditional PCNN, Johnson transforms the two-dimensional image into one-dimensional time series signal equation through PCNN iteration [28]. In order to solve the problem of complex PCNN structure and inconvenient application, Zhang proposed a simplified PCNN model and use PCNN to extract the fault feature of the bearing [29]. At present, fault diagnosis is only used for fault feature extraction, which needs further study.

Support vector machine (SVM) is an artificial neural network based on structural risk minimization criterion [30]. Through function mapping, the low-dimensional and non-separable problems in linear space are converted into a linear problem with high latitudes, and the optimal hyperplane is constructed. SVM has the characteristics of small sample dependence, unique solution, and global optimization. It can effectively solve multi-class classification problems.

4 Soft Computing

Zadeh takes the lead in proposing the concept of soft computing. It is a combination of fuzzy logic, neural computing, evolutionary computation, and ordinal optimization. Soft computing can deal with imprecise and uncertain information and has the ability of perceiving, reasoning, and learning in the environment. Because of its advantages of easy handling and robustness, the ensemble of neural network and other methods are widely used in the field of fault diagnosis [31].

Literature [32] combines fuzzy neural network with rough set, wavelet transform, and least square weighted fusion algorithm. Literature [33] combines the neural

network with SVM implements for fault classification. Literature [34] gives two ways of combining neural networks with expert systems. The wavelet neural network combines the good time–frequency localization characteristic of wavelet transform and the self-learning function of neural network. Fuzzy neural network for fault detection and classification is proposed [35]. Neural network is also combined with genetic algorithm and simulated annealing algorithm. Steve [36] applies genetic algorithm to network weight training of artificial neural network. Some scholars also introduce simulated annealing algorithm to Boltzmann machine, which makes it possible to classify and learn algorithms and expands the application prospect of artificial neural network in fault diagnosis.

5 Conclusion and Prospect

5.1 Conclusion

This paper gives a comprehensive overview of the development of artificial neural networks for fault diagnosis. Neural network theory provides a new idea and method for fault detection and diagnosis of complex nonlinear dynamic systems. Neural networks used for system fault diagnosis are mainly used in the following aspects: first, the neural network is used to identify specific fault modes; second, the neural network is used as the state estimator for fault diagnosis by using the idea of analytic redundancy; thirdly, using artificial neural network to produce or evaluate the residual; and fourthly, the nonlinear error of observer can be compensated by neural network with adaptive error compensation.

5.2 Prospects

Only one kind of theory or one method makes it difficult to realize the fault diagnosis absolutely, accurately, and timely in complex condition, but the comprehensive application of various methods can take into account the advantages of many methods and can improve the real-time and accuracy of fault diagnosis. Therefore, the organic integration and integration of various methods will become the main direction of the development of online fault diagnosis technology.

Mixed fault diagnosis of complex systems. Complex systems often consist of devices at different levels and have diverse functions and complicated information interaction. System complex composition, complex fault-related relationship, and complex time series on information flow bring huge challenges to fault diagnosis and recognition. However, how to ensure that this information can be effectively used to improve the accuracy of fault diagnosis and reduce false alarm rate, and failure should be further studied.

Precision and efficiency have always been the goal of the algorithm. With the emergence need of remote fault diagnosis, distributed fault diagnosis, embedded real-time fault diagnosis, and portable fault diagnosis, the real-time, diagnostic accuracy and calculation efficiency of the artificial neural network algorithm under limited computing resources is a major problem for the current fault diagnosis. Therefore, it is an

important research direction to improve the structure and performance of the neural network and improve the structure and performance of the neural network on the basis of the high-speed, parallel, and exponential capacity of the computer or chip.

Compared to fault diagnosis, fault prediction is more challenging. However, the local minima and overlearning problems of neural networks lead to low prediction accuracy and are applied only in the fields of battery and mechanical equipment. Therefore, constructing artificial neural network theory and engineering application method suitable for fault prediction is also an important research direction in the future.

Acknowledgements. The 207th Research Institute of NORINCO GROUP, Project 61603006 Supported by NSFC.

References

1. Liu, R.: Design for distributed electronic system BIT. In: SSME 2016, pp. 715–720 (2016)
2. Subbaraj, P., Kannapiran, P.: Artificial neural network approach for fault detection in pneumatic valve in cooler water spray system. *Int. J. Comput. Appl.* **9**(7), 43–52 (2010)
3. Demetgul, M., Unal, M., Tansel, I.N., Yazcioglu, O.: Fault diagnosis on bottle filling plant using genetic-based neural network. *Adv. Eng. Softw.* **42**(12), 1051–1058 (2011)
4. Lei, Y., Li, N., Lin, J., He, Z.: Two new features for condition monitoring and fault diagnosis of planetary gearboxes. *J. Vib. Control* (2013). <https://doi.org/10.1177/1077546313486284>
5. Ding, X.Q., Sun, J., Yuan, Y.B.: BP neural network-based fault diagnosis methods improvement. *Power Grid Technol.* **22**(11), 62–63 (1998)
6. Adedigba, S.A., Khan, F., Yang, M.: Dynamic failure analysis of process systems using neural network. *Process Safety Environ. Prot.* 529–543 (2017)
7. Ma, C.G., Li, Q., Sun, D.L., et al.: research on intelligent fault diagnosis method for a certain equipment based on neural network. *Ship Electron. Eng.* **37**(10), 104–107 (2017)
8. Lei, Y.G., He, Z.J., Yan, Y., Hu, Q.: Mechanical fault diagnosis model based on feature evaluation and neural network. *Acad. J. Xi'an Jiaotong Univ.* **40**(5), 558–56 (2006)
9. Leger, R.P., Garland, W.J., Poehlman, W.F.S.: Fault detection and diagnosis using statistical control charts and artificial neural networks. *Artif. Intell. Eng.* 35–47 (1998)
10. Davey, N., Hunt, S.P., Adams, R.G.: High capacity recurrent associative memories. *Neurocomputing* **62**, 459–491 (2004)
11. Li, P.H., Jiang, J.C., Qiu, B.M., et al.: Hopfield neural network fault diagnosis method for analog circuits under wavelet energy description. *J. Chongqing Univ.* **37**(7), 136–146 (2014)
12. Fischer, A., Igel, C.: Training restricted Boltzmann machines: an introduction. *Pattern Recogn.* **47**(1), 25–39 (2014)
13. He, X.H., Wang, D., Li, Y.F.: A Novel bearing fault diagnosis method based on Gaussian restricted Boltzmann machine. *Math. Probl. Eng.* 1–9 (2016)
14. Yang, C., Wang, Z.W.: Fault diagnosis method of rolling bearings based on Elman neural network. *Bearing* **5**, 49–52 (2010)
15. Yang, R., Cheng, H., Gui, L.: airborne stabilized platform servo system fault analysis and diagnosis method. *Electron. Optics Control* **24**(2), 64–68 (2017)
16. Jun, P.I., Huang J B, Fault diagnosis of aero engine based on IPSO-Elman neural network. *J. Aerosp. Power* **32**(12), 3031–3038
17. Shen, Y.X., Ji, Z.C., Jiang, J.G.: Review of artificial intelligence method for fault diagnosis of motor. *Micro motor* (2), 39–42 (2004)

18. Penman, J., Yin, C.M.: Feasibility of using unsupervised learning, artificial neural networks for the condition monitoring of electrical machines. *IEE Proc. B* **141**, 317–322 (1994)
19. Wang, H., Zheng, H., Azuaje, F.: Self-adaptive neural networks based on a Poisson approach for knowledge discovery. In: *Proceedings of the IJCAI 2007*, pp. 1101–1106 (2007)
20. Demetgul, M., Tansel, I.N., Taskin, S.: Fault diagnosis of pneumatic systems with artificial neural network algorithms. *Expert Syst. Appl.* **36**(7), 10512–10519 (2009)
21. Fernando, H., Surgenor, B.: An unsupervised artificial neural network versus a rule-based approach for fault detection and identification in an automated assembly machine. *Robot. Comput.-Integr. Manuf.* **43**, 79–88 (2017)
22. Guan, C.B., Li, Q.Y., Yang, L.J.: Application of CPN neural network in analog circuit fault diagnosis. *Ship Electron. Eng.* **28**(55), 5–8 (2008)
23. Schmidhuber, J.: Deep learning in neural networks: an overview. *Neural Netw.* **61**, 85–117 (2014)
24. Jia, F., Lein, Y.G., Lin, J., et al.: Deep neural networks: a promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mech. Syst. Signal Process.* **72–73**, 303–315 (2016)
25. Bengio, Y., Courville, A., Vincent, P.: Representation learning: a review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**, 1798–1828 (2013)
26. Liu, M.R., Li, Z., He, Y.G., et al.: Fault diagnosis of analog circuits based on LMD multiscale entropy and extreme learning machine. *J. Electron. Measurement Instrument.* **31** (4), 530–536 (2017)
27. Wang, X.Y., et al.: Research on fault diagnosis technology of aeronautical rotating rectifier based on extreme learning machine. *Electr. Autom.* 219–222
28. Johnson, J.L., Padgett, M.L.: PCNN models and applications. *IEEE Trans. Neural Netw.* **10** (3), 480–498 (1999)
29. Zhang, Y.Q., Zhang, P.L., Ren, G.Q.: Application of simplified PCNN in color feature extraction for wear particle images. *Chin. Internal Combust. Engine Eng.* **34**(5), 69–75 (2013)
30. Nieto, P.J.G., García-G, E., Lasheras, F.S., Juez, F.J.D.C.: Hybrid PSO-SVM based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability. *Reliab. Eng. Syst. Safety* **138**, 219–231 (2015)
31. Sakthivel, N.R., Binoy, B., Nair, V., Sugumaran: Soft computing approach to fault diagnosis of centrifugal pump. *Appl. Soft Comput.* **12**, 1574–1581 (2012)
32. Dong, L.X., Xiao, D.M., Liang, Y.S., Liu, Y.L.: Rough set and fuzzy wavelet neural network integrated with least square weighted fusion algorithm based fault diagnosis research for power transformers. *Electr. Power Syst. Res.* **78**(1), 129–136 (2008)
33. Thukaram, D., Khincha, H.P., Vijaynarasimha, H.P.: Artificial neural network and support vector machine approach for locating faults in radial distribution systems. *IEEE Trans. Power Deliv.* **20**(2), 710–721 (2005)
34. Powell, M.J.: Radial basis functions for multivariable interpolation: a review. In: *Algorithms for Approximation*. Clarendon Press, pp. 143–167 (1987)
35. Quteishat, A., Lim, C.P.: A modified fuzzy min-max neural network with rule extraction and its application to fault detection and classification. *Appl. Soft Comput.* **8**(2), 985–995 (2008)
36. Billings, S.A., Zheng, G.L.: Radial basis function network configuration using genetic algorithms. *Neural Netw.* **8**(6), 877–890 (1995)



Virtual Samples for Cloud Classification via Supervised Learning

Shuang Liu^{1,2}, Mei Li^{1,2}, Zhong Zhang^{1,2}, Mingzhu Shi^{1,2}, and Xiaozhong Cao³(✉)

¹ Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin, China

{shuangliu.tjnu, limeitjnu, zhong.zhang8848}@gmail.com, shimingzhu1@163.com

² College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin, China

³ Meteorological Observation Centre, China Meteorological Administration, Beijing, China

xzhongcao@163.com

Abstract. Convolutional neural networks (CNNs) have been widely used in image classification task, which is based on the huge amount of image samples. However, the insufficiency of cloud sample numbers brings obstacles to classify clouds using CNNs. In this paper, we propose to apply Wasserstein generative adversarial network (WGAN) to generate virtual cloud samples via supervised learning. Afterward, we fine-tune a deep CNN model to evaluate the classification performance under different number of virtual cloud samples. The experimental results demonstrate the feasibility of the proposed method.

Keywords: Convolutional neural networks · Generative model · Discriminative model · Cloud classification

1 Introduction

Clouds are closely related to the daily activities of the human being and accurate cloud classification has always been pursued in both the research field and practical application. A number of researchers adopt various methods for automatic cloud classification. Liu et al. [1] regarded the cloud sequences as dynamic texture and described the cloud in a tensor manner, and also proposed the super-pixel segmentation algorithm to obtain superpixels could adaptively [2]. Heinle et al. [3] presented to utilize statical features, such as color and texture, to describe clouds. Buch et al. [4] used binary trees to distinguish the cloud between clear sky and other cloud types. Kazantzidis et al. [5] applied multifeatures for cloud classification, including statistical color, textural features, solar zenith angle, cloud coverage as well as the existence of raindrops in cloud images. Recently, convolutional neural networks (CNNs) have shown an outstanding performance

in image classification and identification tasks [6–9]. Inspired by that, many researchers [10–13] applied CNNs to extract deep CNN-based features for automatic cloud classification and the promising results validate the great potential of CNNs for cloud classification.

However, the scarcity of the cloud sample quantities has been a bottleneck for the further development of the cloud classification. The problem is that superior properties cannot be learned for classifiers by using a limited number of cloud samples. Recently, generative adversarial network (GAN) [14] and its multiple extensions [15, 16] have brought significant impact in the research field, which could generate a substantial amount of images. However, the existing generating methods are based on unsupervised learning and have difficulty in labeling the generated samples.

In this paper, we aim to classify cloud images by utilizing virtual cloud samples which are generated via supervised learning. The rest of the paper is organized as follows. Section 2 introduces the virtual sample generating model, the cloud classification network, and the implementation details. Section 3 presents the experimental results and Sect. 4 makes a conclusion for this paper.

2 Generating Virtual Samples via Supervised Learning

In this section, we first introduce the Wasserstein generative adversarial network (WGAN) [17] that is used to generate the virtual cloud samples. Then, we introduce the classification network, and finally, we present the implementation details.

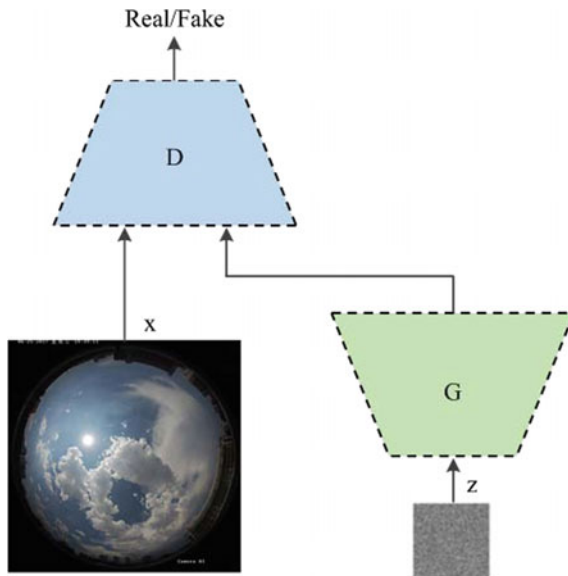


Fig. 1. Illustration of the process flowchart for G and D

2.1 WGAN

The WGAN consists of a generative model (G) and a discriminative model (D). G is used to generate fake images and D is responsible for discriminating whether the input is real or not. Figure 1 illustrates the process flowchart of G and D . Concretely, G takes a n -dimensional noise z as input and outputs an image $G(z)$. D takes an image $x/G(z)$ as input and outputs a probability value $D(x)/D(G(z))$.

G contains five deconvolutional layers and the kernel size for each layer is 4×4 . A batch normalization layer and a ReLU layer follow the first four deconvolutional layers, respectively. The tanh function is used on the output of the last deconvolutional layer. D contains five convolutional layers and the kernel size for each layer is 4×4 . The output of the first convolutional layer is connected to the Leaky ReLU. The next three convolutional layers are followed by a batch normalization layer and a Leaky ReLU layer, respectively.

2.2 Classification Network

We modify the ResNet-50 [8] as the classification network. The modifications include several aspects. First, we replace the last average pooling layer with an adaptive average pooling layer which is simply used as the transition step to connect the convolution layer to the fully connected (FC) layer. Second, we remove the last FC layer and add two new ones and the second one is with the number of N neurons. Herein, N is the number of cloud classes. In addition, between the two added FC layers, we also add a batch normalization layer, a Leaky ReLU layer and a dropout layer sequentially. Note that, the softmax function is applied to the outputs of the last FC layer and the cross-entropy loss is utilized to measure the performance of the classification network.

2.3 Implementation Details

In this subsection, we first introduce the implementation details of G and D . Then, we present the implementation details of the classification network.

We aim to generate virtual cloud images for each class via supervised learning. That is to say, we only use one class of the cloud samples to train G and D at each time. When we obtain the trained G , we use it to generate the virtual cloud images. However, the number of real cloud images in each class is 2000, which runs the risk of overfitting to train G and D . To overcome this problem, we first resize the images to 120×120 , then we crop the images twice. In the first time, the images are centrally cropped to 80×80 and in the second time, the randomly cropped images are with the size of 64×64 . We also apply the random horizontal flip strategy after each cropping process. Meanwhile, each image is normalized by Gaussian distribution with the RGB mean values and standard deviation values computed on a specific class before it is fed into D . The weights in deconvolutional layers, convolutional layers, and batch normalization layers are subjected to the uniform distribution, and the bias in batch normalization layers

is initialized to 0. The dimensions of noise z are fixed to 100, and the noise obeys to the standard normal distribution. The coefficient of leakage in Leaky ReLU is set to 0.2.

The purpose of G is to generate images as real as possible. While the purpose of D is to distinguish the fake images from the real images to the best of its ability. The training process of G and D is a two-player minimax game. Specifically, we need to maximize the loss of D and minimize the loss of G at the same time. We adopt the RMSProp to optimize the training process. The training epoch is set to 100. The learning rate and weight decay are both set to 0.00005. Meanwhile, during each iteration, the parameters in D are clamped to the range of -0.01 to 0.01 . To improve the efficiency of the training process, the G starts to be trained after every 100 iterations of D .

For the training phase of the classification network, we fine-tune the parameters of ResNet-50 which is pretrained on the ImageNet. The parameters (weights and bias) in the two added FC layers are subjected to the uniform distribution. The weights and bias in the added batch normalization layer are initialized to 1 and 0, respectively. The cloud images in the training set are first resized to 300×300 and then centrally cropped to 256×256 . Afterward, they are normalized by Gaussian distribution with the RGB mean values and standard deviation values calculated on all the classes in the training set. The stochastic gradient descent (SGD) [6] with momentum is adopted to update the parameters. The momentum value is set to 0.9. The training epoch is set to 50 with the batch size of 32. The learning rate is set to 0.001, except the added FC layers which are set to 0.01. The coefficient of leakage in the added Leaky ReLU is set to 0.1 and the drop rate in the added dropout layer is set to 0.5. It should be noticed that the learning is reduced to one-tenth of that after 30 epochs.

During the test process, the cloud images in the test set are resized to 256×256 . The other preprocessing operations are in accordance with those in the training stage. Afterward, they are fed to the well-trained classification network to obtain the classification accuracies.

3 Experimental Results

The cloud samples are taken by a sky camera with fisheye lens. According to the International cloud classification system criteria published in the World Meteorological Organization (WMO), and the visual similarity in practice, we classify the cloud samples into seven classes, i.e., cumulus, cirrus, and cirrostratus, cirrocumulus and altocumulus, clear sky, stratocumulus, stratus, and altostratus, cumulonimbus, and nimbostratus. The sample number in each class is 2000. Hence, there are 14000 real cloud samples in the cloud dataset. We randomly select 1300 samples from each class as training set and the remaining 700 samples in each class as the test set. We use the generative model (G) to generate 2000 virtual cloud samples for each class. Figure 2 shows the real and virtual cloud samples in each class. From the figure, we can see that the generated images possess many properties of the corresponding class. We equally partition the virtual

cloud samples into four parts for each class. We first train the classification network without the virtual cloud samples. Then, we merge one part of the virtual cloud samples with the real ones in the training set and shuffle them to train the classification network. The emergence process is conducted four times and the test set is used to evaluate the classification performance of each well-trained network.

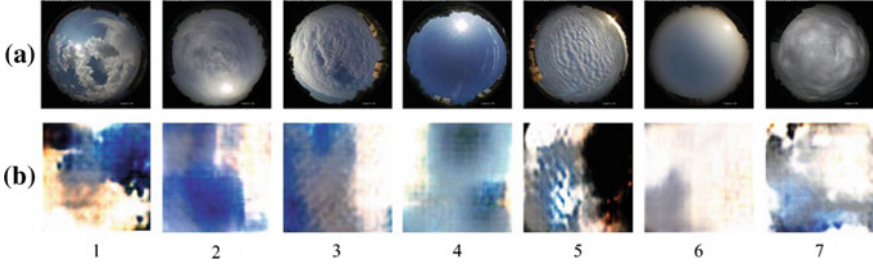


Fig. 2. Cloud samples of the real images and virtual images in each class. Set **a** represents the real cloud samples and set **b** represents the virtual cloud samples. The Arabic numbers in the bottom denote the corresponding label of the cloud

Table 1. Classification accuracies (%) using different number of virtual cloud samples in each class

Number	Accuracy (%)
0	89.63
500	89.90
1000	86.92
1500	88.06
2000	89.35

The classification results with different virtual cloud sample numbers in each class are listed in Table 1. From the table, we can see that compared with the classification results with no virtual cloud samples, the classification results with virtual cloud samples are competitive. Especially, the classification accuracy achieves to 89.90% with 500 virtual cloud samples in each class. The results demonstrate that it is feasible to generate virtual cloud samples under supervised learning to solve the sample scarcity problem.

4 Conclusion

In this paper, we have proposed to use virtual samples for cloud classification. The virtual cloud samples are generated by WGAN via supervised learning. We have evaluated the classification performance under different number of virtual

cloud samples in each class and the experimental results have validated the effectiveness of the proposed method.

Acknowledgements. This work was supported by National Natural Science Foundation of China under Grant No. 61501327, No. 61711530240, and No. 61501328, Natural Science Foundation of Tianjin under Grant No. 17JCZDJC30600 and No. 15JCQNJC01700, the Fund of Tianjin Normal University under Grant No.135202RC1703, the Open Projects Program of National Laboratory of Pattern Recognition under Grant No. 201700001 and No. 201800002, the China Scholarship Council No. 201708120039 and No. 201708120040, and the Tianjin Higher Education Creative Team Funds Program.

References

1. Liu, S., Wang, C., Xiao, B., Zhang, Z., Cao, X.: Tensor ensemble of ground-based cloud sequences: its modeling, classification, and synthesis. *IEEE Geosci. Remote S.* **10**, 1190–1194 (2013)
2. Liu, S., Zhang, L., Zhang, Z., Wang, C., Xiao, B.: Automatic cloud detection for all-sky images using superpixel segmentation. *IEEE Geosci. Remote S.* **12**, 354–358 (2015)
3. Heinle, A., Macke, A., Srivastav, A.: Automatic cloud classification of whole sky images. *Atmos. Meas. Tech.* **3**, 557–567 (2010)
4. Buch, K.A., Sun, C.H., Thorne, L.R.: Cloud classification using whole-sky imager data. In: 5th Atmospheric Radiation Measurement Science Team Meeting, pp. 19–23. San Diego, CA, USA (1995)
5. Kazantzidis, A., Tzoumanikas, P., Bais, A.F., Fotopoulos, S., Economou, G.: Cloud detection and classification with the use of whole-sky ground-based images. *Atmos. Res.* **3**, 557–567 (2012)
6. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems*, pp. 1097–1105. Lake Tahoe, Nevada, USA (2012)
7. Marmanis, D., Datcu, M., Esch, T., Stilla, U.: Deep learning earth observation classification using ImageNet pretrained networks. *IEEE Geosci. Remote S.* **3**, 557–567 (2016)
8. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778. Las Vegas, Nevada (2016)
9. Zheng, Z., Zheng, L., Yang, Y.: A discriminatively learned CNN embedding for person reidentification. *ACM T. Multim. Comput.* **14**, 13:1–13:20 (2017)
10. Ye, L., Cao, Z., Xiao, Y.: DeepCloud: ground-based cloud image categorization using deep convolutional features. *IEEE T. Geosci. Remote* **55**, 5729–5740 (2017)
11. Shi, C., Wang, C., Wang, Y., Xiao, B.: Deep convolutional activations-based features for ground-based cloud classification. *IEEE Geosci. Remote S.* **14**, 816–820 (2017)
12. Zhang, Z., Li, D., Liu, S., Xiao, B., Cao, X.: Cross-domain ground-based cloud classification based on transfer of local features and discriminative metric learning. *Remote Sens.* **10**, 8 (2017)
13. Liu, S., Li, M.: Deep multimodal fusion for ground-based cloud classification in weather station networks. *Eurasip J. Wirel. Comm.* **1**, 48 (2018)

14. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: *Advances in Neural Information Processing Systems*, pp. 2672–2680. Montreal, Canada (2014)
15. Mirza, M., Osindero, S.: Conditional generative adversarial nets. [arXiv:1411.1784](https://arxiv.org/abs/1411.1784) (2014)
16. Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. [arXiv:1511.06434](https://arxiv.org/abs/1511.06434) (2015)
17. Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein GAN. [arXiv:1701.07875](https://arxiv.org/abs/1701.07875) (2017)



Deep Learning for Optical Character Recognition and Its Application to VAT Invoice Recognition

Yu Wang, Guan Gui^(✉), Nan Zhao, Yue Yin, Hao Huang, Yunyi Li,
Jie Wang, Jie Yang, and Haijun Zhang

College of Telecommunication and Information Engineering, Nanjing University
of Posts and Telecommunications, Nanjing 21003, China
guiguan@njupt.edu.cn

Abstract. Optical character recognition (OCR) is considered as one of long-term and hot research topics due to the fact that OCR technique can change the documents from paper to computer-readable format by consistently growing. However, the recognition accuracy of current OCR technique is required to improve some special applications such as in reimbursement of value-added tax (VAT) invoices. This paper proposes two OCR techniques by using deep convolutional neural network (CNN) and residual network (ResNet), respectively. According to our test dataset, the formerly proposed techniques can reach up to 97.08%, while the latter can increase to 99.38%.

Keywords: Optical character recognition · Value-added tax invoices · Deep learning · Convolutional neural network · Residuals network

1 Introduction

Optical character recognition (OCR) refers to the process of recognition text data from image containing text that is not editable by computer [1]. It generally has four steps: digitization, preprocessing, segmentation, and feature extraction [1].

- Step 1. Digitization is the first step of OCR, and we adopt camera or scanner to convert paper document containing handwriting or machine text to digital images, which will be used in the next steps of preprocessing.
- Step 2. Preprocessing generally includes binarization, tilt correction, and many more. Binarization is the first step of preprocessing to convert color images to black and white images, which is aimed at decreasing the amount of information and speeding up recognition. Tilt correction is correcting text tilt due to shooting or scanning before segmentation and recognition [2].
- Step 3. Segmentation is splitting character sequence into individual characters; it is an essential step in OCR, the effect of which directly affects the final result of OCR [2].

Step 4. Feature extraction is the extraction of single-character feature that distinguishes themselves from other characters, and it is the most essential step of OCR.

OCR has many engineering applications, such as license plate recognition and business card recognition. We also want to introduce OCR to reimburse value-added tax (VAT) invoices, due to that VAT is the most important tax in China, which is a tax levied on the added value of product or service at every stage of production or distribution, but we generally reimburse VAT invoices manually, which has many disadvantages, such as wasting time, being inefficient, and high error rate. However, traditional feature extraction methods, such as pattern matching [3] and structural analysis [4], are not very effective, especially in face of Chinese characters and reimbursement invoices require more accurate OCR that traditional OCR cannot do. Deep learning can complete this task and provide more precise OCR, if you own enough datasets [5].

In this paper, we creatively apply Sobel operator to segmentation and creating datasets. Data augmentation was also used to expand the datasets, aiming to train more effectively and reduce over-fitting [6]. We adopt deep learning to feature extraction in OCR and apply it to recognizing the amount of words, which is an economic term and the most important part of VAT invoices. There are 16 categories of characters, which are generally used in amount of words, as shown Fig. 1 has been modified.

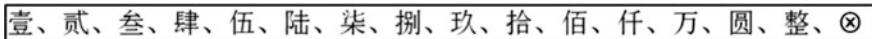


Fig. 1. Sixteen characters commonly used in VAT invoices and corresponding English explanations

This paper is organized as follows. In Sect. 2, we introduce Sobel operator and show how to apply Sobel operator to segmentation and creating our datasets. In Sect. 3, two algorithms of deep learning are described. Section 4 shows the experimental results of VAT invoice recognition. We conclude the paper in Sect. 5.

2 Sobel Operator and Production of Datasets

2.1 Sobel Operator

Sobel operator is one of the most important operators in image processing and plays an important role in machine learning and computer vision. The main role of Sobel operator is edge detection, and smooth suppression of noise is another function. Sobel operator weights the gray values of the neighborhood of every point in the image. This weight value will reach the extreme at the edge, so we can apply it to detecting the edge of the image [7].

Sobel operator has a pair of 3×3 matrices. Convolving them with the image, we can get the approximate the gradient of the image intensity function in the horizontal and vertical directions [8].

$$\mathbf{G}_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * \mathbf{A} \quad (1)$$

$$\mathbf{G}_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{A} \quad (2)$$

The gradient magnitude is shown as follows:

$$\mathbf{G} = \sqrt{\mathbf{G}_x^2 + \mathbf{G}_y^2} \quad (3)$$

2.2 Production of Datasets

The production of datasets is one of the most important parts of the whole experiment, which can affect the result of networks training. The steps of making the datasets are as follows.

2.2.1 Positioning

Positioning the location of the amount of words is the first step in production of datasets. VAT invoices have standard templates, and all characters and seals must be printed or covered in the specified location. Otherwise, this VAT invoice is invalid. So their positions are relatively fixed. In this case, it is easy to intercept the amount of words and get character sequences by fixed template, which record position information. The result is shown in Fig. 2.



Fig. 2. The image of the amount of words obtained from the whole VAT invoice with a fixed template. In addition, characters in this image are \otimes , two, ten thousand, two, hundred, nine, ten, five, yuan and integer, and the image represents 2295 yuan

2.2.2 Preprocessing

Before segmentation, we applied image binarization and Sobel operator to process the intercepted character sequence. Image binarization aims to reduce the amount of information and speed up processing [9]. Sobel operator is to reduce the failure rate of segmentation, due to that many Chinese characters are made up of two parts and if splitted directly, the character will be splitted into two parts. However, after applying the Sobel operator, we can detect the edge of a character and turn the character into a whole, bringing about increasing the rate of correct segmentation (Fig. 3).

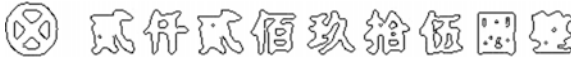


Fig. 3. The image processed by image binarization and Sobel operator

2.2.3 Segmentation

- Step 1. Project the processed character sequences into horizontal and vertical directions, respectively, and draw histograms of their corresponding pixel distributions, as is shown in Fig. 4. From their histograms, we can easily location every character.
- Step 2. Split horizontally according to the histogram that projected the image to the vertical direction. The result is shown in Fig. 5.
- Step 3. Split vertically based on the histogram that projected the image horizontally. The final segmentation result is shown in Fig. 6.

Using this method, we can make correct segmentation and about 2500 images of individual characters were collected, which were divided into 16 categories, everyone ranging from 100 to 200 images that are datasets. We divide the dataset into two parts of training datasets and test datasets. Every character in the test datasets has 30 pieces of images, and the rest are the training datasets.

2.2.4 Data Augmentation

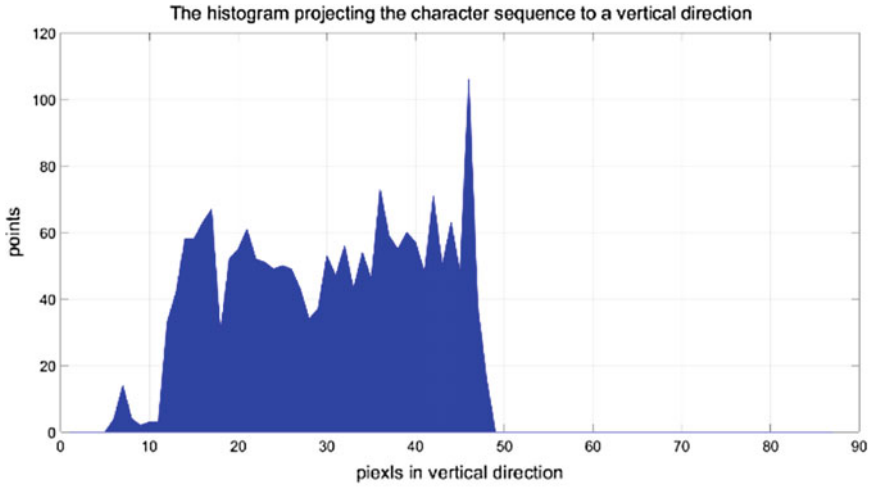
However, the number of every type of image in the training datasets is not equal and some of types do not have enough data to train, which can lead to over-fitting. To reduce over-fitting, the most commonly used method is data augmentation, which artificially enlarges the datasets using label-preserving transformations [10]. We use random crop, shift, and rotation to, aiming to more effective training, and the number of images in every category is expanded to about 400 pieces, which is enough to train.

3 Two Deep Learning Algorithms

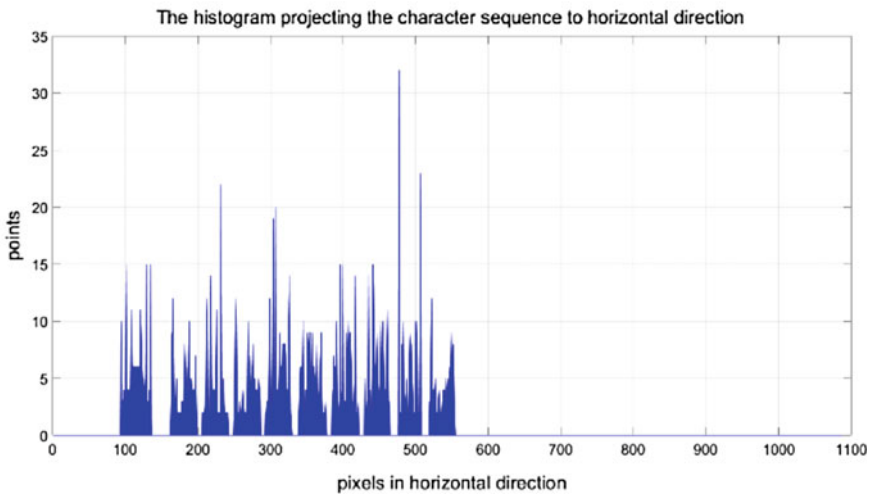
Deep learning is a branch of machine learning. It is an algorithm that attempts to abstract data characteristics using multiple processing layers that consist of complex structures or multiple nonlinear transformations. In this paper, the deep learning framework used is CNN, as a result of its better results in image and speech recognition. CNN is a kind of feedforward neural network, which consists of three basic parts of convolutional layer, pooling layer, and fully connected layer [10].

- Convolutional Layer

The convolutional layer is the heaviest and most basic component of a convolutional neural network and consists of several convolutional units. The purpose of convolutional layer is to extract the different characteristics of the input.



(a) The histogram projecting the character sequence to vertical direction.



(b) The histogram projecting the character sequence to horizontal direction.

Fig. 4. Two histograms projecting a character sequence to vertical and horizontal directions



Fig. 5. The image was splitted vertically based on the histogram projected the image horizontally, and characters in this image are same with characters in Fig. 2



Fig. 6. The image was splitted vertically by histogram and it is also the result of the final segmentation, and characters in this image are same with characters in Figs. 2 and 5

- Pooling Layer

Pooling layer is down-sampling in nature. Pooling layer can reduce the input dimension, thereby reducing the amount of parameters and simplifying the calculation; it also decreases the possibility of over-fitting. Pooling layers are generally inserted periodically between two adjacent convolution layers.

- Fully Connected Layer

Fully connected layer is generally the last layer of the network, the main function of which is classifier.

3.1 Deep CNN

In this paper, the structure of this network is shown in Table 1. Parametric Rectified Linear Unit (PReLU) has been chosen as activation function for every applicable layer, except the last output layer selecting Softmax, which is suitable for multiple classification problems, as activation function. In addition, we use categorical cross-entropy as loss function [10].

Table 1. The type of layer in the deep CNN, along with the activation function and data dimension of each layer

Layer	Dimension	Activation
Input	$48 \times 48 \times 3$	\
Convolutional 2D	$64 \times 5 \times 5$	PReLU
Max pooling 2D	5×5	\
Convolutional 2D	$64 \times 3 \times 3$	PReLU
Convolutional 2D	$64 \times 3 \times 3$	PReLU
Average pooling 2D	3×3	\
Convolutional 2D	$128 \times 3 \times 3$	PReLU
Convolutional 2D	$128 \times 3 \times 3$	PReLU
Average pooling 2D	3×3	PReLU
Flatten	\	\
Dense	1024	PReLU
Dropout (0.5)	\	\
Dense	1024	PReLU
Dropout (0.5)	\	\
Dense	16	Softmax

3.2 ResNet

Residual network is the classification task champion of ILSVRC2015 [11], created by Kaming He et al. ResNet, introducing a structure named residual block [11], solves the problem that even if you increase the number of network layers and make the network deeper, you cannot improve the accuracy of the network, due to degradation problem [11]. In this paper, we use a 110-layer residual network.

4 Experimental Result

In this section, we present our experiment results. We use MATLAB R2017a to complete our datasets created. Deep CNN and ResNet, relying on Keras2.1.5 and Tensorflow1.6.0, were trained and tested by two NVIDIA GeForce GTX 1080 Ti GPUs.

The time of every training epoch of Deep CNN ranges from 48 s to 52 s, but ResNet spends about 6.5 min per training epoch in the same devices and environment, as shown in Fig. 7. The accuracy of deep CNN reaches to above 95% after the tenth training epoch, which is so quick, but it rarely improves during the subsequent training epoch. However, the accuracy of ResNet can continuously rise and reach up to more than 99%, though it improves slowly, due to the deeper network structure of ResNet.

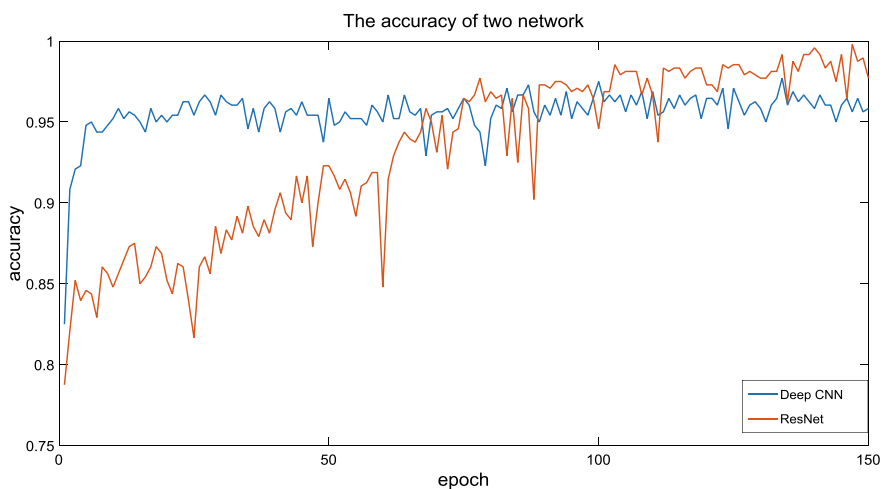


Fig. 7. Accuracy changes of deep CNN and ResNet in test datasets

We use the confusion matrix to show the recognition accuracy in Fig. 8. Confusion matrix, also known as the possibility or error matrix, is a specific matrix used to visualize the performance of the algorithm, generally supervised learning [12]. Every column represents the predicted value, and every row represents the actual category. Recognition accuracy of Deep CNN and ResNet is, respectively, 97.08% and 99.38%.

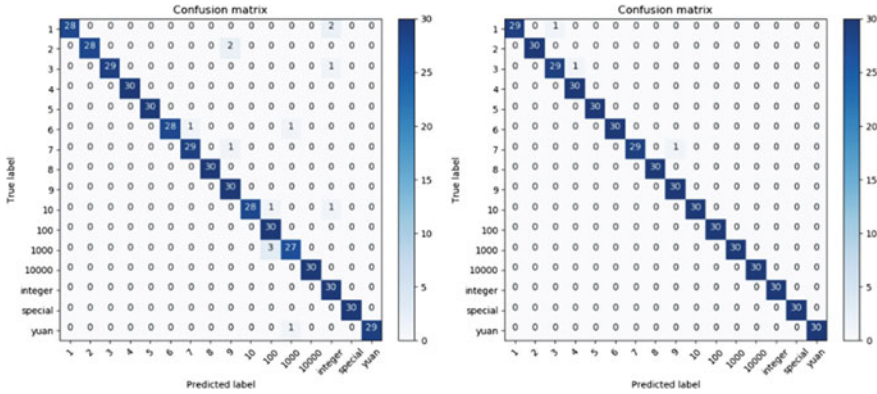


Fig. 8. There are two confusion matrices, with deep CNN on the left and ResNet on the right. In the confusion matrices, ‘1’, ‘2’, ‘3’, ‘4’, ‘5’, ‘6’, ‘7’, ‘8’, ‘9’, ‘10’, ‘100’, ‘1000’, ‘10,000’, ‘integer’, ‘special’ and ‘Yuan’ represent individually one, two, three, four, five, six, seven, eight, nine, ten, hundred, thousand, ten thousand, integer, ⊗ and Yuan in VAT invoices

5 Conclusion

In this paper, we have demonstrated that algorithms of deep learning, which own outstanding image classification capability, can be applied to OCR, especially the recognition of amount in words in VAT invoices. We apply Sobel operator and other image processing techniques to create datasets used to train our network, aiming to solve the problem of amount of words recognition in VAT invoices and achieve high recognition accuracy in this problem.

The deep learning methods are applicable to OCR and can be applied to the recognition of amount of words in VAT invoice with high efficiency and accuracy, because of that they can automatically extract the features of every class of pictures from the datasets, if you have huge amount of data. Deep learning has indicated its strong power and broad application prospects in field of OCR.

References

1. Modi, H., Scholar, P.G., Parikh, M.C.: A review on optical character recognition techniques. *Int. J. Comput. Appl.* **160**(6), 975–8887 (2017)
2. Sawant, A.S., Chougule, D.G.: Script independent text pre-processing and segmentation for OCR. In: International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO), pp. 1–5 (2015)
3. Mohammad, F., Anarase, J., Shingote, M., Ghanwat, P.: Optical character recognition implementation using pattern matching. *Int. J. Comput. Sci. Inform. Technol.* **5**(2), 2088–2090 (2014)
4. Yi, C., Tian, Y.: Scene text recognition in mobile applications by character descriptor and structure configuration. *IEEE Trans. Image Process.* **23**(7), 2972–2982 (2014)

5. Shi, B., Bai, X., Yao, C.: An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(11), 2298–2304 (2017)
6. Wigington, C., Stewart, S., Davis, B., Barrett, B., Price, B., Cohen, S.: Data augmentation for recognition of handwritten words and lines using a CNN-LSTM network. In: In 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), pp. 639–645 (2017)
7. Vairalkar, M.K.: Edge detection of images using Sobel operator. *Int. J. Emerg. Technol. Adv. Eng.* **2**(1), 291–293 2012
8. Tabatabai, A.J., Mitchell, O.R.: Edge location to subpixel values in digital imagery. *IEEE Trans. Pattern Anal. Mach. Intell.* **6**(2), 188–201 (1984)
9. Gupta, M.R., Jacobson, N.P., Garcia, E.K.: OCR binarization and image pre-processing for searching historical documents. *Pattern Recogn.* **40**(2), 389–397 (2007)
10. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: International Conference on Neural Information Processing Systems, pp. 1–9 (2012)
11. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
12. Ohta, M., Takasu, A., Adachi, J.: Retrieval methods for English-text with miss recognized OCR characters. In: International Conference on Document Analysis and Recognition, pp. 950–956 (1997)



Deep Learning Based Detection Method for SDN Malicious Applications

Chi Yaping, Yu Yuzhou^(✉), and Yang Jianxi

Beijing Electronics Science & Technology Institute, Beijing, China
gloryperson@gmail.com, 1245497865@qq.com

Abstract. SDN is a new type of network architecture. The core technology of the SDN is to separate the control plane of the network device from the data plane so as to achieve flexible control of network traffic. Such structure and characteristics have put forward higher requirements on the security protection capability of the SDN controller. However, there are still less researches on malicious applications for the SDN network architecture. This article aims at this problem, based on the analysis of the existing malicious application detection methods and on deep learning technology proposed by a detection method for SDN malicious applications. Finally, under the TensorFlow deep learning simulation environment Keras, 30 SDN malicious samples were studied and tested. The experimental data show that the detection rate of this method for malicious applications can reach 89%, which proves the feasibility and scientificity of the program.

Keywords: SDN · Malicious applications · Deep learning

1 Introduction

Malware is one of the major threats to today's computer systems. Malware is constantly evolving, and its complexity is getting higher [1]. At the same time, there are more and more types of attacks against SDN networks. However, most of the existing SDN-related research focuses on using of software-defined network control and forwarding separation features and the network architecture to solve the traditional malicious attacks, and less research on malicious applications specifically targeted at SDN.

In [2], a dynamic analysis environment (sandbox) was set up to detect malicious applications. However, this detection system also has its obvious limitations, because the SDN controller itself in this system is also vulnerable to the SDN model of threats and attacks. If the controller is compromised, the system will not be able to properly perform malware analysis functions. And the original intention of this scheme is to detect malicious code for ordinary computers instead of malicious code for SDN controllers. Lee et al. [3] propose a framework SHIELD for automatically analyzing SDN malicious applications, which examines the API calls of SDN applications and summarizes existing benign and malicious applications and extracts their behavioral

characteristics to detect malicious behaviors, but it is obvious. The disadvantage is that the detection of malicious applications that can hide their own behavior is very poor. Röpke [4] proposed two protection schemes: one is application isolation and domain adoption method, and the sandbox is used to place important modules on the control in different domains and isolate each other to ensure that one of them is not affected by other infections. The other solution is to add a proxy between the controller and the switch. The proxy controller functions as a policy checker so that the proxy controller judges each time the controller sends a flow table. Whether the flow table issued by it is reasonable or not is a normal instruction or a malicious behavior generated by a malicious application, and then whether or not to isolate the related application according to the condition. Reference [5] refers to the isolation of SDN applications through virtualization to protect the control plane, or to take strict measures to separate malicious SDN applications. At the same time, the policy checking mechanism is allowed to operate independently of the SDN controller or a behavioral policy detection module is added at the control layer to give a certain level to each behavior and operation of the related application. The malicious application and protection can be detected by judging whether the application behavior is malicious or not. The purpose of the controller. Such a detection method does not analyze the structure of the malicious application itself, but only waits for the malicious behavior to trigger after the sample analysis and extraction rules, so the accuracy and timeliness of the time-liness are detected of unknown malicious applications, etc. It is obviously deficient. Reference [6] introduced the detection of malicious applications on the Android system, and pointed out that the extracted features include Android installation program, file MD5 value, package name, startup activity name, permission information, and other static features, which did not go deep into the malicious application internal attention. In the case of function calls, these static features may be forged and it is difficult to accurately describe a malicious application.

This article addresses the security issues of the SDN controller itself, especially the application layer security that is now weak in protection capabilities. We start from the function calls inside each application, extract features, and combine the idea of artificial intelligence about deep learning with SDN's malicious application detection to provide a new method for detecting SDN malicious applications.

2 The Difference Between SDN's Malicious Applications and Ordinary Malicious Applications

The working principle and trigger mechanism of SDN applications and traditional applications are different. SDN applications are usually activated by events such as flow table in forwarding layers, but general malicious applications such as Android malicious are activated by the user or a specific time or automatically. These differences also determine that we should adopt different detection methods (Table 1).

Table 1. Differences between malicious applications and malicious applications for SDN

	SDN malicious applications	General malicious applications
Existing form	Independent existence	Independent or in the form of plug-ins
Way for spreading	The Internet	All methods including the Internet
Affected objects	SDN controller	Host or network
Trigger method	OpenFlow table	Designated or computer user
Call interface	Restful API and Ad Hoc API	Specific system API
The level of danger	Harmful	Uncertain
Operating platform	New network platform	Traditional platform, smart mobile

3 Deep Learning Based Malicious Application Detection

The convolutional neural network directly takes the original data as input, automatically learns the representation of features from a large amount of training data, and has the features such as local connection, weight sharing, and pooling operations, which can effectively reduce the complexity of the network and reduce the number of training parameters. It can also make model to have some features of invariability to translation, distortion, and scaling.

The research shows that the features obtained by convolution neural network are more discriminative and generalization than those designed artificially. Feature expression is the research foundation of computer vision. How to use convolution neural network to learn, extract, and analyze the characteristic expression of information, so as to obtain more general features with better discriminant performance and better generalization performance, and will have a positive impact on the whole computer vision and even more extensive fields.

3.1 Malicious Application Detection Scheme Design

The main idea of this method is to transform SDN malicious applications into samples for deep learning. The industry pointed out that deep learning plays an important role in the field of image recognition. At present, the best system is based on deep learning, so this method will transform all benign and malicious applications into samples in the form of images and then imported into the deep learning framework for deep learning.

The scheme has three advantages:

- (1) In the past, the static feature extraction of malicious code mainly focuses on its behavior, label, package name, activity name, permission information, and other features that are easy to be hidden or forged. This scheme is a high reliability, hard to forgery sample for the control flowchart of malicious applications (Control Flow Graph, CFG diagram). This kind of high reliability is difficult to forge samples greatly increase the authenticity and usability of the sample.
- (2) Deep convolutional neural networks are now recognized as the most outstanding deep learning framework in the field of image recognition.

- (3) Convolutional neural network as a deep learning model is with the ability to layer learning characteristics. Studies have shown that the features learned through convolutional neural networks have stronger discriminative ability and generalization ability than artificial design features (Fig. 1).

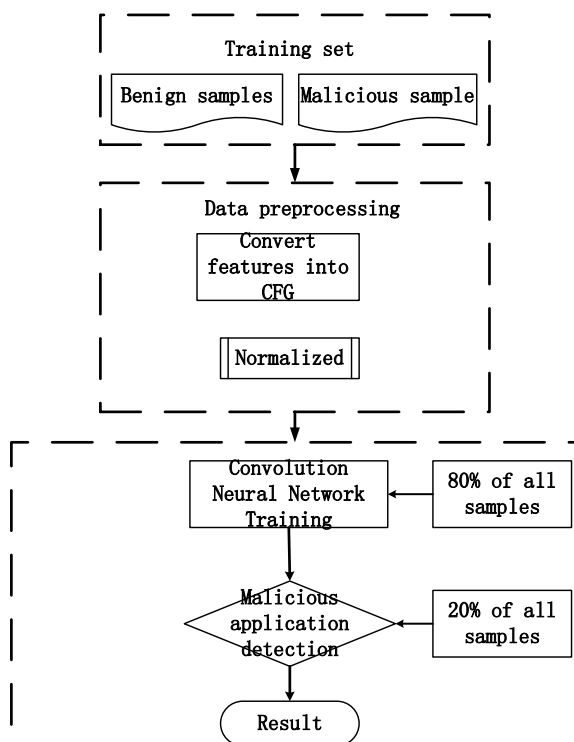


Fig. 1. Mainly introduces the workflow of the method. Next, we will focus on the generation of malicious application samples and the construction and training of convolutional neural networks. The sequential convolutional neural network uses a sequential model

4 Method Implementation

4.1 Related Software Introduction

Graph-easy is a Perl module for graphical DSLs that provide customers with an easy-to-understand, highly readable graphical description language and supports Graphviz, VCG, and GDL files in a variety of imported formats.

Keras is a high-level neural network API. Keras is written in pure Python (Python version: Python 2.7–3.6) and supports TensorFlow, Theano, and CNTK backends. Keras is highly modular, minimalistic, and extensible, supporting CNN and RNN (Table 2).

Table 2. Experimental environment configuration

Experimental environment	Development environment
Operating system	Windows10
Processor	Inter i5-4300
RAM	8 GB
Programming language	Python 3.6
Deep learning framework	TensorFlow

Sample source:

- (1) 12 samples found from the GitHub open source community.
- (2) 20 samples shared by the research team in Reference 3.

4.2 Experimental Implementation

- (1) First, we divide the application into two categories (benign, malicious), and IDA reverse tool, call IDC module, processing the original sample, generate Graph Description Language (GDL) file.

This article took into account the following points when generating GDL files:

1. Simplify the process, the number of iterations, and function calls to the same function that does not exceed two times.
 2. The picture of this method is just an undirected figure. It only cares about the structure of the entire program and does not care about the calling relationship between functions.
 3. In order to simplify the call relationship, display the call relationship of all function interfaces and ignore part of the call for the local library functions.
 4. The function entry point is set to green, and the main function entry point is set to blue. We pay special attention to the functional modules peculiar to SDN malicious applications just like `DataChangeListener()`, `FlowRuleModification()`, `SwitchFirmwareMisuse()`, and `ReplayingKnownAttack()`; the main function call function is set to pink. Set the API of the SDN controller to be light green
- (2) Graphical transformation model Graph-easy is used to convert GDL files into CFG flowcharts (as shown in Fig. 2).
 - (3) Write a script to normalize and batch the generated flowcharts so that they become $256 * 256$ PNG format samples for deep learning.
 - (4) Based on Google's deep learning framework TensorFlow and Keras, build a deep convolutional neural network. Based on previous experience, debug various parameters.

First of all, there is no fixed method for setting various parameters in a convolutional neural network, and there is no so-called optimal method. Therefore, the following parameters are selected after referring to the setting method given in [15, 16]. The optimal combination of efficiency and accuracy is coupled with their own understanding of the set.

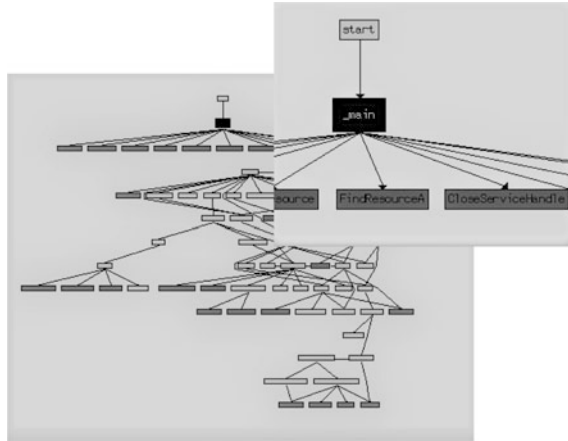


Fig. 2. The original picture is a color picture, and the color setting rule is as mentioned in the fourth point of step (1). In Fig. 2, the larger image is a complete CFG image, and the small image is a partial enlarged view of the main function to illustrate

Stride is the step size, filter is the filter (also known as the convolution kernel), kernel_size is the size of the convolution kernel, and padding is equal to zero when the

Table 3. Convolutional neural network parameters used in this paper

	Stride	Padding	Filter	Kernel_size
Convolution layer 1	1	Same	10	3 * 3
Pooling layer 1	2	0	10	2 * 2
Convolution layer 2	2	Same	20	3 * 3
Pooling layer 2	2	0	20	2 * 2
Convolution layer 3	1	Same	40	3 * 3
Pooling layer 3	2	0	40	2 * 2
Fully connected layer	2	0	120	

Note Regarding the choice of step size, if the step length of the convolutional layer is too long, it will lose a large number of features, too small will lead to overfitting, so the step length is 2. The step size of the pooling layer and the parameters of the convolution kernel are based on previous experience. Regarding the selection of the filter, the basis for the setting is mainly the image processing capability of the experimental platform. Regarding the size of the convolution kernel, the size of the mainstream convolution kernel now introduced is usually 3 * 3, 5 * 5, so a 3 * 3 convolution kernel is used, and the calculation of formula (2) is also considered. The result of the process can be an integer so there is also a 2 * 2 convolution kernel. The choice of padding is also to enable the overall structure to be more reasonable and optimized.

padding is equal. At this time, a value of 1 is assigned. When it is valid, it means that it is not filled.

out_p represents the output after the pooled layer, and out_c represents the output through a convolutional layer (Table 3).

$$out_p = in_{length} / stride \quad (1)$$

$$out_c = \frac{in_{length} - kernel_size + padding}{stride} + 1 \quad (2)$$

- (5) Put all sample numbers in the same directory, put 80% of the samples into the deep learning framework as a training set for learning, and use 20% of the samples as test sets to test the accuracy of the model.

4.3 Comparison of Experimental Results

Result analysis:

In Table 4, loss represents the loss value, acc represents the accuracy of the training set, and val_acc represents the accuracy of the test set. The loss represents that the loss value is greater than 1, indicating that the loss of information is serious and there is a certain overfitting phenomenon, but it tends to become smaller gradually. The accuracy of this method is set at 0.8966. If the number of samples and the number of cycles are increased, the result will be more ideal, however. The continuous development of the deep convolutional neural network makes the number of neural networks more and more, the longer the calculation time, the more the resources are consumed, and the accuracy rate is correspondingly improved. However, the overfitting phenomenon may be more serious, so the convolution of this paper. On the basis of VGG and AlexNet, the neural network adjusts the design and parameters of each layer according to the characteristics of the method, taking into account the accuracy and efficiency, and achieves better results (Table 5).

Table 4. Comparison of different deep learning frameworks

Framework	Loss	Acc	Val_loss	Val_acc
AlexNet	0.7821	0.5000	0.6957	0.4643
VGG	5.0369	0.6875	3.3348	0.7931
Described in Table 2	0.6897	0.6429	0.3365	0.8966

Table 5. Different methods for the same malicious application in average detection times

	OpenDaylight (s)	Floodlight (s)
SHIELD	27.85	45.26
Described in Table 2	4.42	3.52

Compared with the SDN malicious application detection method mentioned in [3], the detection accuracy rate of SDN malicious application is not mentioned in [3], only given the average detection time, so the results are obviously not accurate enough and rigorous. Since there are few documents related to SDN malicious application detection; after a comparative analysis based on existing data, this experiment can better complete the malicious application detection function.

We can achieve a certain goal in the detection of malicious code of SDN through a deep learning method and achieve an ideal result. However, because there are still few relevant samples on SDN malicious applications, the highest accuracy rate is only about 89%.

5 Concluding Remarks

SDN is a new type of network architecture, but its research on security is still relatively backward. We provide a new method for the detection of SDN malicious applications, which has important practical significance for the security protection of SDN.

However, because there are few malicious applications of samples for SDN networks, the detection accuracy is not high enough. With the progress of follow-up research, after collecting a large number of SDN malicious application samples, we can continue to optimize and perfect the detection method and verify the method. At the same time, due to the complex logic relationship of the application program, the control flow graph is very complex and the features are too scattered, which is not conducive to deep learning. Next, we will remove some noncritical calls and retain the key processes to improve the accuracy.

Acknowledgements. This research was financially supported by the National Key R&D Program of China (No: 2018YFB1004101).

References

1. Zhang, Y., Pan, X., Liu, Q.: APT attacks and defenses. pp. 1–7 (2017-08-10). <https://doi.org/10.16511/j.cnki.qhdxxb.2017.21.024>
2. Ceron, J.M., Margi, C.B., Granville, L.Z.: MARS: an SDN-based malware analysis solution. In: Computers and Communication, pp. 525–530. IEEE (2016); Foster, I., Kesselman, C.: The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufmann, San Francisco (1999)
3. Lee, C., Shin, S.: SHIELD: an automated framework for static analysis of SDN applications. In: ACM, pp. 29–34 (2016)
4. Röpke, C.: SDN Malware: Problems of Current Protection Systems and Potential Countermeasures. Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn (2016)
5. Röpke, C., Holz, T.: SDN Rootkits: subverting network operating systems of software-defined networks. In: Research in Attacks, Intrusions, and Defenses, pp. 339–356. Springer International Publishing (2015)

6. Lee, S., Yoon, C., Shin, S.: The smaller, the shrewder: a simple malicious application can kill an entire SDN environment. In: ACM International Workshop on Security in Software Defined Networks & Network Function Virtualization, pp. 23–28. ACM (2016)
7. Dai, H., Dai, B., Song, L.: Discriminative embeddings of latent variable models for structured data (2016)
8. Cheung, S., Fong, M., Porras, P., et al.: Securing the software-defined network control layer (2015)
9. Röpke, C., Holz, T.: On network operating system security. *Int. J. Netw. Manag.* **26**(1), 6–24 (2015)
10. Feng, Q., Zhou, R., Xu, C., et al.: Scalable graph-based bug search for firmware images. In: ACM Sigsac Conference on Computer and Communications Security, pp. 480–491. ACM (2016)
11. Xu, X., Liu, C., Feng, Q., et al.: Neural network-based graph embedding for cross-platform binary code similarity detection. In: National Center for Biotechnology Information (2017). <http://www.ncbi.nlm.nih.gov>



Survey of Big Data Application Technology on Multimedia Data of Public Security

Huibo Li, Yanan Jiang^(✉), Yunxiang Yang, Jing Guo, Xiaocheng Hu, Ke Guo, Bo Zhang, and Jing Cheng

China Academy of Electronic Information Technology, Beijing 100041, China
jiang_yinan@126.com

Abstract. The era of multimedia big data has a profound and extensive impact on the field of public security. The application of multimedia data and big data technology has brought new opportunities to the construction of public security system, as well as new challenges. This paper summarizes the new characteristics of various public security risk events, such as violent terrorist attacks, serious criminal offences, major group events, and network crimes, and analyzes the main problems existing in the application of big data technology in the field of public security. The progresses and trends of some essential technologies are analyzed.

Keywords: Multimedia data · Big data · Public security · Association analysis

1 Introduction

With the fleeting progress of computer technology, network communication, and digital media, the technology of obtaining, generating, processing, and analyzing multimedia data has also been developed rapidly. The information society has entered the era of multimedia big data. Multimedia big data includes but is not limited to text, image, graphics, audio, video, social, and sensor data that is highly valuable in decision-making. The application of multimedia big data has brought momentous changes to the fields of finance, transportation, medical and health, education, safety, and so on [1–3]. In the field of public security, which is crucial to the national security system, this change is more profound and extensive. Through satellite communications, video surveillance, sensor equipment, mobile Internet devices, and other ways, mass multimedia big data has been generated. It is fundamentally changing the environment, countermeasures, and even the way of thinking of public security problems [4, 5].

The popularity of multimedia big data and the immediacy of communication make the problem of public security more complex and diversified [6, 7]. It brings new characteristics to the public security risks, including violent terrorist attacks, serious criminal offences, severe group events, and network crimes. The main new characteristics can be summarized as follows:

- (1) Dissemination. In the era of multimedia big data, the generation, dissemination, and acquisition of information are extremely convenient, which improves the efficiency of the social and economic operation. On the other hand, the extensive

dissemination and the instant propagation of information can lead to the rapid spread of false information such as rumors and panic. Moreover, the tremendous panic and anger would easily induce group events [8, 9].

- (2) Concealment. Benefited from the high degree of liberalization and virtualization of the data space, various kinds of activities that may endanger public security are conducted through the Internet, including Internet fraud, online theft, the communication of criminals, and so on. The concealment of these activities is mainly embodied in two aspects. The first is the concealment of the data generation, because the important information related to the criminal activity is often fragmented in the massive raw multimedia data. The second is the concealment of the data transmission. Compared with the point-to-surface transmission mode, point-to-point and even point-to-circle transmission mode make it difficult to track the path and scope of information dissemination [10].
- (3) Accessibility. With the fully developed and wide application of multimedia technology and network technology, information dissemination is not restricted by space and almost has no boundaries. The impact of the broad outreach on public security is also reflected in two aspects. One is the wide range of participation. The criminals can communicate, control, and command remotely, which drastically reduces the cross-border and trans-regional crime cost. The other one is the wide range of influence. The spread of public opinion and group events have a wider audience, which means the problems become more difficult to control and management.

The characteristics of mass multimedia data for public security risk often exist and intertwined in the same public security event, which causes difficulty to implement the traditional prevention and control measures effectively. There is an urgent need to study latest technologies to prevent, mitigate, and predict the evolution of public security risk. Therefore, it is an effective means to deal with the new challenges of public security risk by using big data technology to detect, integrate, and analyze the public security-related multimedia information from mass multidimensional fragmented data, and to identify public security risk with a forward-looking view.

2 Analysis of the Application Problems of Multimedia Data in Public Security

Facing with the massive multimedia data, the key of reducing public security risk lies in the comprehensive collection, fusion, understanding, mining and analysis of public security-related information, the discovery of the potential associated relations within the raw data, the reveal of the patterns of public security events, and the accurate prediction of the evolution situation of public security events.

Focusing on these crucial issues, there remain some application problems of multimedia data in public security:

(1) Insufficient data collection

The concealment and complexity of public security risk result in the high requirement of data for risk perception. First, to identify all kinds of potential risks, it is necessary to collect multi-source heterogeneous data that crossing both virtual and physical spaces. Second, fine granularity of data is the prerequisite for rapid identification of potential risk sources and promptly disposal of public crisis. At present, in terms of big data collection in public security, abundant data of human activities, production and facility operation that related to public security cannot be captured promptly and effectively or cannot be transferred to the decision-making system in time.

(2) Incomplete data fusion and integration

At present, the collected multimedia data in public security have typical features of big data which include a wide variety of sources, different structures, and huge volume. But in practice, the information cannot be shared and integrated since these data are often widely scattered among various departments, institutions, and enterprises. The so-called information isolated islands become a common phenomenon. For example, the public safety management system is often operated based on the “divide and conquer” principle. And in many cities, several or even a dozen security platforms exist simultaneously in a safety city system without interconnection and integration according to practical services. Accordingly, there is a lack of intelligent technologies to conduct data cleaning and data fusion when facing the multi-source heterogeneous data that crossed networks and applications in a public security system.

(3) Imperfect data association analysis

The difficulty of applying public security big data lies not only in the huge volume but also in the hidden complex relations between each other. Only by digging the inherent laws and relations, can we find the logical consistency behind the scattered information, and then accurately analyze and predict the evolution trend of public security events [11]. For example, it is normal that a person bought a pressure cooker or a clock or a steel ball. But it is definitely abnormal if this person brought all these things together. However, the current public security multimedia data and its application have the problems of fragmented information, low-value density, low utilization of data resources, and limited means of data analysis. To exploit and search for clues in massive, scattered, chaotic, and seemingly unrelated data, it is necessary to introduce intelligent data association and analysis techniques comprehensively.

(4) Inefficient assistant decision-making

Through multimedia big data to assist the relevant decision-making of public security risk, the transformation from “business experience driven” to “data quantification

drive” can be realized. It not only provides quantitative scientific basis for emergency decision-making but also provides strong support for early prediction and early warning [12].

To analyze and process mass public security multimedia data and then assist in decision-making, we need to establish effective prediction and early warning models based on data fusion understanding and association analysis, then complete the “data—information—knowledge—decision” conversion in a short time. At present, in the field of public security, prediction and early warning models for multimedia data and assistant decision-making tools are still in the initial stage of application, and there is a lack of systematic application with practical effect and operability.

3 Progresses and Trends of the Application Technologies of Multimedia Big Data in Public Security

In view of the mentioned application problems, this section focuses on two key issues, which are fusion and understanding of multi-source heterogeneous multimedia data and association analysis of mass multidimensional multimedia data in public security, to analyze the current progresses and trends of the application technologies.

3.1 Fusion and Understanding of Multi-source Heterogeneous Multimedia Data

The traditional methods of data fusion and understanding based on static and surface features cannot adapt to the demand for deep intelligent understanding and computing applications of multimedia data content. The emergence of big data provides an opportunity for the deep modeling of multimedia data content and the intelligent semantic understanding. Because of the characteristics of a wide range of data sources, different data structures, numerous data categories, long time span of data generation, and uneven data quality, the application in public security demands higher requirements for the organization and management of multimedia data. Currently, in the field of public security, the fusion and understanding of multi-source heterogeneous multimedia data are not systematic and operational in practice. Therefore, the well-established technologies like big data fusion technology from other fields can be adopted.

The universal mode of big data fusion is shown in Fig. 1 [13]. The critical supporting technologies needed in this mode include schema/ontology alignment, entity linking, conflict resolution, and relation inference and evolution [14–16]. The technology of schema/ontology alignment is used to deal with heterogeneity of ontology and heterogeneity of data sources. Entity linking consists of named entity recognition and record linkage and entity association, which are the basis of big data fusion. Conflict resolution currently focuses on the authenticity screening and evolution of entity level. Relation inference and evolution is used to extend and complement knowledge base and discover hidden knowledge.

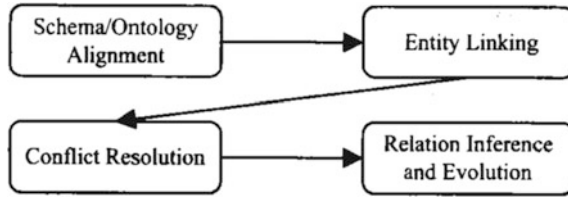


Fig. 1. Universal mode of big data fusion [13]

Drawing on the framework of multi-sensor data fusion technology, we can also divide the public security data fusion into data-level fusion, feature-level fusion, and decision-level fusion, as shown in Fig. 2 [17]. The framework divides data fusion into low-, middle-, and high-level fusion according to the type of processed information. The low-level fusion is data-level fusion, and the main purpose is to integrate variables and signals collected by different sensors. The middle-level fusion is feature-level fusion, in which the main features need to be extracted from the collected signals and processed. High-level fusion is decision-level fusion. The information recognition and integration can be realized eventually by applying various complex decision algorithms.

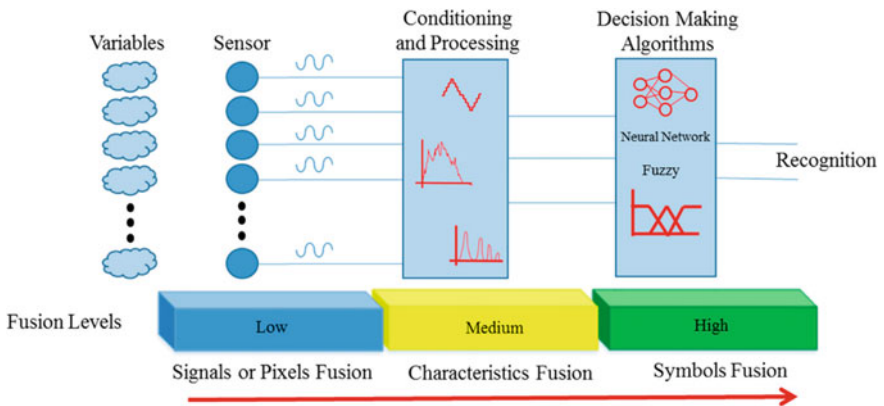


Fig. 2. Summarized flowchart of a sensor fusion system [17]

For the realistic demands of dealing with mass heterogeneous multimedia data in the field of public security, the key technical methods of multimedia data fusion and understanding are mainly concentrated in the middle and high levels, namely, the feature-level fusion and decision-level fusion. As a typical technique for characteristic learning, machine learning has been successfully applied to the fields of speech recognition, image processing, search, and recommendation through automatic learning of the multilayer representation of data, such as supervised/unsupervised learning and deep neural networks [18, 19]. These technologies can also be applied to feature extraction, label generation, and information fusion of multimedia big data in public security.

3.2 Association Analysis of Mass Multidimensional Multimedia Data

The association analysis technology of big data is used to mine the association relationship among mass multi-source heterogeneous multimedia data, by combining applied mathematics, graphics, information science, and other subjects with citation analysis in documentary metrology, co-occurrence analysis, and association rules learning technologies. To make it more effective, we need to apply the mining further on the relationship diagram or knowledge map.

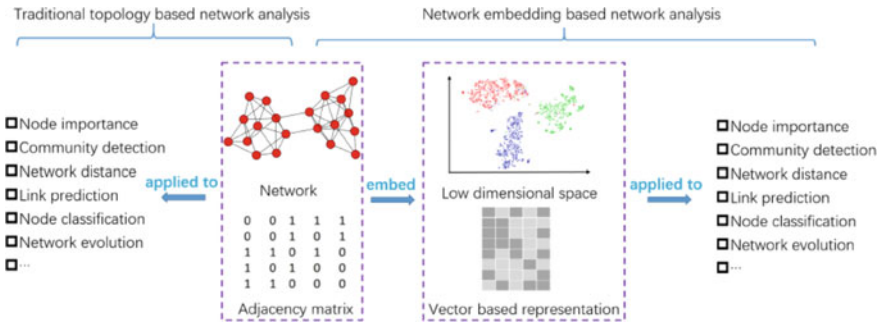


Fig. 3. A comparison between network topology-based network analysis and network embedding-based network analysis [20]

Currently, the main research direction is to improve the robustness of information inference, improve the ability of mining semantic related objects, and improve the dynamic incremental updating ability of the relation network, among which information inference is widely used to discover implicit knowledge, and inference function is usually accomplished by extensible rule engine. At present, the results of information inference are easily influenced by a partial change in object attributes or relations, so it is necessary to study the stable information inference algorithm based on the holographic data of the object. The related object search algorithms are mainly based on traditional connection relations or basic attributes, such as the classic PageRank algorithm. Therefore, it is necessary to establish semantic association models to mine semantically related objects. For a relation network, the original network structure needs to be updated dynamically when the nodes in the relation network appear to be increased, deleted, changed, or the type of links and link modes are changed. Because of the large-scale and heterogeneity of multimedia data, the computation cost of dynamic updating process is very high, so it is necessary to improve the efficiency of the network update mechanism in the algorithms.

In addition, the related individuals and the relationship between individuals in the field of public security can be abstracted as a complex social network composed of multilayer interaction among individuals. As the network has a strong and flexible representation and association ability, it is the most natural and direct way to model perceive multimedia big data of public security with inherent relevance. In order to analyze and process large-scale complex social networks efficiently, a method of

network embedding is proposed, as shown in Fig. 3 [20]. This method transforms the network topology into vector space, and the nodes in the network correspond to the sample points in the vector space, while the edges of the network are replaced by the distance metric of the vector space. The characterization of large-scale networks in vector space can effectively reduce the complexity. On this basis, distributed parallel computing can be easily proceeded, and machine learning algorithm can also be used to learn and analyze networked multimedia data directly.

4 Conclusion

The era of multimedia big data introduces new characteristics of dissemination, concealment, and accessibility to the public security risks, including violent terrorist attacks, serious criminal offences, severe group events, and network crimes. Meanwhile, the application of massive multimedia data and big data technology in the field of public security creates new opportunities for the construction of public security system. However, there remain some application problems including insufficient data collection, incomplete data fusion and integration, imperfect data association analysis, and inefficient assistant decision-making. Focusing on two key issues, which are fusion and understanding of multi-source heterogeneous multimedia data and association analysis of mass multidimensional multimedia data in public security, we analyze the current progresses and trends of the application technologies. It has reference significance for the application and research of public security multimedia big data technology.

Acknowledgements. This paper is supported by Beijing NOVA Program (Z181100006218041) and National Key R&D Program of China (2017YFC0820106).

References

1. Lazer, D., Kennedy, R., King, G., Vespignani, A.: Big data. The parable of Google Flu: traps in big data analysis. *Science* **343**(6176), 1203 (2014)
2. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., et al.: Big data: the next frontier for innovation, competition, and productivity. *Analytics* (2011)
3. Simon, P.: *Case Studies: The Big Rewards of Big Data*. Wiley (2015)
4. Williams, M.L., Burnap, P., Sloan, L.: Crime sensing with big data: the affordances and limitations of using open source communications to estimate crime patterns. *Br. J. Criminol.* **57**(2) (2017). <https://doi.org/10.1093/bjc/azw031>
5. Wang, H., Kifer D., Graif, C., Li, Z. (eds.): Crime rate inference with big data. In: *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016)
6. Dhaka, P., Johari, R. (eds.): CRIB: cyber crime investigation, data archival and analysis using big data tool. In: *International Conference on Computing, Communication and Automation* (2017)
7. Zhang, Q.: The research for the problem of big data application in current crime prevention and control. *Southwest University of Political Science and Law* (2016)

8. Sakaki, T., Okazaki, M., Matsuo, Y.: Tweet analysis for real-time event detection and earthquake reporting system development. *IEEE Trans. Knowl. Data Eng.* **25**(4), 919–931 (2013)
9. Wang, L.: *Research on Influence and Dynamic Evolution Analysis of Internet Public Opinion Events*. Tianjin University (2015)
10. Morton, J.F.: *Next-Generation Homeland Security: Network Federalism and the Course to National Preparedness*. US Naval Inst Press (2014)
11. Pedrycz, W., Chen, S.M.: *Information Granularity, Big Data, and Computational Intelligence*. *Studies in Big Data*, vol. 8 (2015)
12. Janssen, M., Voort, H.V.D., Wahyudi, A.: Factors influencing big data decision-making quality. *J. Bus. Res.* **70**, 338–345 (2017)
13. Meng, X., Du, Z.: Research on the big data fusion: issues and challenges. *J. Comput. Res. Dev.* **53**(2), 231–246 (2016)
14. Shvaiko, P., Euzenat, J.: Ontology matching: state of the art and future challenges. *IEEE Trans. Knowl. Data Eng.* **25**(1), 158–176 (2012)
15. Zhao, L., Ichise, R.: Ontology integration for linked data. *J. Data Semantics.* **3**(4), 237–254 (2014)
16. Matteini, A., Matteini, A., Isele, R., Becker, C., Becker, C. (eds.): *LDIF-linked data integration framework*. In: *International Conference on Consuming Linked Data* (2011)
17. Luo, R.C., Chang, C.C., Lai, C.C.: Multisensor fusion and integration: theories, applications, and its perspectives. *IEEE Sens. J.* **11**(12), 3122–3138 (2011)
18. Mitchell, T.M.: *Machine Learning*, pp. 417–433. McGraw-Hill (2003)
19. Witten, I.H., Frank, E.: *Data mining: practical machine learning tools and techniques*. *ACM SIGMOD Record.* **31**(1), 76–77 (2005)
20. Wang, D., Cui, P., Zhu, W. (eds.): *Structural deep network embedding*. In: *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016)



Pedestrian Detection Based on Deep Neural Network in Video Surveillance

Bo Zhang, Ke Guo^(✉), Yunxiang Yang, Jing Guo, Xueying Zhang, Xiaocheng Hu, Yinan Jiang, and Xinhai Zhang

China Academy of Electronics and Information Technology, China Electronics Technology Group Corporation, Beijing, China
guoke101@outlook.com

Abstract. Pedestrian detection is an essential and challenging problem in machine vision and video surveillance signal processing. To handle the high cost of training-specific discriminative classifier for pedestrian detection, we focus on the learning of suitable features for pedestrian detection representation. A deep neural network is presented in this paper to resolve the above issue. Our pedestrian detection method has several appealing properties. First, the learning of features is much more efficient under the configuration of the proposed framework due to the reduction of training classifier. Second, a K-Nearest Neighbor (KNN) method is adopted to solve the comparison between the regions of interest and the templates. Third, due to the less dependency of the classifier, the performance across different datasets overcomes most traditional ones. Finally, we perform extensive comparison across different public datasets and compared with corresponding benchmarks.

Keywords: Pedestrian detection · Deep neural network · Auto-encoding · Dimension reduction · KNN

1 Introduction

Pedestrian detection is a special case of object detection. Although object detection has been quickly developed for the last several decades as a fundamental component of machine vision, pedestrian detection still has lots of challenges in the progress of research, such as occlusion, variations, and outdoor environment. Several attempts have been made to improve the performance of the detecting of pedestrians under video surveillance scenes [5, 19, 22]. This paper has presented an approach for pedestrian detection under various video surveillance scenes. We have presented a novel model to detect human object which is to be distinguishable from other sorts of objects. Many recent papers have used deep neural networks to recognize the pedestrian from the crowded scenes and to detect them from videos and images [14, 18, 20]. They typically defined the process of pedestrian detection as two parts, extraction for Regions of Interest (ROIs) candidates, and discriminative classification of true positive ones from the others. The proposed method has a much different flavor from the approaches mentioned above. Our pedestrian detection algorithm mainly focuses on the deep learning of the representation for pedestrians. In this paper, we apply a fully connected

convolutional deep neural network based on auto-encoder to learn deep features of pedestrians by using large datasets at multiple scales to achieve high detection rates and low false positive rates. Our detection method builds on the auto-encoding work [12] and distance similarity. Novel contributions of this paper include the following: (i) we consider the pedestrian detection issue as a feature extraction task, rather than performing two components extraction and classification orderly; (ii) the weights of the deep neural network being updated during the training phrase would be employed to extract the representation of the candidate ROIs, which makes it possible to extract candidates' features online; and (iii) the distribution of the true positives and false positives is described by the distance between features, which enables the detection method has a better generalization ability compared with traditional classifiers.

The rest of this paper is organized as follows. Section 2 introduces the related work about pedestrian detection problem. In Sect. 3, the proposed detection algorithm is described. The experiments and results are discussed in Sect. 4. We conclude in Sect. 4.

2 Related Work

The field of pedestrian detection has been paid much attention since the wide application of video surveillance and automatic driving. Methods in pedestrian detection can be divided into two categories: traditional handcrafted feature-based methods and deep learning feature-based ones. Viola et al. [19] proposed a detector using both motion and appearance information; the detector is trained through AdaBoost to detect a pedestrian. The combination of multiple sources of handcrafted features in a single model has been proven to be robust and efficient. Dalal et al. [3] developed grids of Histograms of Oriented Gradient (HOG) descriptors and linear SVM-based human detection and the experimental results had shown that HOG features outperformed most other handcrafted features for pedestrian detection. Part-based models are broadly adopted in resolving the pedestrian and object detection issue [2, 8, 18]. A discriminatively trained part-based model is proposed by Felzenszwalb et al. [8]; bounding boxes around objects are trained to detect and localize objects from generic categories in static images. To achieve that, a collection of parts is arranged in a deformable configuration, latent variables are predefined for the part locations, and parts filter is set at twice resolution of the root filter. Azizpour et al. [2] utilized deformable part-based models with additional supervision and explored strong supervised models in terms of annotated object parts; the experiments show that significant improvements had been achieved compared with other detectors such as LSVM [8].

As for deep learning features, it has recently boosted the progress of many fields in machine vision including pedestrian detection [1, 14, 15, 18, 20]. A discriminative deep model is presented by Ouyang et al. [13] for detecting pedestrian with occlusion handling, with which the visibility relationship can be learned among overlapping parts at multiple layers. A region proposal network (RPN) that shares full-image convolutional features with the whole detection network is introduced by Ren et al. [17, 21], which are trained end-to-end to generate high-quality region proposals. In [1], Angelova et al. proposed a real-time approach by exploiting the efficiency of cascade

classifiers with the accuracy of deep neural networks; the proposed approach is able to run in real time and achieves high accuracy. Tian et al. [18] paid attention to the transfer learning of learned features of ConvNet to pedestrians and developed the DeepParts which consists of extensive part detectors. Apart from these detection and segmentation work, most of the object detection methods are facing the general object, not specified for pedestrian or human, such as Regions with Convolutional Neural Networks (R-CNN) [10] and YOLO [16] et al.

Based on the above traditional handcrafted features and deep features trained by deep learning neural networks for pedestrian, various classification methods including statistical methods and deep learning discriminative approaches are employed. In this paper, we use the deep auto-encoding techniques to learn the representation for the regions for pedestrians and non-pedestrian. Alternatively, we calculate the distance between the learned features of candidates and templates rather than training a new classifier to distinguish the true positive candidates from negative ones.

3 The Proposed Algorithm

3.1 Siamese Network Based on Auto-encoding

We perform the Siamese network learning in the training process, where the positive ROIs are from the annotations of the training set, and the negative ones extracted by sliding windows on multiple scales. On one hand, auto-encoder as the basic component is to use the current network to infer a coding representation of the input data, through encoder and decoder to learn the optimal value of weights and biases. The optimization goal is to reconstruct the original image using the learned coding representation. An auto-encoder is to make the output as possible as close to the input, and the parameters of the network are usually optimized by the backpropagation algorithm. On the other hand, the Siamese network structure enables the weights that could be shared during the training of the coding for two input images. The structure of the Siamese network is presented as in Fig. 1, where \mathbf{X}_1 and \mathbf{X}_2 denote the two separate input raw image patches, W represents to the sharing weights, $AC(\mathbf{X}_1)$ and $AC(\mathbf{X}_2)$ denotes to the auto-encoding vectors for the original input images. The final output is the distance between

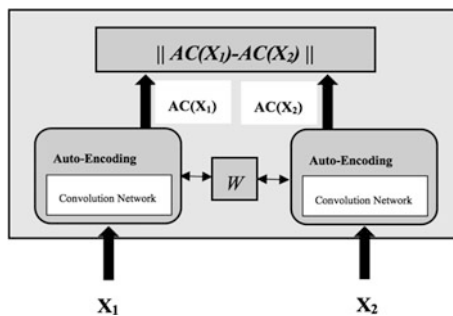


Fig. 1. The network structure of the proposed algorithm

these two input image regions. This network is trained to minimize the global loss of the total output. The global loss is calculated based on the minimization of distance between true positive samples and on the maximization of distance between true positive sample with false positive sample or true negative sample.

3.2 Distance-Based Classification

In the training step, image pairs of true pedestrian regions with annotation are put into the deep model as positive samples, and pairs of true pedestrian region and false ones are taken as negative samples. Through the optimization of the network weights, the features of true pedestrian regions are much closer and features between true pedestrian regions and false pedestrian regions are distributed far away in the encoded new feature space.

In the test step, sliding window is widely used in ROI extraction for object detection problem [5, 6, 17, 21], and we adopted the sliding window technology in multiple resolution levels to get ROI candidates. These ROI candidates are further put into the deep model as inputs. After the encoding features for candidates are extracted, the distance between the candidates and the annotated true positive sample are measured to verify whether the candidate ROI is pedestrian or not. The distance-based classification can be viewed as K-Nearest Neighbor like process. Due to the large amount of the candidates, hashing similarity [9] is additionally performed to achieve high efficiency during the comparison.

4 Experiment

4.1 Datasets and Configuration

The training sets in Caltech Pedestrian [5] dataset are selected as the training set, and the performance is validated on the testing set. Moreover, the INRIA [3] and ETH [7] datasets are often used for verifying the generalization capability of the models; our algorithm is also tested on these two datasets.

Following the settings in this section, our model is trained on the positive and negative images in the Caltech training set. The models are evaluated on the partial testing images in Caltech, INRIA, and ETH, evaluated by Missing Rate. We chose to use the 3×3 filters with four convolutional layers, and the channel size is set to 16, 8, and 8. Rectified Linear Unit (ReLU) [11] is employed as the activation function of hidden units and 2×2 max-pooling is used. The full connected structure is adopted to build the convolutional layers and the downsampling layer to compose the encoder, the decoder is composed of deconvolution layers and upsampling layer. As shown in Fig. 1, to achieve the optimization goal, the coding features are connected to the final decision layer.

4.2 Experimental Results

To investigate the proposed algorithm in terms of candidate quality, experiments are performed and evaluated by the recall rates under different Intersection over Union (IoU) thresholds. On average, 1, 4, or 100 candidates per image are evaluated. Figure 2 shows that in general our detection algorithm performs better than its counterparts, HOG + SVM [3], and ACF [4]. With 100 proposals per image, more than 70% recall has been achieved by our proposed detection at an IoU of 0.5.

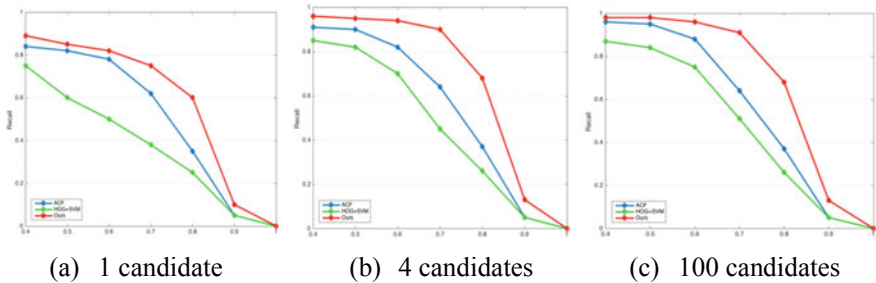


Fig. 2. Number of candidates and corresponding recall (Ours vs. HOG + SVM [3] vs. ACF [4])

The comparison of HOG + SVM, ACF, and our algorithm is evaluated by missing rate (MR) across the Caltech, INRIA, and ETH datasets; all the indices are listed in Tables 1, 2, and 3. From the detection results, we can figure out that our proposed pedestrian detection method has a robust performance and good generalization across extensive datasets.

Table 1. Comparison of different detection methods on the selected subset of Caltech

Method	MR (%)	Ref.
HOG + SVM	47.1	[3]
ACF	33.6	[4]
Ours	22.5	Ours

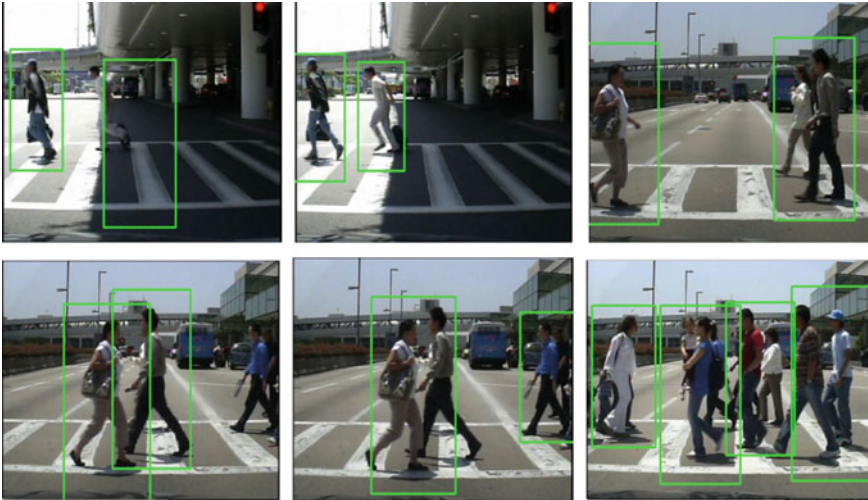
Table 2. Comparison of different detection methods on the selected subset of INRIA

Method	MR (%)	Ref.
HOG + SVM	39.3	[3]
ACF	31.2	[4]
Ours	21.5	Ours

Table 3. Comparison of different detection methods on the selected subset of ETH

Method	MR (%)	Ref.
HOG + SVM	43.2	[3]
ACF	33.1	[4]
Ours	21.9	Ours

From the examples of the pedestrian detection results shown in Fig. 3, it is easy to figure out that the disadvantages existing in the proposed method: (i) For some frames in the video, the location of the segmented candidates may be not corresponding to the exact area of the target pedestrian; (ii) When there are several targets being nearby, the number of the segmented candidates may be inaccurate, and this problem would become more prominent in the case of the increment of pedestrian; (iii) The pedestrian far away from the camera would have a larger probability of missing detected. Additionally, the resolution is chosen manually which is highly dependent on the initial parameters setting. To promote the usage in intelligent multicamera video surveillance and real-world application such as automatic driving, the problems mentioned above still have to be resolved.

**Fig. 3.** Examples of the pedestrian detection results on Caltech pedestrian dataset

5 Conclusion

In this paper, we present a novel pedestrian detection deep neural network framework via auto-encoding and distance calculation. The Siamese framework is adopted to optimize the weights and parameters of the network. Detection performance is improved through the learning of representation for true positives and non-pedestrian regions. To handle the high cost of training-specific discriminative classifier for pedestrian detection, we target on the learning of suitable features for pedestrian detection representation. Extensive experimental evaluation shows that the proposed framework can significantly improve the detection accuracy with high recalls and low missing rate. Our approach is more efficient than the comparison methods. In the future, we would focus on improving the accuracy of the detected pedestrian number and pedestrian detection under low-resolution scenes.

Acknowledgements. This paper is supported by Beijing NOVA Program (Z181100006218041) and National Key R&D Program of China (2017YFC 0820106).

References

1. Angelova, A., Krizhevsky, A., Vanhoucke, V., Ogale, A.S., Ferguson, D. (eds.): Real-time pedestrian detection with deep network cascades. In: BMVC (2015)
2. Azizpour, H., Laptev, I. (eds.): Object detection using strongly-supervised deformable part models. In: European Conference on Computer Vision. Springer (2012)
3. Dalal, N., Triggs, B. (eds.): Histograms of oriented gradients for human detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005). IEEE (2005)
4. Dollár, P., Appel, R., Belongie, S., Perona, P.: Fast feature pyramids for object detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **36**(8), 1532–1545 (2014)
5. Dollar, P., Wojek, C., Schiele, B., Perona, P.: Pedestrian detection: An evaluation of the state of the art. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**(4), 743–761 (2012)
6. Dollár, P., Wojek, C., Schiele, B., Perona, P. (eds.): Pedestrian detection: a benchmark. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009). IEEE (2009)
7. Ess, A., Leibe, B., Van Gool, L. (eds.): Depth and appearance for mobile scene analysis. In: 2007 IEEE 11th International Conference on Computer Vision (ICCV 2007). IEEE (2007)
8. Felzenszwalb, P.F., Girshick, R.B., McAllester, D., Ramanan, D.: Object detection with discriminatively trained part-based models. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(9), 1627–1645 (2010)
9. Gionis, A., Indyk, P., Motwani, R. (eds.): Similarity search in high dimensions via hashing. *Vldb* (1999)
10. Girshick, R., Donahue, J., Darrell, T., Malik, J. (eds.): Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2014)
11. Krizhevsky, A., Sutskever, I., Hinton, G.E. (eds.): Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems (2012)

12. Masci, J., Meier, U., Cireşan, D., Schmidhuber, J. (eds.): Stacked convolutional auto-encoders for hierarchical feature extraction. In: International Conference on Artificial Neural Networks. Springer (2011)
13. Ouyang, W., Wang, X. (eds.): A discriminative deep model for pedestrian detection with occlusion handling. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE (2012)
14. Ouyang, W., Wang, X. (eds.): Joint deep learning for pedestrian detection. In: 2013 IEEE International Conference on Computer Vision (ICCV). IEEE (2013)
15. Paisitkriangkrai, S., Shen, C., van den Hengel, A.: Pedestrian detection with spatially pooled features and structured ensemble learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(6), 1243–1257 (2016)
16. Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (eds.): You only look once: unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2016)
17. Ren, S., He, K., Girshick, R., Sun, J. (eds.): Faster R-CNN: towards real-time object detection with region proposal networks. In: Advances in Neural Information Processing Systems (2015)
18. Tian, Y., Luo, P., Wang, X., Tang, X. (eds.): Deep learning strong parts for pedestrian detection. In: Proceedings of the IEEE International Conference on Computer Vision (2015)
19. Viola, P., Jones, M.J., Snow, D. (eds.): Detecting pedestrians using patterns of motion and appearance. IEEE (2003)
20. Zeng, X., Ouyang, W., Wang, X. (eds.): Multi-stage contextual deep learning for pedestrian detection. In: 2013 IEEE International Conference on Computer Vision (ICCV). IEEE (2013)
21. Zhang, L., Lin, L., Liang, X., He, K. (eds.): Is faster R-CNN doing well for pedestrian detection? In: European Conference on Computer Vision. Springer (2016)
22. Zhang, S., Benenson, R., Omran, M., Hosang, J., Schiele, B. (eds.): How far are we from solving pedestrian detection? In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2016)



Robust Model for Chinese License Plate Character Recognition Using Deep Learning Techniques

Amr Abdussalam^{1,2,3(✉)}, Songlin Sun^{1,2,3}, Meixia Fu^{1,2,3},
Yasir Ullah^{1,2,3}, and Safwan Ali^{1,3}

¹ School of Information and Communication Engineering, Beijing, China
helloamr2016@yahoo.com

² Key Laboratory of Trustworthy Distributed Computing and Service, Beijing, China

³ Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract. Character recognition and classification is considered one of the most important parts of current LPR systems. Because of low recognition quality and poor robustness of traditional character recognition techniques, those techniques were gradually replaced by powerful deep learning modules such as convolutional neural networks. Convolutional neural networks (CNNs) show satisfying ability in character recognition and outperform most of other available models. Since Chinese license plates contain Chinese characters in addition to ordinary alphanumeric characters, a robust, powerful, and efficient CNN is needed to accomplish character recognition task efficiently. In this paper, we have proposed an efficient CNN model based on Darknet architecture to perform character recognition. Through convolutional and max pooling layers, features of input character images will be extracted and then sent to softmax layer for classification. To avoid overfitting problem in the training process, the dropout regularization technique is adopted. We have used a dataset of 84,000 character images for training and testing our model. The experimental result shows satisfactory outputs and eventually achieves test accuracy of 99.69%.

Keywords: Character recognition · CNN · License plate recognition · Dataset · Darknet · Softmax

1 Introduction

License plate recognition system (LPR) [1, 2] is a key component of intelligent transportation system ITS. It has been extensively used in numerous applications such as traffic law enforcement, unattended parking lots, and so on. Although there are many techniques and methods which have already been proposed for recognizing Chinese license plates, the existence of Chinese characters in these plates still makes the recognition task more challenging and cumbersome. The special characteristics of Chinese characters stimulated many researchers to modify existing algorithms and invent new methods to increase and enhance performance and accuracy of Chinese LPR systems. Character recognition and classification module is an important

component of LPR systems. This module is responsible for identifying and classifying all characters of license plates. Therefore, improving and optimizing performance and accuracy of character recognition modules will surely enhance and increase performance and accuracy of whole LPR systems.

In recent years, researchers have implemented many techniques for license plate character recognition. Template matching [3–5] is considered the most popular method used in character recognition area. In template matching, similarities between template characters and input characters are calculated. Template character with highest similarity with input character will be selected as recognition outcome. Another method used for character recognition is the feature matching method [6, 7]. In feature matching method, a comparison between extracted character features and standard features is performed, and characters are identified by similarity.

In addition, several CNN [8, 9]-based character recognition models have been suggested. One of these models is proposed in [10]. It consists of seven-layer network architecture with three convolutional layers, two pooling layers, and two fully connected layers. A coarse-to-fine strategy is used for extracting features of input character images. Another CNN-based character recognition model called shared hidden layers CNN (SHL-CNN) is introduced in [11]. In this model, a CNN model which has shared hidden layers and two distinct softmax layers for Chinese and alphanumeric characters have been implemented to solve the problem of Chinese characters shortage.

The contribution in this paper is that we have proposed an efficient CNN model based on Darknet architecture [12] to perform character recognition task. Proposed architecture contains 18 convolutional layers, five max pooling layers, and one softmax layer which achieved satisfactory results in character recognition. Dropout regularization technique is used to avoid overfitting during training process. We have employed a dataset of 84,000 character images for training and testing our CNN model. Experimental result shows satisfactory results of our CNN model with high accuracy and excellent performance.

2 Proposed Model

Generally, Chinese license plates contain 31 kinds of Chinese characters, 10 kinds of numbers, and 24 kinds of English letters (except “O” and “I”). Typically, every Chinese license plate contains only one Chinese character representing the province issuing the license plate, while the remaining characters of license plate are alphanumeric characters. Typically, for one Chinese license plate, the ratio between Chinese characters and alphanumeric characters is almost 1:6. So, it is obvious that number of samples of alphanumeric characters is often much larger than number of samples of Chinese characters. Therefore, powerful, efficient, and robust character classification architecture is required for identifying Chinese characters and classifying them correctly. Our CNN-based character recognition model uses robust architecture with excellent feature extraction ability to identify, distinguish, and classify Chinese and alphanumeric characters into 65 classes.

2.1 Architecture of CNN Character Recognition Model

Architecture of the proposed character recognition CNN is based on the “Darknet” architecture proposed in [12]. Darknet architecture consists of a sequence of convolutional and max pooling layers in which the numbers and parameters were set in a manner that accelerates training process and improves feature extraction. Our proposed character recognition CNN architecture uses 18 convolutional layers and five max pooling layers, and will be given the name “Darknet-18”. A full description of Dark-18 architecture is shown in Table 1.

Table 1. Description of Darknet-18 architecture

Type	Filters	Size/Stride	Output
Convolutional	32	3×3	96×96
Max pool		$2 \times 2/2$	48×48
Convolutional	64	3×3	48×48
Max pool		$2 \times 2/2$	24×24
Convolutional	128	3×3	24×24
Convolutional	64	1×1	24×24
Convolutional	128	3×3	24×24
Max pool		$2 \times 2/2$	12×12
Convolutional	256	3×3	12×12
Convolutional	128	1×1	12×12
Convolutional	256	3×3	12×12
Max pool		$2 \times 2/2$	6×6
Convolutional	512	3×3	6×6
Convolutional	256	1×1	6×6
Convolutional	512	3×3	6×6
Convolutional	256	1×1	6×6
Convolutional	512	3×3	6×6
Max pool		$2 \times 2/2$	3×3
Convolutional	1024	3×3	3×3
Convolutional	512	1×1	3×3
Convolutional	1024	3×3	3×3
Convolutional	512	1×1	3×3
Convolutional	1024	3×3	3×3

The outputs of the last convolutional layer of Darknet-18 must be reshaped and flattened to a single vector which can be used as an input to the softmax layer. Softmax layer of this CNN has 65 classes representing all characters of Chinese license plate (31 Chinese characters, 10 numbers, and 24 English letters). A dropout regularization technique is used after the flattened vector to reduce overfitting and enhance training and accuracy of CNN model. The whole structure of the character recognition CNN is depicted in Fig. 1.

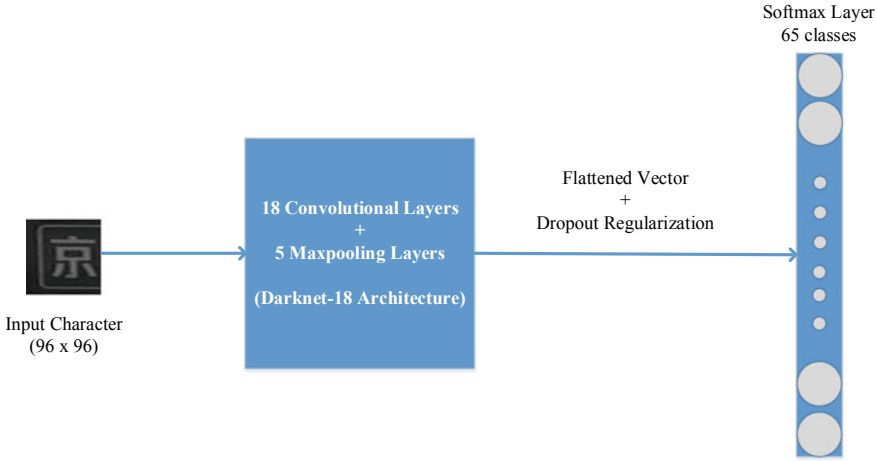


Fig. 1. Structure of character recognition CNN model

2.2 Training of Character Recognition CNN Model

In the training process of character recognition CNN, cross-entropy loss function is used to measure differences between predicted and true classes. Adam optimizer is used to optimize loss function. The learning rate for Adam optimizer is 0.0001 which gives better optimization by experiment. The training batch size was 50 images per batch, and the training process lasts for more than 600 epochs.

3 Experimental Results

In this section, we will describe the dataset used for training and testing proposed character recognition CNN module. In addition, the classification results and accuracy of character recognition CNN will also be introduced.

3.1 Used Dataset

For achieving reliable and satisfying results, datasets used for training and testing character recognition CNN must be prepared well and labeled correctly. Character samples were taken directly from Chinese license plate images. License plate images were preprocessed before splitting them into individual characters. First, RGB to grayscale conversion and skew correction were performed on license plates, and then those license plates were divided manually into their individual characters. Character sample images were reshaped into 96×96 and converted into grayscale to simplify

structure of character recognition CNN model. After that the labels were added to each character image so that they become ready for training and testing CNN. The labeling process was performed manually with the assistance of MATLAB software.

The total dataset used for training and testing character recognition CNN module is about 84,000 character images. About 70,000 character images were used as a training set, and 14,000 character images were used as a testing set. We must point out that the selection process of the training set images and testing set images was performed randomly to avoid subjective effects. Table 2 and Fig. 2 show details about distribution of used dataset.

Table 2. Distribution of characters used in dataset

	Training set	Testing set	Total set
Letter	21,019	4235	25,254
Number	38,981	7765	46,746
Chinese	10,000	2000	12,000
Total	70,000	14,000	84,000

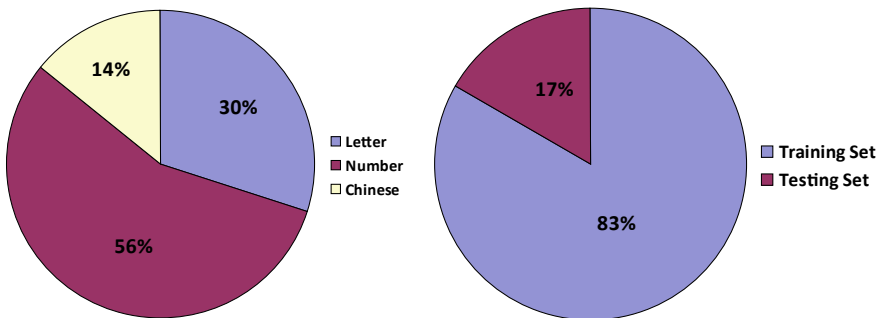


Fig. 2. Composition of Dataset

3.2 Achieved Results

The test set used for evaluating accuracy of character recognition CNN model consists of 14,000 character images. According to testing results, the maximum achieved test accuracy of character recognition CNN is 99.69%—(13957/14000). This maximum accuracy is achieved after more than 550 epochs of training. The results show that among 14,000 test images, 13,957 images were truly evaluated by the character recognition CNN, while only 43 images were falsely evaluated. Figure 3 shows test accuracy of character recognition CNN for every 50 epochs of training.

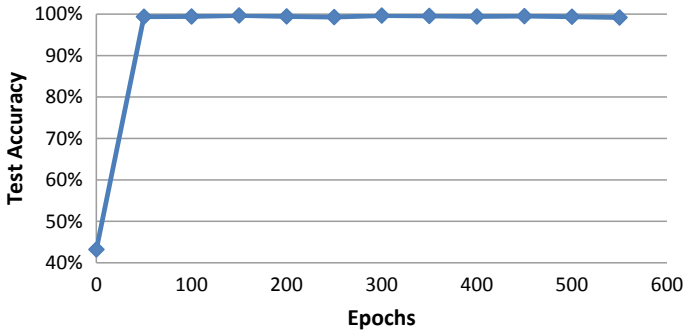


Fig. 3. Test accuracy versus training epochs

4 Conclusion

In this paper, we have introduced a robust, powerful, and efficient character recognition CNN-based model. By adopting architecture similar to Darknet-19 and using dropout regularization technique, learning and feature extraction ability of our CNN model have been greatly improved. Experimental results show that our character recognition CNN model has achieved satisfying classification ability and high recognition accuracy so that it can be used to increase performance of current LPR systems.

In future, our work can be improved by adopting dataset augmentation methods such as blurring and rotating sample characters to increase the size of datasets and improve test accuracy.

References

1. Ahmad, I.S., Boufama, B., Habashi, P., Anderson, W., Elamsy, T.: Automatic license plate recognition: a comparative study. In: 2015 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Abu Dhabi, pp. 635–640 (2015)
2. Bakhtan, M.A.H., Abdullah, M., Rahman, A.A.: A review on license plate recognition system algorithms. In: 2016 International Conference on Information and Communication Technology (ICICTM), Kuala Lumpur, pp. 84–89 (2016)
3. Tsukumo, J.: Handprinted Kanji character recognition based on flexible template matching. In: Proceedings of the 11th IAPR International Conference on Pattern Recognition. Vol. II. Conference B: Pattern Recognition Methodology and Systems, The Hague, pp. 483–486 (1992)
4. Baten, R.A., Omair, Z., Sikder, U.: Bangla license plate reader for metropolitan cities of Bangladesh using template matching. In: 8th International Conference on Electrical and Computer Engineering, Dhaka, pp. 776–779 (2014)
5. Xing, X.H., Guo-Hua, G.U.: Method of quickly recognizing vehicle plate based on pattern matching and characteristic dot matching. *Optoelectr. Technol.* **23**(4), 268–270 (2003)
6. Muhammad, A., Haswadi, H., Dewi, N., Habibollah, H.: A review on feature extraction and feature selection for handwritten character recognition. *Int. J. Adv. Comput. Sci. Appl.* **6**(2) (2015)

7. Rahman, C.A., Badawy, W., Radmanesh, A.: A real time vehicle's license plate recognition system. In: 2003 Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, pp. 163–166 (2003)
8. Girshick, R.: Fast R-CNN. In: 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, pp. 1440–1448 (2015)
9. He, K., Zhang, K., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, pp. 770–778 (2016)
10. Yao, D., Zhu, W., Chen, Y., Zhang, L.: Chinese license plate character recognition based on convolution neural network. In: 2017 Chinese Automation Congress (CAC), Jinan, pp. 1547–1552 (2017)
11. Liu, Y., Huang, H.: Car plate character recognition using a convolutional neural network with shared hidden layers. In: 2015 Chinese Automation Congress (CAC), Wuhan, pp. 638–643 (2015)
12. Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, pp. 6517–6525 (2017)



A Multi-label Scene Categorization Model Based on Deep Convolutional Neural Network

Gaofeng Zhao¹, Wang Luo^{1(✉)}, Yang Cui¹, Qiang Fan¹,
Qiwei Peng¹, Zhen Kong¹, Liang Zhu², and Tai Zhang³

¹ NARI Group Corporation (State Grid Electric Power Research Institute),
Nanjing, China

luowang@sgepri.sgcc.com.cn

² State Grid Hunan Electric Power Company Limited, Changsha, China

³ State Grid Sichuan Electric Power Company Limited, Chengdu, China

Abstract. Being one of the most fundamental embranchments of deep learning theory, scene categorization technology has been extensively researched because of its great value in engineering application, especially in the field of remote monitoring and intelligent fault detection. To bridge the gap between theoretical accuracy and practical performance of relevant classification models which is mainly caused by nonstandard labeling information, this paper builds a normative dataset composed of 10,000 high-quality manual labeled images from the power sector, and proposes a high-performance multi-label classification model utilizing deep convolutional neural network (CNN) inspired by Inception-v4 [1] on this basis. Experiments demonstrate that the model proposed achieves an accuracy of 94.125% on the test set and thus can be deployed into practical intelligent surveillance scenarios.

Keywords: Multi-label · Scene categorization · CNN

1 Introduction

Before the prevalence of deep CNN, there were several mature business solutions in image classification based on machine learning algorithms, such as the successful application of LeNet in handwriting character recognition [2]. However, restrained by the deficiency of computing resources and the absence of efficient model training means, large-scale application of automatic scene categorization technology was not a preferred solution in intelligent surveillance, for example, manual fault detection has been a predominant monitoring means in power sector for decades, which is of unsatisfying real-time capability and accuracy.

Owing to the productive development of deep learning theory, training deep neural networks with billions of undetermined parameters becomes feasible [3], and the breakthrough in convolutional network structure significantly improve the performance of classification model, laying a solid theoretical foundation for the commercial application of scene categorization models. Furthermore, with the large-scale implementation of high-definition cameras and automatic line inspection equipment like

unmanned aerial vehicles, building diverse high-quality dedicated image dataset is now practicable and economical.

This paper focuses on the automatic scene categorization problem in the power sector. Compared to general image classification tasks, scene categorization in the power sector is challenged by the redundancy of homogeneous images and the disproportion of label distribution. Besides, high demands of efficiency, accuracy, flexibility, scalability, and robustness are imposed to lower the cost of deployment and maintain categorization performance in a complicated environment.

To meet the requirements, this paper proposed a multi-label image classification model utilizing convolutional network structure enhanced by inception modules, which meets the state-of-the-art performance. Compared with the existing methods, the model proposed is reinforced from three aspects. First, as the quality of train set is a significant determinant of classification accuracy [4], we established a manually labeled dataset selected from more than 30,000 images of 12 typical scenes in the power sector. As shown in Fig. 1, all images are scaled into 400×400 pixels with multiple scene labels. Second, considering the absolute number of images in our dataset is far less than that of some public dataset like ImageNet, we doubled the number of a portion of high-quality images to expand our dataset, as well as to mitigate the impact of label distribution imbalance to strengthen the generalization ability of the classification model. Lastly, we adopted the inception module, an efficient sparse convolution structure, to reduce computational complexity remarkably while maintaining excellent classification performance.



Fig. 1. Example images of our multi-label dataset

2 Related Work

This section presents a concise review of the evolution of CNN and inception module that have yielded extraordinary achievements in image classification, oriented by the development of scene categorization technology.

Scene categorization. In traditional machine learning methods, as same as the other image classification tasks, the mainstream approach of scene categorization uses Support Vector Machine (SVM) to learn classification rules [5] from low-level and mid-level hand-crafted features, such as gradient histogram and speed-up robust features [6]. However, due to the limited characterization ability of hand-crafted features,

known as semantic gap, traditional methods are not competent to handle intra-class variations [4]. With the ascendance of deep learning theory in image classification field, convolutional network structure gradually becomes predominant in scene categorization models, which extracts high-level semantic features of images automatically through efficient model training algorithms like stochastic gradient descent and thus gives categorization models powerful expression capabilities [7], resulting in their excellent classification performance and robust generalization ability.

CNN. Inspired by the pioneering work of Hubel and Wiesel in neural network [8], Fukushima constructed an original structure of Artificial Neural Network (ANN), Neocognitron [9], which is an embryonic form of CNN. However, due to the deficiency of feature extraction means and model training algorithm, Neocognitron has not become a mainstream solution. In the 1990s, Lecun et al. summarized previous researches and accordingly proposed a typical CNN structure characterized by local receptive field, pooling layer, and shared weights [2], which was the best solution in handwriting character recognition. In 2006, Hinton et al. demonstrated an efficient training algorithm in feedforward neural networks evolved from restricted Boltzmann machine [3], laying a solid foundation of deep neural network. Afterward, Krizhevsky et al. built a deep CNN named AlexNet that achieved unprecedented classification accuracy in ImageNet in 2012 [10], marking the maturity of CNN.

Inception. The typical structure of CNN that continuous convolutional layers with pooling and local contrast normalization layers are followed by several fully connected layers has been established by LeNet-5 [11], and to acquire superior classification capability, most recent researches concentrate on deepening network structure by adding convolutional layers or utilizing the structure of network in network [12]. Considering deeper network tends to significantly increase the computational complexity, Szegedy et al. focuses on broadening the width of the network by proposing a module named inception, in which dense convolutional structure is replaced by sparse one as Fig. 2 shows to accelerate training process. Subsequently, Ioffe et al. proposed three improved versions of inception, Inception-v2 adopts batch normalization layer to inhibit internal covariate shift [13], Inception-v3 increases the nonlinear level of the network via factorization technique [14], and Inception-v4 introduces residual connection structure for further improvement model performance [1].

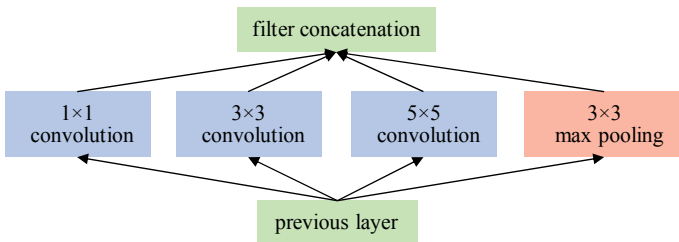


Fig. 2. The structure of inception module [11]

3 Proposed Method

This section demonstrates the details of the proposed scene categorization model, including the physical structure of convolutional network and the definition of loss function, as well as the concrete process of building a high-quality train image dataset.

Dataset. Since there is no existing high-quality dataset consisting of images from the power sector, we establish a dedicated database to train the categorization network, for the quality of dataset is critical to model the performance. We select 10,000 images with 12 kinds of typical scenes from more than 30,000 samples captured by high-definition surveillance device, and each image is tagged by specialists with a 12-digit label indicating the presence or absence of each scene. To increase the variety of our dataset while maintaining acceptable computational complexity, we resize the images into 400×400 pixels and generate 299×299 crops randomly during training phase. Moreover, as Table 1 shows the range between the absolute number of different labels is so large that will cause severe under-fitting or over-fitting problem, so we redouble some of the selected images to obtain uniformly distributed labeled images. Notably, the doubling operation is after the division of train set and test set to guarantee their independence.

Table 1. The number of images with different labels

Label	Original number	Resulting number	Label	Original number	Resulting number
Transmission line	2741	2741	Bird nest	258	1806
Insulating switch	1949	1949	Tree	1850	1850
Illegal building	160	1600	Frozen wire	196	1960
Unclosed switch	188	1880	Meter	1419	1419
Rusty device	170	1700	Transformer	711	1422
Insulator flashover	155	1550	Substation	4276	4276

CNN architecture. The architecture of our deep CNN is inspired by Inception-v4 network [1], a state-of-the-art classification model. As Fig. 3 shows the physical structure of our scene categorization network is similar to typical CNN structure that several stacked convolutional layers optional with enhancement technologies like pooling and Rectified Linear Unit (ReLU) activation are followed by a few fully connected layers, furthermore, to improve categorization accuracy, standard convolutional layers are replaced by network-in-network modules including three kinds of inception components and two kinds of reduction components, which are evolved from the original inception module as shown in Fig. 2. These variations of the original inception module are reinforced by the implementation of residual connection [15] and more flexible choices of convolutional kernels, as Table 2 shows, while the absolute number of convolutional kernels in different network-in-network components is

notably large, due to the well-organized network architecture and the preference of small convolutional kernels like 1×1 and 1×3 ones, the model is of high training efficiency and outstanding classification accuracy.

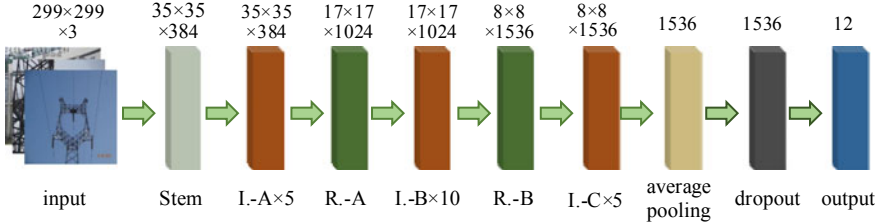


Fig. 3. The architecture of our deep CNN. I is short for inception, R is short for reduction

Loss function. To overcome the learning slowdown problem caused by the derivative term of traditional loss functions like quadratic cost with Sigmoid activation function, we adopt cross-entropy cost function to accelerate gradient backpropagation. Assuming that $Y = \{y_1, y_2, \dots, y_{12}\}$ denotes the actual label of the input image X , in which $y_n = 1$ indicates X belongs to the category n while $y_n = 0$ indicates the opposite, $A = \{a_1, a_2, \dots, a_{12}\}$ denotes the output of our classification model when the input image is X , we can define the cross-entropy loss as the following Eq. 1.

$$Loss = -\frac{1}{12} \sum_{i=1}^{12} [y_i \ln a_i + (1 - y_i) \ln(1 - a_i)] \quad (1)$$

Since the output of our model can be arbitrary values, binarization is essential to gain the meaningful predictive label of the input image $\hat{A} = \{\hat{a}_1, \hat{a}_2, \dots, \hat{a}_{12}\}$, so we choose 0 as the threshold according to the normal practice, the binarization equation can be written as Eq. 2.

$$\hat{a}_n = \begin{cases} 0, & a_n < 0 \\ 1, & \text{otherwise} \end{cases} \quad (2)$$

4 Experiment

In this section, we illustrate the deployment details of our scene categorization model and analyze the experimental results to evaluate it.

Deployment details. The size of the input image of our scene categorization model should be 299×299 pixels, and to strengthen noise resistance, we resize the images into 400×400 pixels and sample their 299×299 sub-images randomly as input images via crop generation technology, rather than input the entire resized images to our network directly. Moreover, ReLU activation function is adapted to avoid output

Table 2. Number of convolutional kernels in different network in network components, cons. is short for convolutional kernels

	Inception-A	Inception-B	Inception-C	Reduction-A	Reduction-B
Number of cons.	1088	2272	3264	1056	1056

saturability, dropout layer with a drop rate of 50% is adopted to strengthen generalization ability, and our dataset is divided into two independent parts as train set and validation set, with a proportion of 90 and 10% separately. To efficiently train the network, the initial learning rate is set to 0.00001 with step updating strategy, momentum is set to 0.95, weigh decay is set to 0.0005, and batch size is set to 32. Besides, to accelerate the training process, we choose high-performance GPU to train the network.

Results analysis. During the test phase, accuracy is strictly defined as the percentage of test images that are correctly categorized in each of the 12 labels, and after 100,000 iterations of train, we obtain a competent scene categorization model that achieves an overall accuracy of 94.125%, exceeding the 90% baseline of the technique request in the field of intelligent surveillance. Moreover, when considering the categorization performance of single scene, as Table 3 shows, the categorization accuracy rises dramatically to near 98%, even 100% in insulator flashover and frozen wire, this result is counterintuitive but can be exemplified by the sources of our image dataset. As been discussed, we choose images mainly from surveillance data of power sector, and similar sources will inevitably lead to strong correlation between different images, more specifically, though stochastic partition yields nonoverlapping train set and test set, the close resemblance between them results in their virtual overlap, for example, the contents of different insulator flashover images are almost the same in practice, partially explaining its 100% test accuracy. In ordinary classification tasks, the strong correlation between train set and test set will reduce the credibility of experimental results, however, our scene categorization network is a dedicated model instead of a universal

Table 3. The accuracy and missing alarm rate of each label

Label	Accuracy	Missing alarm rate	Label	Accuracy	Missing alarm rate
Transmission line	0.999375	0.00155	Bird nest	0.999375	0.00000
Insulating switch	0.988750	0.00852	Tree	0.986250	0.00866
Illegal building	0.974375	0.00743	Frozen wire	1.000000	0.00000
unclosed switch	0.986875	0.00068	Meter	0.998750	0.00125
Rusty device	0.980625	0.00753	Transformer	0.988750	0.01010
Insulator flashover	1.000000	0.00000	Substation	0.999375	0.00062

one, it will only be deployed into surveillance scenes in power sector, meaning that its input images are confined to highly correlated sample images with similar backgrounds, therefore, the experimental results do demonstrate our model is a competent dedicated model in power sector.

5 Conclusion

This paper establishes a dedicated image dataset of 10,000 high-quality images from the power sector, and proposes a multi-label scene categorization model on this basis. This categorization model is a concrete application of classical algorithms of deep learning theory such as error backpropagation and stochastic gradient descent, as well as is inspired by cutting-edge achievements like Network in Network architecture and inception module. Since our categorization model reaches an outstanding classification accuracy of 94.125% with strong robustness and scalability, it is feasible to deploy it into intelligent surveillance and fault detection scenarios in power sector, and it can be applied into other classification tasks conveniently via transfer learning technique.

Acknowledgements. This research was supported by the project as follows: Science and Technology Project of SGCC “Research on feature recognition and prediction of typical ice and wind disaster for transmission lines based on small sample machine learning method”.

References

1. Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.A.: Inception-v4, inception-ResNet and the impact of residual connections on learning. In: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, pp. 4278–4284 (2017)
2. Lecun, Y., Bottou, L., Bengio, Y., Ha, P.: Gradient-based learning applied to document recognition. *Proc. IEEE*, 1–46 (1998)
3. Hinton, G.E., Osindero, S., Teh, Y.-W.: A fast learning algorithm for deep belief nets. *Neural Comput.* **18**(7), 1527–1554 (2006)
4. Zhu, J., Liao, S., Lei, Z., Li, S.Z.: Multi-label convolutional neural network based pedestrian attribute classification. *Image Vis. Comput.* **58**, 224–229 (2017)
5. Yuan, L., Chen, F., Zhou, L., Hu, D.: Improve scene classification by using feature and kernel combination. *Neurocomputing* **170**, 213–220 (2015)
6. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: SURF: speeded up robust features. *Comput. Vis. Image Underst.* **110**, 346–359 (2008)
7. Nogueira, K., Penatti, O.A.B., Jefersson, A.: Towards better exploiting convolutional neural networks for remote sensing scene classification. *Pattern Recognit.* **61**, 539–556 (2017)
8. Hubel, B.Y.D.H., Wiesel, A.D.T.N.: Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *J. Physiol.* 106–154 (1962)
9. Fukushima, K.: Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern.* **36**(4), 193–202 (1980)
10. Krizhevsky, A., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **25**(2), 1–9 (2012)

11. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. *Comput. Vis. Pattern Recognit.* 1–9 (2015)
12. Lin, A., Chen, Q., Yan, S.: Network in network. *Comput. Sci.* 1–10 (2013)
13. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. 448–456 (2015)
14. Szegedy, C., Vanhoucke, V., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. 2818–2826 (2015)
15. He, Z., Zhang, X., Ren, S., Sun, S.: Deep residual learning for image recognition. 770–778 (2015)



Discriminative Structured Dictionary Learning for Face Recognition

Ying Zhu^(✉)

College of Telecommunication and Information, Nanjing University of Posts and
Telecommunications, Nanjing 210003, China
1016010606@njupt.edu.cn

Abstract. For a few years, plenty of face recognition algorithms, such as deep learning, have been in hot pursuit with the trend of the technology, while dictionary learning algorithm is still out of the woods for the sake of its higher robustness to occlusion and light. In this paper, we propose to learn a discriminative structured dictionary with constraint named as multi-label to suppress representations for different classes, as well as Laplacian Eigenmaps to encourage the representations for the same class to be close to each other. Demonstrated by the results of the experiments, our proposed dictionary learning methods intend to achieve better classification performance and higher computational efficiency compared to the existing algorithms.

Keywords: Dictionary learning · Supervised learning · Face recognition · Laplacian Eigenmaps

1 Introduction

In recent years, face recognition [1] technology has been gradually applied in many fields such as identity authentication [2] and remote control because of its accuracy and convenience. Sparse representation [3] is a newly developing method that is extensively used in face recognition, which is based on the prior condition that the representation of signals on a suitable basis is sparse. This method has better robustness to noise and occlusion compared to some traditional clarification algorithms.

Many of methods [4, 5] have been provided for sparse coding using defined transformation matrices. In [5], joint Haar and Gabor basis is used as a dictionary, and the obtained sparse code is used as features to train the SVM classifier to recognize faces and numbers. However, the dictionary does not include the semantic information of the original signal that people are more interested in. The sparse representation-based classification (SRC) method is proposed in [6], and it has made a significant breakthrough in face recognition field for the reason that it creatively adopts the sparse linear combination of the training samples to represent the test images. Nevertheless, it is time consuming for SRC to sparsely encode with a large dictionary, thus it is extremely essential to learn a compact and discriminative dictionary.

Based on the aforementioned researches, a discriminative structured dictionary learning algorithm for face recognition is presented in this paper. The algorithm integrates the reconstruction error and regularization into the objective function so that the learnt dictionary contains both reconstruction and discriminative characteristics. In order to improve the accuracy of classification, and obtain a stronger ability to distinguish, we can enlarge the gap between different classes and, on the other hand, narrows the gap between similar classes. That is, let the representation coefficients of signals from the same class concentrated on their specific atoms and shared atoms, while the other atoms need to be suppressed to a certain degree. The experimental results on The Extended Yale face and The Yale Face demonstrate that the proposed approach can indeed achieve promising performance in face classification.

The organization of this paper is as follows. First, in Sect. 2, we briefly introduce the general dictionary learning model, and then review the related work on it. The objective function of the proposed algorithm is clarified in detail in Sect. 3. We conduct extensive experiments on two classical face datasets to demonstrate the effectiveness of our model in Sect. 4. Finally, we conclude our paper in Sect. 5.

2 Background

2.1 Dictionary Learning Model

Dictionary learning algorithms [7] are generally divided into two steps: the calculation of sparse coding X when dictionary D is known and the update of dictionary D when sparse coding X is calculated. Considering the general dictionary learning model, [8] which is expressed as

$$\min_{D, x_i} \sum_{i=1}^m \|y_i - Dx_i\|_2^2 + \lambda \sum_{i=1}^m \|x_i\|_1 \quad (1)$$

where $Y = [y_1, \dots, y_n] \in R^{m \times n}$ is a set of input test signals, $D = [d_1, \dots, d_k] \in R^{m \times k}$ donates the learnt dictionary, and $X = [x_1, \dots, x_n] \in R^{k \times n}$ is the sparse codes based on D .

The first cumulative term in the above formula indicates that the first goal of dictionary learning is to restore the samples as much as possible by the linear combination of dictionary matrix and sparse representation. The second term shows that the sparse codes X should be as sparse as possible. Unfortunately, the optimization problem of l_0 -norm is an NP-hard problem, and l_1 -norm has been proved theoretically to be the optimal convex approximation of l_0 -norm and that regularization of l_1 -norm is much easier to obtain sparse solutions, as a result of which l_1 -norm is extensively adopted instead of l_0 -norm. In general, the fundamental purpose of the dictionary learning algorithm is to achieve the dictionary and the sparse representation of the input samples that satisfy the above optimization problems.

So far, the sparse representation of signals is transformed into solving the sparse regular optimization problem. The overall strategy for solving the optimization problem is to alternately and iteratively optimize the learnt dictionary D and the sparse

representation X . When D is fixed (so-called sparse coding), it is an l_1 -regularized least squares problem, and usually it gets solved by the way of LASSO regularization [9], such as Proximal Gradient Design method. When X is fixed (so-called dictionary learning), it is an l_2 -norm constraint least squares problem. Plenty of approaches have been applied on it such as KSVD method [10, 11] and stochastic gradient algorithm.

2.2 Related Work

In order to obtain some classification characteristics, a great deal of algorithms have been proposed by the way of restricting dictionary or sparse encoding or both of them. Mairal et al. [9] applied softmax discriminative loss function to the reconstruction error of signal and dictionary in 2008. Yang et al. [12] applied Fisher discriminant to sparse coding in 2011. Both methods made sparse coding have smaller interclass spacing and larger outer-class spacing. In 2010, Ramirez et al. [8] add a penalty of irrelevance to the subdictionary to make the dictionaries of different classes independent of each other. By learning a sub-dictionary for each class, these algorithms can reconstruct the corresponding classes well, but they do not have such characteristics for other classes. Then, the reconstruction errors between signals and subclasses are used to classify them. But the algorithm of this model is especially time consuming when the number of classes is large.

Some other methods classify signals by training classifiers. Wang et al. [7] proposed in 2010 to learn a dictionary, and then train an SVM linear classifier with sparse coding as a feature to classify images. In 2008, Pham et al. proposed an algorithm putting the learning parameters of classifiers in the objective function and optimizing the dictionary learning by the feedback of classification accuracy. These two algorithms separate dictionary learning from the training of classifier, so that it is difficult to obtain global optimal solution. In 2011, Zhang et al. [10] proposed that the dictionary and classifier be combined to be solved by K-SVD model, but this method cannot have promising discrimination. It is proved that the sparse coding obtained by Jiang et al. [11] in 2011 increases the regularity of sparse coding discrimination, so as to ensure that similar sparse coding has a similar representation. This method classifies the detected signals directly with the classifier we have learned.

3 Optimization of the Objective Function

Supposing dictionary learning is expressed as an optimization problem, the simplest form of expression can be formulated as

$$\min_{D, x_i} \sum_{i=1}^m \|y_i - Dx_i\|_2^2 + \beta \sum_{i=1}^m \|x_i\|_2^2 \text{ s.t. } \|d_k\| = 1, \forall k \quad (2)$$

where $Y \in R^{m \times p}$ represents the facial image samples and the superscript means belonging to the c th class, $D \in R^{m \times n}$ denotes the dictionary to be learnt, $X \in R^{n \times p}$ means the sparse codes corresponding to the input image signals. The first term represents the reconstruction error of training samples, while the left part denotes the

sparse regularization constrained by coefficient β . As we mentioned above, l_p -norm applied in sparse regularization is always set as l_0 or l_1 [3] as a general rule. In order to optimize the classification efficiency without depressing the discriminative ability of classification, we employ a simpler l_2 -norm instead in our proposed algorithm.

We propose two new methods: the multi-label suppression and the Laplacian Eigenmaps to make the learnt dictionary more discriminative and constructive for classification. Inspired by [13], we will not only focus on the shared atoms of all the classes, but specific atoms that mostly bound up with some specific classes in our method. Thus, the dictionary learning model we proposed is formulated as

$$\min_{D, X} \sum_{c=1}^C \left\{ \|Y^c - DX^c\|_F^2 + \beta \sum_{j=1}^{N^c} \|X_j^c\|_2^2 \right\} + \lambda \sum_{c=1}^C \|P^c X^c\|_F^2 + \gamma \text{tr}(X \tilde{L} X^T) \quad (3)$$

s.t. $\|d_k\|_2 = 1, \forall k$

This is a general model of dictionary learning with multi-label suppression together with Laplacian Eigenmaps, where the two last terms with parameters λ and γ represent them, respectively. $P^c \in R^{K \times K}$ is adapted to the algorithm for the sake of multi-label suppression with scalar λ , and P^c is defined as follows:

$$P^c(m, n) = \begin{cases} 1, & m = n \text{ and } m \in \mathcal{L} \setminus (\mathcal{L}^0 \cup \mathcal{L}^c); \\ 0, & \text{elsewhere,} \end{cases} \quad (4)$$

where $c = 1, 2, \dots, C$ with $P^c(m, n)$ indicates the (m, n) th entry of P^c . Then

$$(P^c X_i^c)(m) = \begin{cases} x_i(m), & m \in \mathcal{L} \setminus (\mathcal{L}^0 \cup \mathcal{L}^c); \\ 0, & \text{elsewhere,} \end{cases} \quad (5)$$

in which $(P^c X_i^c)(m)$ and $x_i(m)$ denotes the m th components of the corresponding vectors, respectively. This kind of constraint makes large coefficients to basically locate on their relevant atoms, thus suppresses the occurrence of those large coefficients at label-specific atoms.

As regard of Laplacian Eigenmaps part, it is assumed that the total graph consisted of C groups that are separated from one another. Thus, the formula of the normalized Laplacian matrix can be derived as

$$\tilde{L} = \begin{bmatrix} \tilde{L}^1 & & & \\ & \tilde{L}^2 & & \\ & & \ddots & \\ & & & \tilde{L}^c \end{bmatrix} \in R^{N \times N} \quad (6)$$

where $\tilde{L}^c \in R^{N^c \times N^c}$ ($c = 1, 2, \dots, C$) denotes the normalized Laplacian of the subgraph that is mapped with the c th class. For the reason that the N^c training samples of the c th class, as well as each of the two vertices are all connected, \tilde{L}^c will be deduced as

$$\tilde{L}^c = \frac{1}{N^c - 1} \begin{bmatrix} N^c - 1 & -1 & \cdots & -1 \\ -1 & & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & N^c - 1 \end{bmatrix} \quad (7)$$

Thus, we can finally reformulate the objective function in (2) as below

$$\min_{D, X} \sum_{c=1}^C \left\{ \|Y^c - DX^c\|_F^2 + \beta \sum_{j=1}^{N^c} \|X_j^c\|_2^2 + \lambda \|P^c X^c\|_F^2 + \gamma \text{tr}(X^c \tilde{L}^c (X^c)^T) \right\} \quad (8)$$

s.t. $\|d_k\|_2 = 1, \forall k$

4 Experiments and Analyze

In this chapter, we will proceed several simulation experiments on two commonly used face databases to compare our proposed algorithm with developed dictionary methods including SRC [13], FDDL [9], KSVD [10], and DLSI [12]. To gain more insights into how the proposed method works, discussion and analysis of the function of the promoted multi-label suppression and Laplacian Eigenmaps will be shown at the end of this section.

4.1 Datasets

The Yale database is a facial set with only 165 grayscale images from 15 individuals, and that means each person owns 11 images. All the images are resized into 576-dimensional vectors with normalization for representation. In the experiment, we randomly select half of the images for training and the other half for testing.

The Extended YaleB contains about 2414 frontal face images of 38 individuals, about 64 for each individual. It is more challenging for a reason that the original signals were cropped to 192×168 pixels, as well as their varying illumination conditions and poses. For the better performance in our experiment, the dataset will be converted as column vectors, normalized into a unit norm, and in the end, projected into a dimension of 300 by the way of PCA. As usual, we randomly split the database into two halves, and that one-half will be used for training while the other is used for testing.

4.2 Accuracy Comparison

Due to the limitation in the size of facial dataset and great changes in facial expression and light, the standard deviation of the experimental results is large and the fluctuation is large. Hence, the experimental results of our algorithm listed in Table 1 are 50 times more that of the independent training splits on an average. The best accuracy result attained by the method we propose is 1.3% and 1.8% higher than that of the linear SVM in two common facial datasets, separately. From the results, we can clearly discover that the global coding classifier outperforms all the other compared algorithms, even the local coding classifier, which proves that our model has strong classification performance.

Table 1. Accuracy on the Yale face and the extended YaleB face dataset separately

Method	Accuracy (%)
<i>Yale face</i>	
DLSI	72.7
D-KSVD	73.2
LC-KSVD	73.6
FDDL	77.2
SVM	94.42
Ours	95.73
<i>Extended YaleB face</i>	
DLSI	71.4
D-KSVD	72.9
LC-KSVD	75.2
FDDL	79.4
SVM	92.8
Ours	94.68

Moreover, for the sake of further estimate, the algorithm is proposed, and we evaluated the above methods under more difficult conditions where only 32, 20, or even 10 training samples of each class is used for learning, while the rest of the data is used as the test set. The result of the direct comparison is listed in Table 2, which indicates that our method still achieves the highest recognition accuracies compared with the competing methods under all situations. As shown in the above experimental results, the global coding classifier still over performs the local coding classifier in classification accuracy. Analyzing the reasons for this situation, the scale of class-specific dictionaries being relatively small, as well as the collaboration of these sub-dictionaries brings up the satisfactory results, which is consistent with the observation.

Table 2. Accuracy on the extended YaleB dataset

Method	Accuracy (%)		
Number of training set	10	20	32
DLSI	83.08	91.36	94.04
D-KSVD	80.53	88.77	94.39
LC-KSVD	81.57	92.36	94.92
FDDL	–	92.04	–
Ours	93.16	96.48	97.93

Another noticeable feature of the results is that even when fewer training samples were adopted, the accuracy we obtained is still much higher than that of the other dictionary methods with more training samples. The trend goes that as the number of the training samples gets smaller, the standard deviation obtained in the experiments increases. This kind of phenomenon reveals such a fact that if the scale of the training set is small, the figures that composed to it play a fatal role in recognition.

4.3 Analysis

In order to show the effect of multi-labeling suppression and Laplacian Eigenmaps more clearly, we adopt the method of keeping the other experimental conditions same, and detecting the effects separately by fixing a scalar at zero and changing the other scalar at the same time on the extended YaleB dataset. The role of the two proposed terms in recognition will be reflected in the following two experiments.

Figure 1 illustrates the performance of the proposed multi-label suppression on the condition of scalar Υ set as 0 in the cases of tNum set as 20 and 32 while tNum means the quantity of training samples each class. We can clearly see that the tendency of the curves is consistent, especially when λ is larger than 50. The accuracy gets very surprised improvement, and then it almost appears a plateau. Furthermore, the classification accuracy of the proposed Laplacian Eigenmaps without the multi-label suppression is shown in Fig. 2. On the condition that scalar λ set as 0 in the cases of tNum set as 20 and 32, the improvements brought by the Laplacian Eigenmaps is obviously stated.

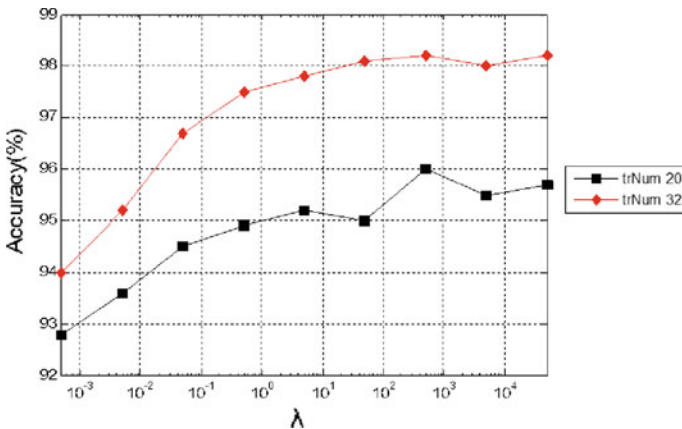


Fig. 1. Accuracy with $\Upsilon = 0$ while tNum = 20 and 32, respectively

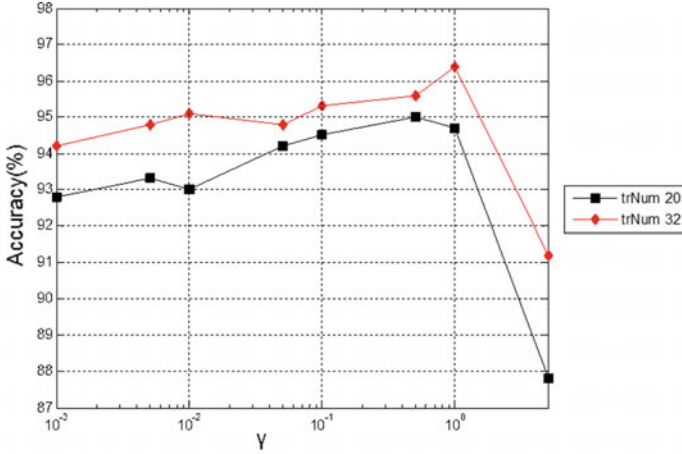


Fig. 2. Accuracy with $\lambda = 0$ while tNum = 20 and 32, respectively

5 Conclusion

In this paper, face recognition is realized by learning a discriminative structured dictionary effectively. In order to make the learnt dictionary contain both reconstruction and discrimination characteristics, the algorithm also learns a linear classifier while learning the dictionary. We propose to add two constraints to the classical dictionary learning algorithm: multi-label suppression used to enlarge the gap between different classes and Laplacian Eigenmaps to make sure that the representations of the same class to be similar to each other. As the supervised learning algorithms goes, most methods adopt l_0 -norm or l_1 -norm as a prior item despite the fact that they are computationally expensive. Thus, l_2 -norm is taken instead in the proposed dictionary learning method to obtain an iteration solution based on closed form. To evaluate our method, we extensively conduct several experiments using two common facial datasets, the results of which demonstrate the advantages of our methods in recognition accuracy and computation reduction. Our work cannot only be applied to face recognition, but also extended to some other classification tasks, such as scene classification and hand-written recognition.

However, although the classification methods based on dictionary learning have achieved promising classification performances, they may still encounter some common problems. Because of the diversity of facial expressions, angles and occlusion degree, new, and significant face data sets will appear constantly after the dictionary is learnt. It would be time consuming if we had to relearn the dictionary. One possible solution is to extend the existing work to online dictionary learning with large-scale dataset, which will be extended as a future work.

References

1. Zhao, W., Chellappa, R., Phillips, P., Rosenfeld, A.: Face recognition: a literature survey. *ACM Comput. Surv. (CSUR)* **35**(4), 399–458 (2003)
2. Turk, M., Pentland, A.: Eigenfaces for recognition. *J. Cogn. Neurosci.* **3**(1), 71–86 (1991)
3. Elad, M., Aharon, M.: Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. Image Process.* **15**(12), 3736–3745 (2006)
4. Hazim Barnouti, N., Sameer Mahmood Al-Dabbagh, S., Esam Matti, W.: Face recognition: a literature review. *Int. J. Appl. Inf. Syst.* **11**(4), 21–31 (2016)
5. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(7), 711–720 (1997)
6. Li, H., Liu, F.: Image denoising via sparse and redundant representations over learned dictionaries in wavelet domain. In: *Proceedings of the 5th International Conference on Image Graphics ICIG 2009*, vol. 15, no. 12, pp. 754–758 (2010)
7. Wang, D., Kong, S.: A classification-oriented dictionary learning model: explicitly learning the particularity and commonality across categories. *Pattern Recognit.* **47**(2), 885–898 (2014)
8. Ramirez, I., Sprechmann, P., Sapiro, G.: Classification and clustering via dictionary learning with structured incoherence and shared features. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 3501–3508 (2010)
9. Mairal, J., Elad, M., Sapiro, G.: Sparse representation for color image restoration. *IEEE Trans. Image Process.* **17**(1), 53–69 (2008)
10. Zhang, Q., Li, B.: Discriminative K-SVD for dictionary learning in face recognition. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2691–2698 (2010)
11. Jiang, Z., Lin, Z., Davis, L.S.: Label consistent K-SVD: learning a discriminative dictionary for recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(11), 2651–2664 (2013)
12. Yang, M., Zhang, L., Feng, X., Zhang, D.: Fisher discrimination dictionary learning for sparse representation. In: *2011 International Conference on Computer Vision*, pp. 543–550 (2011)
13. Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y.: Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(2), 210–227 (2009)



An Improving Data Stream Classification Algorithm Based on BP Neural Network

Baoju Zhang^(✉), Guilin Wang, and Lei Xue

Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
wdxyzbj@163.com

Abstract. With the continuous development of science and technology, many application fields of data belong to the data stream type. Data stream classification is one of the most important analysis methods of data stream processing. The neural network algorithms have no complicated models and reasoning, and have great advantages in data stream classification. In this paper, data stream classification, neural network algorithm, and improved BP neural network algorithm are studied. The neural network toolbox provided by MATLAB is used for data stream classification and simulation.

Keywords: Data stream · Classification · BP neural networks

1 Introduction

In an ever-increasing number of information processing applications, the form of information presentation is no longer a cutoff state, but appears in the form of “streaming”. It is characterized by infinity, unpredictability, non-reproducibility, and changeable. We call them data stream [1].

It requires that processing model of data stream is different from the traditional data processing model. Because the data stream has to be selective [2], it is just a piece of information that stores a portion of the data stream, and then continuously updates the configuration file in the data stream model based on the input data. Finally, the approximate results are provided according to the user requirements [3].

Data stream classification is an important branch of data mining. The traditional classification algorithms mainly include decision tree algorithm, SVM algorithm, Bayesian algorithm, k-nearest neighbor algorithm, and neural network algorithm [4]. This paper mainly chooses neural network algorithm to classify data stream.

Neural network is a mathematical model based on the abstract understanding of biological neurons. BP neural network is a multilayer feed-forward neural network, which is trained according to the error reverse propagation algorithm. It is the most widely used neural network at present. It is also the core part of the forward network and embodies the essence of artificial neural network [5]. It has been experimentally shown that as long as the constructed neural network is large enough, the BP neural network can approximate any continuous function with high accuracy. Data stream classification using BP neural network algorithm, and the neural network algorithm has

strong intelligence, generalization ability is superior to other methods, can use classification software simulation, operable. This dissertation will discuss the basic principle and algorithm process based on the traditional BP neural network algorithm, discover the deficiency, improve, and complete the data stream classification simulation [6]. The specific research process and conclusions will be discussed in the following.

The rest of this paper is organized as follows: In Sect. 2, we introduced the traditional BP neural network algorithm and gave an improved algorithm, Sect. 3 presented the data stream classification experiment based on BP neural network algorithm. Section 4 concluded the paper.

2 BP Neural Network Algorithm

2.1 The Standard BP Neural Network Algorithm Steps

BP neural network consists of three parts: input layer, hidden layer, and output layer [7]. The model of BP neural network is shown in Fig. 1.

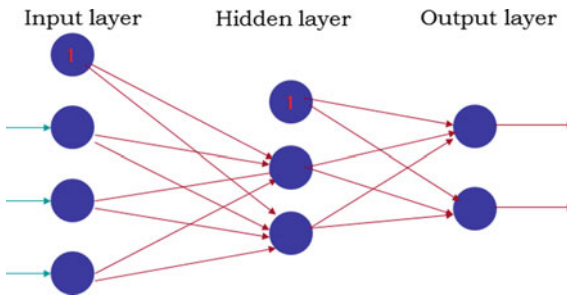


Fig. 1. The model of BP neural network

The calculation process consists of forward propagation and error backpropagation. In the forward propagation process, the input mode is processed layer by layer from the input layer through the hidden unit layer and is turned to the output layer. The state of each layer of the nerve only affects the state of the next layer of neurons. If the output layer does not get the desired output when it outputs, it performs backpropagation and the error signal returns according to the original propagation path by modifying the weight of each neuron, the error signal is minimized. The standard BP neural network learning algorithm [8] is as follows:

Step 1: Network initialization

The initial value of each connection weight is set to the random number within the $(-1, 1)$ interval. Set the error initial value $e = 0$, calculate the accuracy value ε , and the Maximum learning frequency is M .

Step 2: The k th input vector and the corresponding expected output vector are randomly selected.

$$X(k) = (x_1(k), x_2(k), \dots, x_n(k)) \quad (1)$$

$$d_o(k) = (d_1(k), d_2(k), \dots, d_q(k)) \quad (2)$$

Step 3: Calculate the input and output of each neuron in the hidden layer.

Step 4: Using the results of step (3), calculate the partial derivative of the error function for each neuron in the output layer $\delta_o(k)$.

Step 5: Compute the partial derivative of the error function for each neuron in the hidden layer $\delta_h(k)$.

Step 6: Use $\delta_o(k)$ and the output of each neuron in the hidden layer to correct the output layer connection weight.

Step 7: Correct the connection right using $\delta_h(k)$ and the input of each neuron in the input layer

Step 8: Calculate global error E .

$$E = \frac{1}{2m} \sum_{k=1}^m \sum_{o=1}^q ((d_o(k) - y_o(k)))^2 \quad (3)$$

Step 9: Determine whether the model meets the requirements. When the error reaches the preset precision or the learning frequency exceeds the maximum number, the algorithm is terminated. Otherwise, select the next learning sample to return to the third step and enter the next round of learning.

2.2 The Advantage of BP Neural Networks

Nonlinear mapping capability: BP neural network can learn and store a large number of input/output mapping relationship, and there is no need to know in advance mathematical equation describing function relationship, only need to provide the training samples, can complete the multidimensional input nonlinear mapping.

Generalization ability: BP neural network simulation of the human mind has the memory function. If the network has given a sample training, then it sees a strange sample, still can accurately map, and this is the BP neural network generalization ability.

Fault tolerance: When there is a large error in the input sample, BP neural network can be selectively ignored and will not have a great impact on the result. BP neural network has a certain fault tolerance capability.

2.3 The Limitation of BP Neural Networks

The speed of convergence is slow: In order to guarantee the stability of the neural network algorithm, the general fixed learning rate vector should be smaller, which leads to slower speeds of convergence. The more complex problems, the convergence of the longer, the more training time, for the need to deal with the problems of fast, time is too long can lead to the result is not ideal.

It may fall into a local minimum: The traditional BP algorithm uses a gradient descent method, it has a weight adjustment only along the direction of the local improvement, and weights of neural network could be local to the error of plane stop training. So it is obvious that the error is going to be greater.

Network learning is unstable: BP network every learning is, in effect, adjustment of weight, if the learning ability is insufficient, the application will be not ideal, when there is a limit of the trend, technically known as “fitting”, that is to say, because the sample details of neural network learning are too much, it affects generalization.

2.4 The Improvement of BP Neural Networks

For the limitations of the BP neural network algorithm, the researchers explored different improvements. The main purpose of the improvement is usually to speed up the training and avoid falling into a local minimum solution [9]. Common methods are as follows.

Increase momentum item: Each time the weight is adjusted, a small amount of the previous adjustment is added to the adjustment. At present, many BP neural network systems have increased momentum terms to overcome local minimums and improve accuracy. With the introduction of momentum errors, the function of the error surface will be effectively reduced.

Adjust the learning rate adaptively: The learning rate is the step, which is a constant. The best learning speed is difficult to determine. In order to speed up the convergence of the function, the learning rate can be changed in an adaptive manner. The criterion for adjustment is whether the adjustment of the weight really affects the reduction of the error.

Momentum-adaptive learning: Combining these two improved methods together is a momentum-adaptive learning algorithm, which combines the advantages of both and has been widely used in practical applications.

3 Data Stream Classification Experiment

3.1 Data Sources and Normalization

The experimental data comes from the iris data in UCI dataset. There are three categories of iris, namely Iris Setosa, Iris Versicolour, and Iris Virginica, each with 50 sets of data. Each set of data includes four sets of attributes, namely the length of the calyx, the width of the calyx, the length of the petals, and the width of the petals. Because the results of data collection data are often varied, after normalization, the data can be in the same order of magnitude for comparison. Normalization can also accelerate the velocity of gradient descent and improve the accuracy of the results.

3.2 Data Stream Classification

In order to make the data more representative and generic, the data needs to be scrambled and divided into training and testing. The training data are used for training, and the test data are used to test the accuracy of the model. The sample size of the training data should be moderate, too little to be representative of the characteristic

data, too much emphasis on the study burden, and possibly because of too much learning, has produced the fitting.

In this experiment, the training data has 120 groups, while remaining as test data.

3.3 Experimental Parameter Selection and Comparison

In practical application, it is especially important to establish a proper neural network. It mainly includes setting the number of input layer and hidden layer, node number of output layer, transfer function, training function, parameter setting, etc.

3.3.1 Determination of Number of Hidden Layer Nodes in BP Neural Networks

The number of input layer nodes is 4, the number of output layer nodes is 3, and there is no clear formula for the number of neurons in the hidden layer. The heuristic method can be used. (Refer to the formula $m = \sqrt{p+q} + a$, where m is the number of hidden layer nodes, p is the input layer node, q is the output layer node, and a is any integer between 1 and 10.)

The number of hidden layer nodes is set to 10 first. The neural network adopts an improved algorithm and adopts a gradient descent algorithm that incorporates momentum and adaptation. The activation function from the input layer to the hidden layer is “logsig”. The transfer function between the hidden layer and the output layer uses a pure linear function. The training times are set. For 1000 times, the target error is 0.01, and the learning rate is 0.01. Then, change the number of hidden layer nodes to 9 and 11, respectively, run 10 times, record the time of running the results and the number of errors in the operation, take the average, the time is accurate to two decimal places. Classification results of different hidden layer nodes are shown in Table 1.

Table 1. Classification results of different hidden layer nodes

Net number		1	2	3	4	5	6	7	8	9	10	Average
9	Time	3.44	3.53	3.63	3.78	3.54	3.61	3.61	3.54	3.75	3.63	3.6
	Error	0	0	0	0	0	0	0	0	1	0	0.1
10	Time	4.03	3.73	3.68	4.14	3.91	3.64	3.72	3.74	3.64	3.54	3.78
	Error	0	0	0	0	0	0	0	0	0	1	0.1
11	Time	3.7	3.61	3.52	3.73	3.51	3.61	3.53	3.6	3.63	3.68	3.61
	Error	0	0	0	0	1	0	1	0	0	0	0.2

The above experiments show that when the number of nodes is 9, the effect is best, the accuracy is higher, and the running time is shorter. Therefore, we set the number of hidden layer nodes to 9.

3.3.2 The Selection of PB Neural Network Maximum Training Times

The number of nodes in the hidden layer of the fixed neural network is 9, and the error function for opening the neural network operation is shown as Fig. 2.

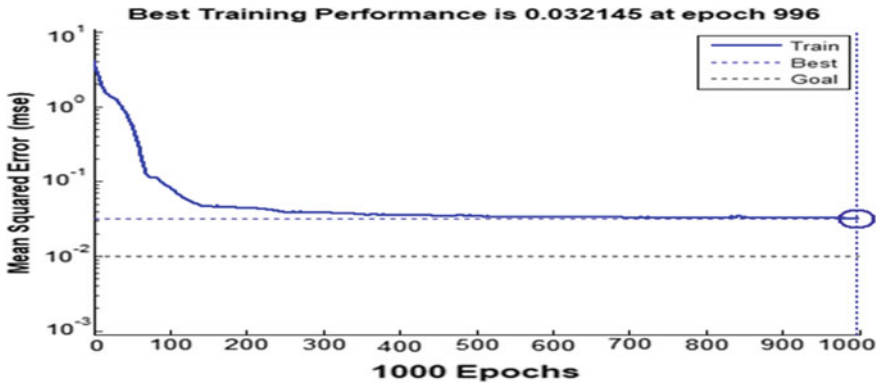


Fig. 2. Neural network error function with node 9

From the error curve, when the maximum number of training times is 996, the best error is 0.032, which does not reach the set target error, and the error between 200 and 1000 changes slowly. Therefore, our maximum number of trainings can be reduced, or the learning rate can be increased so that the function converges quickly. Change the maximum number of trainings to 800, 500, 200, 150, and 100. Each data is run 10 times, and the running time and the average number of errors are recorded. The results are shown in Table 2.

Table 2. Comparison of results of different maximum training times

Maximum number		1	2	3	4	5	6	7	8	9	10	Average
800	Time	3.03	2.91	2.92	3.28	2.85	2.97	3.1	2.73	2.95	3.04	2.978
	Error	0	0	0	0	0	0	0	0	1	0	0.1
500	Time	2.01	1.99	1.94	1.91	1.87	1.86	1.94	1.84	1.85	1.99	1.92
	Error	0	0	0	0	0	1	0	0	0	1	0.2
300	Time	1.26	1.32	1.35	1.29	1.28	1.23	1.24	1.28	1.32	1.35	1.292
	Error	0	0	0	0	1	0	1	0	0	0	0.2
200	Time	1.07	1.01	1	1.02	1.02	1.02	0.96	0.99	0.98	0.93	1
	Error	0	0	0	1	0	0	0	0	0	1	0.2
100	Time	0.7	0.64	0.67	0.67	0.69	0.7	0.69	0.67	0.72	0.7	0.685
	Error	1	1	1	1	1	0	1	0	0	2	0.8

By comparison, the best maximum number of trainings is set to be about 300 times the best, the accuracy is high, and the running time is short. Therefore, this experiment sets the maximum number of trainings to 300.

3.3.3 Comparison Between Standard BP Neural Network and Improved Neural Network Method

Keep the existing parameters unchanged, change the training function, use the standard gradient descent training function “traind”, add the momentum training function “traindm”, adaptive learning rate training function “traindx”. Record experimental results and images. The results are shown in Table 3.

Table 3. Attribute parameter values

Maximum number		1	2	3	4	5	6	7	8	9	10	Average
traind	Time	1.27	1.31	1.26	1.33	1.29	1.44	1.34	1.25	1.24	1.19	1.292
	Error	0	0	1	0	0	1	1	0	1	1	0.5
trainm	Time	1.22	1.31	1.23	1.21	1.23	1.23	1.24	1.21	1.26	1.22	1.236
	Error	0	0	0	0	0	0	1	1	0	1	0.3
traindx	Time	1.28	1.26	1.26	1.28	1.3	1.25	1.27	1.27	1.26	1.28	1.271
	Error	0	0	0	0	0	0	0	1	0	0	0.1

From Table 3, it can be known that the traditional BP algorithm error only converges along the gradient, and sometimes, it easily falls into the local minimum, leading to inaccurate results. By adding the momentum factor, this situation can be avoided, but the running time will increase. To increase the learning step, we use a variable learning rate to make BP neural network training converge faster. Therefore, we use the improved BP neural network algorithm to classify the data stream.

3.3.4 Improved BP Neural Network Classification Result Verification

The program was run using an improved BP neural network algorithm and the drawing function in MATLAB [10] was used to show the error of the iris flower classification result and to show the result of the classification. The final result image of the experiment is shown in Figs. 3 and 4.

It can be seen from the figure that red represents the first kind of flower, blue represents the second kind of flower, green represents the third kind of flower, and the classification accuracy rate reaches 100%.

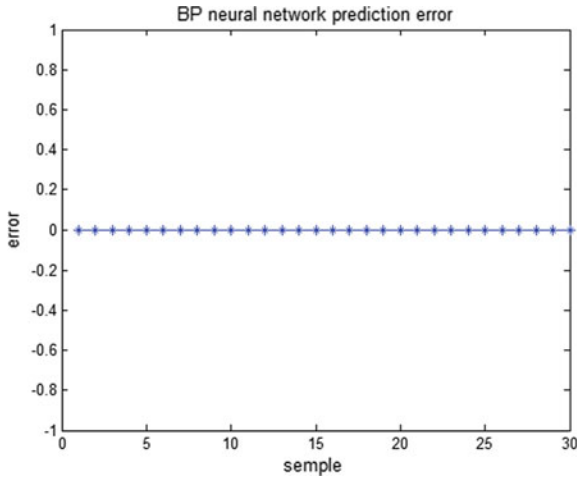


Fig. 3. BP neural network prediction error

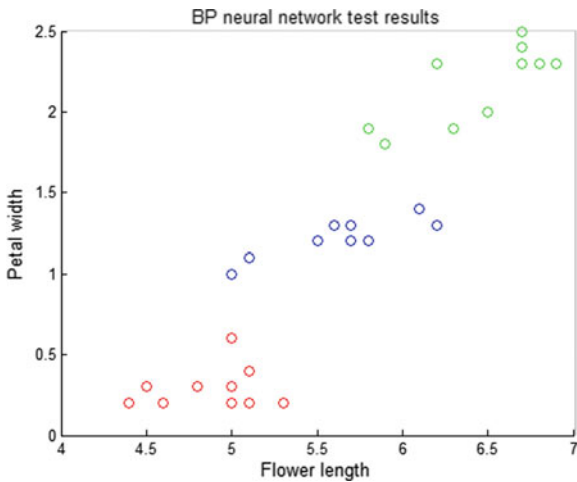


Fig. 4. Schematic diagram of BP neural network classification

4 Conclusions

This paper discusses the basic knowledge of data stream classification and neural network, chooses BP neural network as the main model of classification, and uses MATLABs own neural network toolbox to simulate and realize. Through experimental comparison, the neural network can achieve higher accuracy. The deficiency of this experiment is that it only imitates the classification status of the data stream after processing. When the data stream is dynamic, it needs to undergo data preprocessing to extract dynamic data features. This process is not expressed in this experiment,

but it does not affect the final conclusion: The data stream classification algorithm based on neural network has high accuracy, has strong generalization ability, and can handle nonlinear problems well.

Acknowledgements. This work is supported by Natural Youth Science Foundation of China (61501326, 61401310), the National Natural Science Foundation of China (61731006), and Natural Science Foundation of China (61271411). It is also supported by Tianjin Research Program of Application Foundation and Advanced Technology (15JCZDJC31500), and Tianjin Science Foundation (16JCYBJC16500). This work was also supported by the Tianjin Higher Education Creative Team Funds Program.

References

1. Xie, Y., Xiaolian, Z.: Analysis of data stream classification technology. *Comput. Modern.* **01**, 32–36 (2010)
2. Yanping, Z., Ling, Z.: *Machine Learning Theory and Algorithm*, p. 165. Science Press, Beijing (2012)
3. Nalavade, J.E., Senthil Murugan, T.: THRFuzzy: tangential holoentropy-enabled rough fuzzy classifier to classification of evolving data streams. *J. Central South Univ.* **24**(08), 1789–1800 (2017)
4. Shucheng, H., Yahui, Q.: Review of data stream classification technology. *Appl. Res. Comput.* **26**(10), 3604–3609 (2009)
5. Khaw, J.F.C., Lim, B.S., Lim, L.E.N.: Optimal design of neural networks using the Taguchi method. *Neurocomputing* **7**(3) (1995)
6. Liqun, H.: *Artificial Neural Network Theory, Design and Application*, p. 5. Chemical Industry Press, Beijing (2007)
7. Yudong, H.: *Application of BP neural network in data classification of information system*. China University of Geosciences (Beijing) (2013)
8. Dehu, Q., Jichang, K.: Design of BP neural network. *Comput. Eng. Des.* **02**, 47–49 (1998)
9. Wencheng, W.: Research on adaptive learning rate in BP neural network. *Comput. Sci.* **04**, 48–50 (1995)
10. Bing, L., Haixia, G.: *MATLAB Neural Network Super-Learning Manual*. People's Posts and Telecommunications Press, Beijing (2013)



Facial Fatigue Detection Based on Machine Learning

Dewei Zheng¹(✉), Shaohua Cui², and Chenglin Zhao¹

¹ Key Laboratory of Universal Wireless Communications,
MOE Beijing University of Posts and Telecommunications,
No. 10 Xitucheng Road, Beijing, China
dw.zh@foxmail.com

² China Petroleum Technology & Development Corporation, Beijing, China

Abstract. One of the most important reasons for productivity decline and accidents is work fatigue. Work fatigue research has become more and more important in modern society. This paper proposes a method to detect fatigue, Build new features, propose new compensation methods, and combine the existing models to make the method adapt to the complex environment. As a result, it effectively improves the work fatigue detection efficiency and accuracy under the production environment.

Keywords: Work fatigue detection · Features · Compensation method

1 Introduction

The “overwork” problem has aroused widespread public concern due to its serious hazards. Work fatigue refers to the discomfort that people gradually appear in the work. It is a recognized occupational harmful factor that can endanger the workers’ health and ability to work [1].

At present, there are many methods for work fatigue detection including self-rating method, physiological testing method, and biochemical testing method [2]. It is difficult to choose which method because it is difficult to balance between objectivity and price.

In this study, we refer to the method of driving fatigue detection [3–6] and found some shortcomings of it: (1) The dimension of the feature vector is low. (2) There is no processing for data changes caused by dynamic angles and dynamic distance. (3) The existing method cannot adapt to the discrete short-time windows in production. In view of the above drawbacks, this paper builds new features which are suitable for the discrete short-time window, propose a compensation method to preprocess the data changes. Combine decision tree with SVM for model training. The experiments show that this method can be used to perform work fatigue objectively, efficiently, and accurately.

2 Data Processing

2.1 Feature Extraction

In this paper, AAM algorithm is used to collect the eye and mouth feature points [7], and three new features are proposed on the basis of the five features (as shown in the first five items in Table 1) commonly used in the past.

Table 1. Feature vector parameters

Type	Name	Description
EYE	h_{eye}	Eye height
EYE	$perclos$	Perclos (P80)
EYE	$MECD$	Maximum eye close duration
EYE	BT	Blink times
EYE	$AECS$	Average eye closure speed
EYE	$AREA$	Average area
EYE	$ratio_e$	Eye aspect ratio
MOUTH	$ratio_m$	Mouth aspect ratio

Average Area and Average Eye Aspect Ratio The two new feature $AREA$ and $ratio_e$ are proposed in order to weave the changes on the length of the eye into the eye height.

$$AREA = \frac{\sum h_{eye} \times l_{eye}}{frames} \quad (1)$$

$$ratio_e = \frac{\sum \frac{h_{eye}}{l_{eye}}}{frames} \quad (2)$$

where l_{eye} represents the eye's max length and $frames$ represents the number of frames extracted in the time window.

Mouth Aspect Ratio Yawn is a conditioned reflex deep breathing activity. We defined the $ratio_m$ to describe and determine it as

$$ratio_m = \frac{h_m}{l_m} \quad (3)$$

where h_m and l_m represent the max height and the max width of the mouth.

2.2 Feature Compensation

This paper proposes a compensation method based on distance mapping and Euler angles.

Research shows that the PD of adults is generally constant and between 58 and 64 mm [8]. So, we can quantize the single pixel distance as

$$l_{pixel} = \frac{PD}{l_{pd}} = \frac{PD}{\|l_{pixel} - r_{pixel}\|} \tag{4}$$

where the l_{pixel} and r_{pixel} represent coordinate of the left and right eye.

The Euler angles can be used to describe the head posture, as shown in Fig. 1a., we can infer that h_{eye} and $AREA$ is approximately:

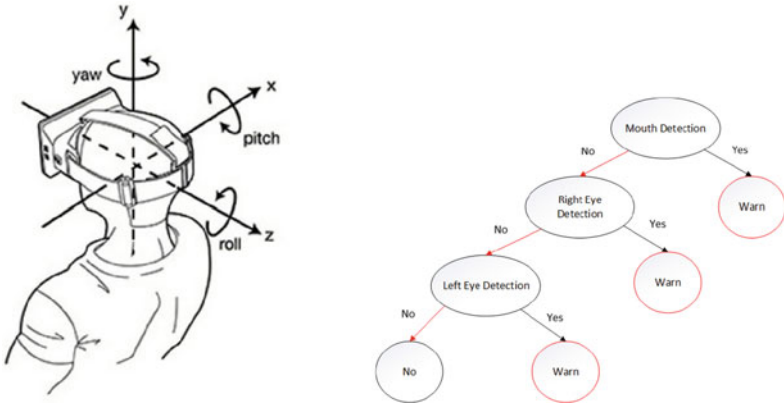
$$h_{eye} = \frac{m \times l_{pixel}}{\cos(\alpha_p)} = \frac{m \times PD}{\cos(\alpha_p) \times l_{pd}} \tag{5}$$

$$area = h_{eye} \times \frac{n \times l_{pixel}}{\cos(\gamma_y)} = \frac{m \times n \times PD^2}{\cos(\alpha_p) \cos(\gamma_y) \times l_{pd}^2} \tag{6}$$

where m, n represents the eye height and width in pixels.

2.3 Decision Tree and SVM

We combined decision tree with SVM in this paper as shown in Fig. 1b in order to improve the accuracy and efficiency when detecting.



(a) α_p for pitch angle, β_r for roll angle, and γ_y for yaw angle

(b) Model

Fig. 1. Euler angles and model

3 Experiments

3.1 Feature Verification

Average Area and Average Eye Aspect Ratio P-tests were performed on 100 groups of the $AREA$ and $ratio_e$. The result is $P_{area} = 6.684 \times 10^{-41}$, $P_{ratio} = 1.185 \times 10^{-47}$, far below 0.001. It shows that there is a significant difference between fatigue and non-fatigue data.

Mouth Aspect Ratio This paper processes the yawning data with single-rule-based classification. If the worker yawns, we can infer that he is in a fatigue state. Fig. 2a shows the changes on $ratio_m$ when yawning and speaking.

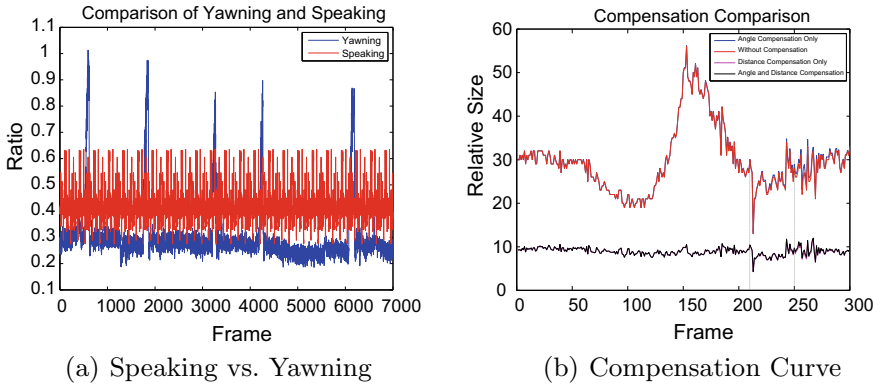


Fig. 2. Mouth and compensation research

From the comparison in Fig. 2a, we can infer that the worker is yawning when $ratio_m$ is greater than 0.7 or more.

3.2 Compensation Function Verification

A video of dynamic distance and dynamic head posture was collected for verification as shown in Fig. 2b. The curve tends to be stable through the compensation.

Do experiments on compensated and uncompensated eye data with fivefold cross-validation method, the results are shown in Table 2.

As we can see from Table 2, the accuracy rate is generally improved through compensation when using linear and polynomial kernel and it is much larger than that of RBF and sigmoid kernel. The compensation does not make any sense when using RBF or sigmoid kernels. It is fully demonstrated that polynomial kernel and linear kernel are more suitable for work fatigue detection in workshop and the compensation method can significantly improve the accuracy rate.

Table 2. Effect of compensation on accuracy (OD: right eye, OL: left eye, OU: both eye)

Kernal	OD	OD compensated	OL	OL compensated	OU	OU compensated
Linear	0.836	0.985	0.857	0.937	0.901	0.971
Polynomial	0.835	0.985	0.857	0.951	0.904	0.975
Sigmoid	0.615	0.615	0.615	0.615	0.615	0.615
RBF	0.615	0.615	0.620	0.620	0.615	0.615

3.3 Model Training

The model is trained by 3000 data in which there are 1170 fatigue samples and 1830 non-fatigue samples. The experiment shows that the precision rate is 99.83% and it increased 3% over traditional SVM method.

4 Conclusion

This paper builds some new features, which are suitable for work fatigue detection in workshop and raises the dimension of the feature vector. The compensation method proposed can effectively deal with the data changes caused by dynamic angle and dynamic distance. It allows the fatigue detection to adapt to the complex environment. Combine decision tree with SVM for model training, and the experiments show that it can improve the accuracy rate significantly.

Acknowledgement. This project is supported by the Project of Intelligent Manufacturing Integrated Standardization and New Model Application.

References

1. Tao, Q., Ji, Y.: Review and comment on overwork in China. *Human Res. Develop. China* (2015)
2. Chen, J.W., Chun-Bo, B.I., Liao, H.J., et al.: Comparative research on measurement methods of work fatigue. *J. Saf. Sci. Technol.* (2011)
3. Azim, T., Jaffar, M.A., Mirza, A.M.: Fully automated real time fatigue detection of drivers through fuzzy expert systems. Elsevier Science Publishers B. V. (2014)
4. Yan, J.J., Kuo, H.H., Lin, Y.F., et al.: Real-time driver drowsiness detection system based on PERCLOS and grayscale image processing. In: *International Symposium on Computer Consumer and Control*, pp. 243–246 (2016)
5. Lee, B.G., Chung, W.Y.: Driver alertness monitoring using fusion of facial features and bio-signals. *IEEE Sens. J.* **12**(7), 2416–2422 (2012)
6. Niu, Q.N.: Research on fatigue driving detection method based on information fusion. Jilin University (2014)
7. Cootes, T.F., Taylor, C.J.: Constrained active appearance models. In: *International Conference on Computer Vision*, pp. 748–754 (2001)
8. Guo, Y.L.: Talking about the interpupillary distance. *Metrol. Meas. Tech.* (2014)



Genetic Algorithm-Based Beamforming Using Power Pattern Function

Shuoguang Wang^(✉), Shiyong Li, and Houjun Sun

Beijing Key Laboratory of Millimeter Wave and Terahertz Technology,
Beijing Institute of Technology, Beijing, China
1120131426@bit.edu.cn

Abstract. To solve the beamforming problems of small-scale array, this paper describes a new method using power pattern function and genetic algorithm. Compared with the traditional antenna beamforming methods which utilize the envelope of the radiation pattern, this method can control all the sidelobes quantitatively and form the desired array pattern. For a small-scale array, its envelope is difficult to be quantified and obtained. The proposed method in this paper is more intuitive and reliable to control the sidelobe levels. Using the power pattern optimization method, it can obtain an improved array pattern and fast convergence in small-scale arrays design. Numerical results are presented to verify the convergence and computational efficiency of the proposed method.

Keywords: Beamforming · Power pattern function · Genetic algorithm · Small-scale arrays design

1 Introduction

The goal of antenna array synthesis is to determine the appropriate physical layout of the array, phases, and amplitudes to produce the radiation pattern close to the desired one. For many applications, there are requirements to suppress the sidelobe level (SLL) of the beam pattern, which can be achieved by adjusting the phases and amplitudes of the elements, just as the Dolph–Chebyshev and Taylor methods do. However, they cannot obtain arbitrary radiation pattern.

Other methods of controlling the array pattern employ evolutionary algorithms, for instance, Simulated Annealing (SA) [1], Genetic Algorithm (GA) [2–8], and Particle Swarm Optimizations (PSO) [1, 9] to optimize the cost function depending on the radiation pattern. Inspired by Darwin’s theory of evolution and the concept of the survival of the fittest ones, GA achieves its goals by simulating the processes of reorganization and mutation of genes in the human evolution. During multiplication and crossbreeding, the fittest is chosen to produce more offspring so that they can homogenize the population.

Genetic algorithm is a global random search optimization algorithm [5], which has been widely used in many domains. In [10], GA is applied to the beamforming of a linear sensor array and its cost function is related to envelope. This method uses Dolph–Chebyshev distribution as a reference, and searches the best pattern in the

global range. However, the GA in [10] adopts the traditional Simple Genetic Algorithm (SGA), which has the problems of low convergence rate, low computational efficiency [11, 12] and poor stability of the calculation results, which is not conducive to improving the performance of beamforming. Moreover, as for small-scale array, it is hard to determine the positions of null, sidelobe, and the ideal beam pattern, which may result in non-convergence.

As for the optimization of the beam pattern, the traditional genetic algorithm uses the envelope as the cost function [12], which is defined by the sum of the part of the pattern outside the limit of the given envelope, the same as [10]. It is very easy to operate for large-scale arrays, but it is not the same case for small-scale arrays. Therefore, this paper presents an improved power pattern calculation model, which can improve the convergence speed and the stability of the optimization results. We use this model to build the cost function which gives the relation between the SLL and the amplitudes. The GA algorithm is applied to minimize the cost function due to the simplicity. As a result, the drawback of the envelop methods is avoided and faster convergence is gained.

The organization of the paper is as follows: Sect. 2 presents how to use the power pattern method to obtain the SLL expression and the optimization model. In Sect. 3, the optimization procedure of the GA is discussed. The simulated results of traditional envelop strategies and our modeling strategies are presented in Sect. 4. In Sect. 5, summaries and conclusions on the performance of the strategies are mentioned.

2 Power Pattern Based Method

When describing the characteristics of the radiation pattern of antennas, engineers use the radiation pattern in many situations. However, in further study, the power pattern method is adopted to analyze the problem. As is known, although more complex, the power pattern shows effectiveness and efficiency when describing several characters, such as the position of SLL, the magnitude of SLL, etc. To simplify the analysis, isotropic elements are adopted in the arrays in the following.

For a linear array (LA), its beam pattern can be written as

$$S(\theta, \varphi) = \sum_{n=0}^{N-1} \dot{I}_n e^{jkz_n \cos \theta} = \sum_{n=0}^{N-1} I_n e^{jn(kd \cos \theta + \alpha)} \quad (1)$$

where \dot{I}_n describes the complex excitation coefficient of n th element, and $\dot{I}_n = I_n e^{jn\alpha}$, α is the phase of the n th element. z_n is the coordinate of the n th element, d is the distance between the antenna elements, and k is the propagation constant. To simplify the expression, we define

$$u = kd \cos \theta + \alpha \quad z = e^{-ju} \quad (2)$$

Using (2) into (1), we obtain

$$S(z) = \sum_{n=0}^{N-1} I_n z^{-n} \tag{3}$$

Take the N elements linear array as an example, the power pattern can be written as

$$|S(z)|^2 = S(z)S^*(z) = \left(\sum_{n=0}^{N-1} I_n z^{-n} \right) \left(\sum_{n=0}^{N-1} I_n z^n \right) = \sum_{m=0}^{N-1} I'_m (z^m + z^{-m}) \tag{4}$$

where

$$I'_m = \sum_{n=0}^{N-1-m} I_n I_{n+m} \quad 1 \leq m \leq N-1$$

$$I'_0 = \frac{1}{2} \sum_{n=0}^{N-1} I_n^2 \tag{5}$$

Make the following substitutions:

$$z + z^{-1} = y, \quad z^2 + z^{-2} = (z + z^{-1})^2 - 2 = y^2 - 2,$$

$$z^3 + z^{-3} = (z + z^{-1})^3 - 3(z + z^{-1}) = y^3 - 3y, \quad z^4 + z^{-4} = \dots \tag{6}$$

Then, (4) can be reformulated as

$$P(y) = |S(z)|^2 = \sum_{m=0}^{N-1} I'_m y^m \tag{7}$$

Here, I'_m is determined by (5). It is widely known that

$$|S(z)|^2 = |S(u)|^2 = 2 \sum_{m=0}^{N-1} I'_m \cos mu \tag{8}$$

Equation (8) implies the extreme points of the radiation pattern. And, rewrite them as follows:

$$\frac{dP(y)}{dy} \frac{dy}{du} = -2 \sin u \frac{dP(y)}{dy} = 0 \tag{9}$$

For a physically achievable array, its power pattern function $P(y)$ has to satisfy:

$$P(y) = 0 \quad -2 \leq y \leq 2 \tag{10}$$

According to the power pattern theory, an array has a power pattern with one of the two following forms, which has excitation of nonnegative symmetric and the maximum number of zero points [13]:

$$\begin{aligned}
 P_o(y) &= \prod_{i=1}^{(N-1)/2} (y + c_i)^2 && (N \text{ is odd}) \\
 P_e(y) &= (y + 2) \prod_{i=1}^{(N-2)/2} (y + c_i)^2 && (N \text{ is even})
 \end{aligned}
 \tag{11}$$

The corresponding beam pattern is given by

$$\begin{aligned}
 S_o(z) &= \prod_{i=1}^{(N-1)/2} (1 + c_i z^{-1} + z^{-2})^2 && (N \text{ is odd}) \\
 S_e(z) &= (1 + z^{-1}) \prod_{i=1}^{(N-2)/2} (1 + c_i z^{-1} + z^{-2})^2 && (N \text{ is even})
 \end{aligned}
 \tag{12}$$

where the coefficient c_i can be obtained directly from I_1 .

Take the nonnegative symmetric array of five elements as an example, if the normalized excitation coefficients are $[1 I_1 I_2 I_1 1]$, then from (3), we can gain:

$$S_o(z) = 1 + I_1 z^{-1} + I_2 z^{-2} + I_1 z^{-3} + z^{-4}
 \tag{13}$$

Compared (12) with (13), c_1, c_2 can be easily obtained and expressed by the amplitudes of elements.

From (9), we can figure out the extreme points, but the null points of $y = -c_1, y = -c_2$ should be cleared out to obtain the position of sidelobe. Though $y = -2$ does not satisfy (9), it is the position of the second sidelobe. They are presented as follows:

$$y_{SL1} = -\frac{1}{2}(c_1 + c_2) \quad y_{SL2} = -2
 \tag{14}$$

Take them into power pattern equation, SLL can be obtained as

$$SSL_i = P(y_{SLi})/P(2) \quad i = 1, 2
 \tag{15}$$

From the above analysis, the power pattern $P(y)$ can be obtained. The position of the sidelobes can be calculated so as to the SLL. When an array has excitation of nonnegative symmetric and the maximum number of zero points, the direct relationship between the amplitudes and the SLL can be obtained accurately.

So, the SLL can be presented as a column vector with amplitudes as variables:

$$sidelobe_{real} = f(In) \quad (16)$$

where In means the amplitudes of elements.

3 Steps of Optimization by GA

In the optimization of small-scale array, the SLL is the only parameters to be known rather than the position of them.

To build a tractable optimization problem, we compute the difference between the SLL gained from amplitudes and the expected SLL, then take the 1 – 2 norm as the cost function. This is expressed as

$$\underset{In}{arg \min} \|sidelobe_{count} - sidelobe_{ideal}\|_2 \quad \text{s.t.} \quad 0 < In < 1 \quad (17)$$

where $sidelobe_{count}$ is the column vector of the computed SLL and $sidelobe_{ideal}$ is the column vector of the expected SLL. GA can be used to solve this problem and find a set of amplitudes which fit the requirement.

The procedure of GA is as follows [14]:

- (1) Generate initial population: The initial population is generated using random number generation, the initial weight vector $W = [\omega_1 \omega_2 \dots \omega_{N-1} \omega_N]$, where $\omega_i = [I_1 I_2 \dots I_{n-1} I_n]^T$ and satisfies the symmetry. N is the population size, and n is the number of array components.
- (2) Calculate and sort individual fitness values: Use the cost function:

$$func_{cost} = \sum \|sidelobe_{count} - sidelobe_{ideal}\|_2 \quad (18)$$

Calculate the fitness value of each individual weight vector and sort them, where $\|*\|_2$ represents 1 – 2 norm of *.

- (3) Choose individuals: Choose a better weight vector $\omega_i = [I_1 I_2 \dots I_{n-1} I_n]^T$, and eliminate relatively backward individuals.
- (4) Cross and alienation: In the new generation of populations, to generate new individuals, in other words, a new weight vector, crossover, and alienation operations are implemented according to the set crossover rate and alienation rate.
- (5) Judge, circle, or end: Put new individuals from the step into Eq. (18), and determine whether the cost satisfies the criterions or whether the maximum number of iterations are reached. If so, comes to the end. If not, turn to step 2 for further optimization. After those criterions are satisfied, the optimized vector $\omega_{opt} = [I_1 I_2 \dots I_{n-1} I_n]^T$ would be obtained.

4 Simulation Results

In order to illustrate the effectiveness of the stated approach, two examples for synthesizing linear array are presented in this section.

4.1 Arbitrary SLL

First, we consider a small-scale linear array composed of $\lambda/2$ -spaced with eight isotropic elements in the array, and use the cost function defined as the sum of parts beyond the upper limits of restricts. It is expressed in (18). The second cost function is defined as what we mentioned in the previous section, and its expression is so as (18).

It is hard to quantify envelope for small-scale arrays, so we can only roughly determine the relative positions of different sidelobes. From $u = -1$ to $u = 0$, the SLL is set to 12, 16, and 20 dB. The GA uses discrete values for amplitude. And, its population size is equal to 500 and there are 500 generations. The stopping criterion is the average relative change in the best cost function is equal to or less than 10^{-5} .

The convergence of the normalized cost function value versus the number of iterations is shown in Fig. 1. The figure shows that the fitness value of cost function 2 starts at a normalized value of about 1 and quickly drops to the value of about 0.02 in about 40 iterations. Comparatively, after about 90 iterations, the mean values of cost function 1 are settled to nearly 0.3. The optimization of the cost function 2 converges faster and better.

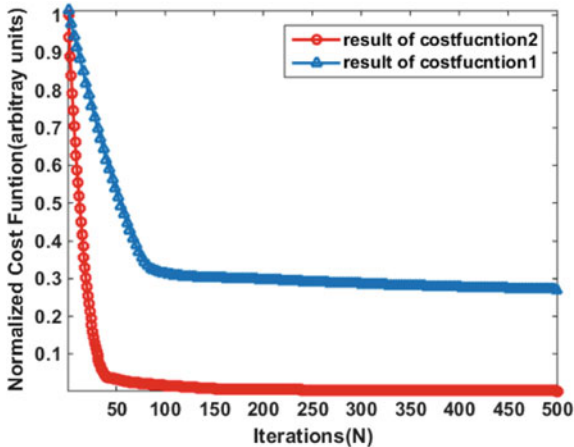


Fig. 1. Convergence of the average normalized cost function versus the number of iterations

Figure 2 shows the comparison of far-field radiation pattern of the best results using different cost function by GA. The results of the cost function 2 fit our requirement better. For small-scale arrays, it is difficult to determine the envelope of the radiation pattern. If the envelope is not set well, the optimization result may not converge. But, when the cost function computed by power pattern is used, a better approximation of the desired pattern will be obtained.

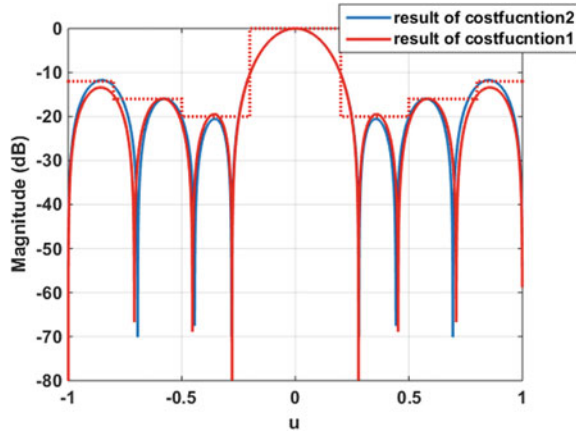


Fig. 2. Comparison of far-field radiation pattern using different cost functions by genetic algorithm

4.2 Chebyshev SLL

In the second example, a small-scale linear array is synthesized. The array is arranged with uniform spacing $d = \lambda/2$. Due to the symmetric structure of this array, only half of the elements in the array are optimized. The ideal amplitude distribution is obtained through the Dolph–Chebyshev method, which is $[0.5799, 0.6603, 0.8751, 1]$. The two cost functions, which are the same as example 1, are determined for GA optimizer.

Figure 3 shows the comparison of the best results of far-field beam pattern using different cost functions by GA for Chebyshev weight. Note that, the SLLs of the synthesized pattern of cost function 1 (after 100 iterations) achieve 16.7, 20, and

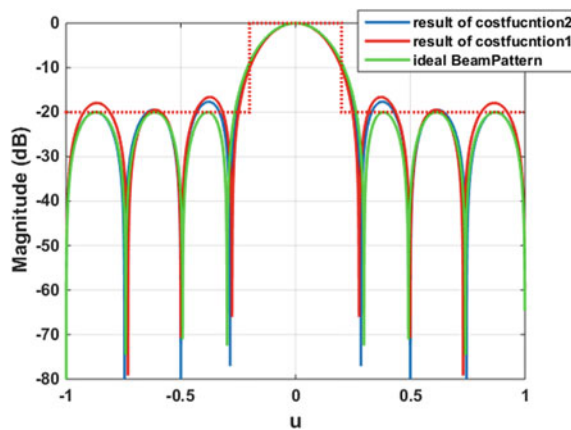


Fig. 3. Comparison of far-field beam pattern using different cost functions by genetic algorithm for Chebyshev weights

18 dB. While the results of cost function 2 are nearly 17.6, 20, and 20 dB (after 50 iterations), which are very close to the result of triple 20 dB optimized by the Chebyshev weight.

5 Conclusion

This paper presents a novel GA-based beamforming method, which applies the power pattern to construct the cost function, in which the SLL can be efficiently determined. The proposed method is much faster than the method of applying envelopes as the cost function to find the amplitudes of elements of small-scale arrays, while maintaining a similar or even better pattern performance. Also, the cost function is easier to realize in small-scale arrays. This advantage has been validated in the examples of synthesizing the small-scale arrays. Applying this model, the algorithm converges faster because we optimize the exact SLL rather than the envelope. However, this method seems difficult when applied to large-scale arrays, because finding the expression of SLL would be intractable.

References

1. Murino, V., Trucco, A., Regazzoni, C.S.: Synthesis of equally spaced arrays by simulated annealing. *IEEE Trans. Signal Process* **44**(1), 119–122 (1996)
2. Rahmat-Samii, Y., Michielssen, E. (eds.): *Electromagnetic Optimization by Genetic Algorithm*. Wiley, New York (1999)
3. Johnson, J., Rahmat-Samii, Y.: Genetic algorithms and method of moments (GA/MOM) for the design of integrated antennas. *IEEE Trans. Antennas Propag.* **47**(10), 1606–1614 (2001)
4. Himdi, M., Daniel, J.: Synthesis of slot coupled loaded patch antennas using a genetic algorithm through various examples. *IEEE Trans. Antennas Propag.* **3**, 1700–1703 (1997)
5. Ares-Pena, F.J., Gonzalez, J.A., Lopez, E., Rengarajan, S.R.: Genetic algorithms in the design and optimization of antenna array patterns. *IEEE Trans. Antennas Propag.* **47**(3), 506–510 (1999)
6. Yan, K.K., Lu, Y.: Sidelobe reduction in array-pattern synthesis using genetic algorithm. *IEEE Trans. Antennas Propag.* **45**(7), 1117–1122 (1997)
7. Tennant, A., Dawoud, M.M., Anderson, A.P.: Array pattern nulling by element position perturbations using a genetic algorithm. *Electron. Lett.* **30**(3), 174–176 (1994)
8. Yeo, B.K., Lu, Y.: Array failure correction with a genetic algorithm. *IEEE Trans. Antennas Propag.* **47**(5), 823–828 (1999)
9. Gies, D., Rahmat-Samii, Y.: Particle swarm optimization for reconfigurable phased-differentiated array design. *Opt. Tech. Lett* **38**(3), 168–175 (2003)
10. Yang, D., Wang, Y.-M., Gou, Y.: Comparative analysis of SA and GA algorithms in beam pattern design. *Comput. Simul.* **25**(8), 323–327 (2008)
11. Keen-Keong, Y., Yilong, L.: Side lobe reduction in array-pattern synthesis using genetic algorithm. *IEEE Trans. Antenna Propag.* **45**(7), 1117–1122 (1997)
12. Brown, A.D.: *Electronically Scanned Arrays MATLAB Modeling and Simulation* (2012)

13. Zhenghui, X., Weiming, L., Wu, R.: Array Antenna Analysis and Synthesis. Beihang University, pp. 102–178 (2011)
14. Man, K.F., Tang, K.S., Kwong, S.: Genetic algorithms: concepts and applications. IEEE Trans. Industr. Electron. **43**(5), 519–534 (1996)
15. Ming, Z., Shudong, S.: Principles and applications of genetic algorithms. National Defense Industry Press, Beijing: Academic, pp. 18–64 (1999)



Short-Text Sentiment Analysis Based on Windowed Word Vector

Dongmei Zhao¹, Yingli Shen¹, Yabo Shen¹, Yong Ma¹, Yun Jin^{1,2},
Shidang Li¹, and Mingliang Gu¹(✉)

¹ School of Physics and Electronic Engineering, Jiangsu Normal University,
No. 101, Shanghai Road, Xuzhou 221116, Tongshan District, Jiangsu Province,
People's Republic of China
mlgu@jsnu.edu.cn

² The key laboratory of children development and learning science,
ministry of education, Southeast University, Nanjing, China

Abstract. The traditional text sentiment analysis directly inputs syntactic feature or word vector to model. It fails to consider the characteristics of time series in sentiment. Considering that product reviews are short text, this paper comes up with the method of windowed word vector and classifier fusion in the decision layer. The results indicate that the proposed method can achieve better performance than several existing methods.

Keywords: Sentiment analysis · Windowed word vector · Classifier fusion · Short text

1 Introduction

The text sentiment analysis aims to explore the emotional tendency of texts. The traditional text sentiment analysis methods are mainly divided into two types: sentiment lexicons method and machine learning method [1]. The early work is mainly based on the lexicon, by constructing sentiment lexicon such as HowNet or WordNet [2]. Keshavarz et al. combined corpora-based and lexicon-based to build adaptive sentiment lexicons in order to improve polarity classification of sentiments [3]. The words in emotional lexicons can clearly reflect a person's emotional tendencies, while emotional lexicons cannot include all emotional words. The method of sentiment analysis based on machine learning is usually through the training of the labeled corpus to generate a tendency classifier and to classify the test text [4]. Currently, more mature classification methods have Support Vector Machine (SVM) [5], Naive Bayes (NB) [6], and Maximum Entropy (ME) [7]. It is necessary for machine learning method to convert the text into a vector form. The dimensions of the traditional Vector Space Model (VSM) [8] are usually thousands or even ten thousand. For short texts such as product reviews, this method can lead to the sparsity of VSM. Mikolov et al. [9] proposed the Word2Vec framework, which can generate word embedding, which represents a text word. It can solve dimensional problems, and discovery the similarity

between words. While the Word2Vec also fails to consider emotional time series information changes. There are also some researches on the sentiment analysis based on the depth study. Certainly, some scholars used the knowledge of the neural network to conduct sentiment analysis [10].

Inspired by the recent success of word vector, this paper proposes a method of windowed word vector. First, adding different size of window on the word vector, in order to consider the characteristics of time series in sentiment. Second, through window operation, the model can obtain different combined features. Then, put these combined features to classifiers fused in the decision layer, and finally uses a grid search algorithm to find the best weight belongs to each classifier.

This paper is organized as follows. The experiment methodology is proposed in Sect. 2. In Sect. 3, we provide the simulation results and the detailed analysis. Finally, conclusions and the future works are drawn in Sect. 4.

2 Method Construction

This section will explore the method construction of short-text sentiment analysis. Our model has two distinctive characteristics: (i) it carries out window operation on the word vector, enabling it to take full account of changes in emotional time series information in short text. (ii) It has classifier fusion at the decision-making level, using the grid search method to find the optimal weight. Furthermore, the sentiment classification results are obtained.

The overall architecture of the experimental methodology is shown in Fig. 1. It consists of several parts: feature selection, word vector training, word vector conversion, window operation, independent classifier model, and classifier fusion at the decision-making level. We describe the details of different components in the following sections.

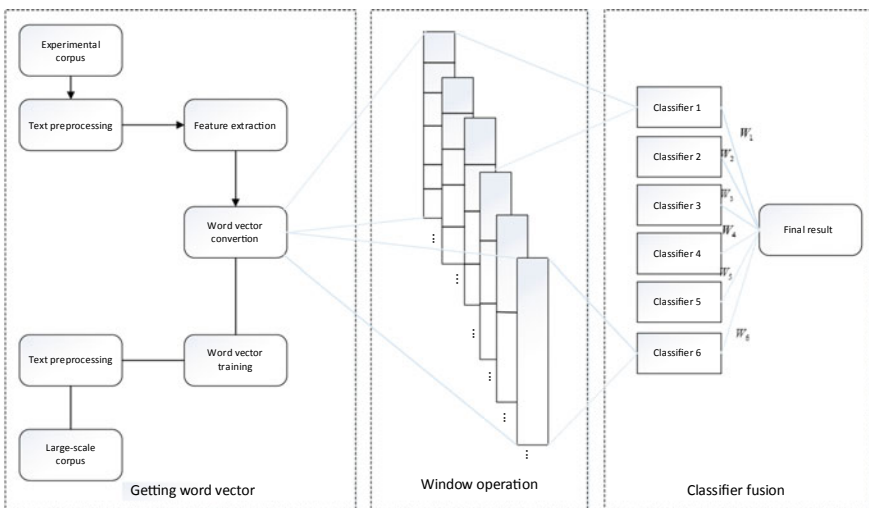


Fig. 1. Methodology adopted in this research

2.1 Windowed Word Vector

This paper uses the Word2Vec [11] to train word vector. In this experiment, the window size is 2, 3, 4, 5, 6, and 7. The windowing operation in Fig. 1 indicates that the former seven words in a sentence are operated. When the window size is 3, a 100-dimensional word vector early used to represent one word, the combined dimension will change from the original 100–300. For the same sentence, we use six different window size independently, so we can get word vector in six-dimensional combinations. Figure 2 shows window operation with a window size of 2, window shift of 1, used for a text with 7 words, then six combined dimensions will be gained.

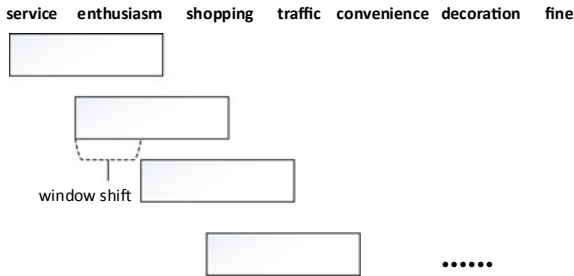


Fig. 2. Window operation, $W = 2$ (W is the window size)

2.2 Classifier Fusion

The classifier fusion diagram is shown in Fig. 3. In this paper, the weights obtained by the grid search algorithm are multiplied by the corresponding probability— P_{ij} by each classifier, and then sum to obtain the final probability— P_j of the corresponding text. A text label with a probability value more than 0.5 is denoted as a positive text. A text

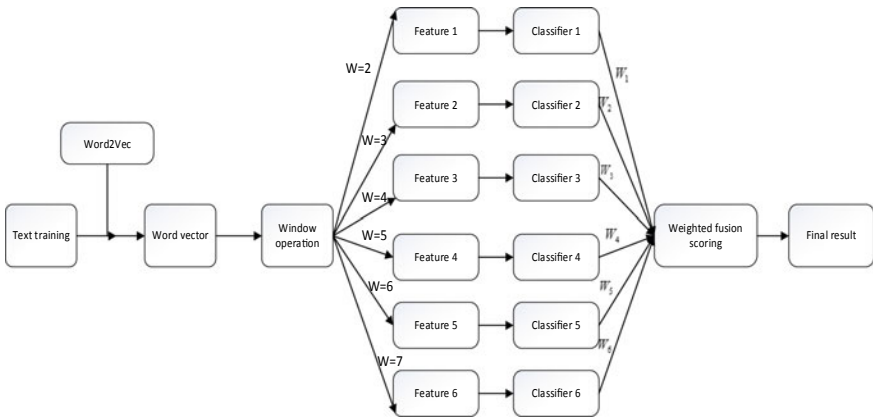


Fig. 3. Classifier fusion diagram

label with a probability value less than 0.5 is marked as a negative text. The final classification formula is as follows:

$$P_j = \sum_{i=1}^M W_i * P_{ij}(i = 1, 2, \dots, M)(j = 1, 2, \dots, N) \quad (1)$$

3 Experiments Methodology

3.1 Feature Extraction

The text data needs to be processed into a form understood by the computer. First, the Jieba (a tool for word segmentation) [12] is used to segment and part-of-speech tagging. This paper selects several features commonly used in text sentiment classification, unigram features, syntactic features, and lexicon-based features. Table 1 gives an example of “The Hotel’s environment is very good” to explain various features, where n denotes: noun, a denotes: adjective, and ad denotes: adverb.

Table 1. Feature representation

Features	Representation
Unigram	Hotel, environment, very, good
Syntactic	Hotel/n, environment/n, very/ad, good/a
Lexicon-based	Good

Text features that contribute a lot to the sentiment classification are preliminarily selected. If the feature’s dimensional space is too large, it will bring great calculations in the experiment. This paper uses Information Gain (IG) and Mutual Information (MI) to select features, which can not only reduce the dimension but also select the most distinguishing features. The experimental data comes from the Chinese emotional mining corpus—ChnSentiCorp, organized by Songbo, Tan [13]. The dataset contains 4,000 hotel review sentences, half positive, and half negative. After word segmentation on hotel reviews, the 2000-dimensional text features were selected by IG and MI.

3.2 Data Processing

In the experiment, the number of words in a sentence cannot be less than 7, due to the maximum window size of 7. Therefore, the sentence with less than 7 words is deleted, and the final experimental corpus includes 2175 review sentences in total, there are the training set, which includes 1957 samples, and the test set includes 218 samples.

The large-scale corpus which crawled from the internet is 15 GB. The dimension of word vector is 100, 200, 300, and 400 after word vector conversion. The experimental framework is shown in Fig. 1. For word vector in the same dimension, add the window size is 2, 3, 4, 5, 6, and 7, respectively. Thus, our model can obtain six different dimensional combinations, which are sent to the later classifiers in the decision layer.

3.3 Classifier Setting

In the experiment, SVMs and Logistic Regression (LR) were used for classifier fusion. SVM is a machine learning method proposed by Vapnik [14], and it has achieved good results in many classification problems. Support Vector Machines (SVMs) has been shown to have better results in sentiment analysis than other methods in practical research and experiments. In order to achieve the best classification results, after lots of experiments, the parameters of six classifiers models are set as in Table 2.

Table 2. Classifier setting

Dimension	w = 2	w = 3	w = 4	w = 5	w = 6	w = 7
100d	SVC ^a	SVC ^a	SVC ^a	LR	SVC ^a	LR
200d	SVC ^a	LR	LR	SVC ^b	SVC ^b	SVC ^b
300d	SVC ^a	SVC ^a	LR	LR	LR	SVC ^b
400d	SVC ^b	SVC ^a	LR	LR	SVC ^b	SVC ^b

Note The SVC^a kernel function is rbf, gamma = 1; SVC^b kernel function is poly, degree = 1, gamma = 1, and coef0 = 10; LR represents a logistic regression model with the default parameters

3.4 Results Analysis

The best classification results are selected as the baseline of the experiment, while the four different dimensional word vector without window. The Acc (Accuracy) on sentiment analysis results about using different dimensional word vector, and different window size in the experiment shown in Table 3. The results in Table 3 can be analyzed from two aspects. First, when the window size is constant, with the increase of the dimension of the word vector, the window size of 2, 3, 5, and 7, all show a significant improvement in accuracy rate. The accuracy rate of 88.1% was achieved at 400 dimensions. Because, a performance advantage of Word2vec is to discover the similarity between words. As the dimension of word vector gradually increases, its performance characteristics are more prominent. High-dimensional word vector compared with low-dimensional word vector can better capture the correlation between words and semantic connections.

Table 3. Acc on sentiment analysis using different dimensions and window

Dimension	Baseline (%)	w = 2 (%)	w = 3 (%)	w = 4 (%)	w = 5 (%)	w = 6 (%)	w = 7 (%)
100d	79.0	84.9	83.9	83.5	82.6	81.7	79.8
200d	79.5	85.3	84.9	84.4	82.6	81.2	80.3
300d	79.5	85.8	84.4	85.3	83.9	82.6	81.7
400d	80.3	88.1	85.3	84.9	84.4	82.1	80.7

Second, when the dimension of word vector is fixed, the correct rate presents a decreasing trend as the window size increases. We can explain this phenomenon from the perspective of the sparsity of feature vector space. For example, when window size of 7 on a 200-dimensional word vector, its dimension changes from 200 to 1400, while the amount of test corpus just only in 100. It is difficult for high-dimensional word vector to accurately capture the correlation between words in feature vector space. This leads to a deterioration of our final emotional classification results.

From the classification performance of each classifier in Table 2, in this research, the weight range is set to (0, 0.5), the step size is 0.05, the grid search method is used to find the optimal weight combination, and then calculate by formula (1) to obtain the final result in Table 4.

Table 4. Evaluation on sentiment analysis after classifier fusion

Dimension	Baseline				Windowed word vector			
	P (%)	R (%)	F1 (%)	Acc (%)	P (%)	R (%)	F1 (%)	Acc (%)
100d	82.3	76.1	79.1	79.0	84.2	87.3	85.7	85.3
200d	83.2	76.1	79.5	79.5	86.1	84.5	85.5	85.3
300d	82.6	76.6	79.6	79.5	88.2	88.2	88.2	88.1
400d	84.2	76.6	80.0	80.3	89.0	88.2	88.6	88.5

From the right part of Table 4, when the word vector dimension is 400, the highest accuracy, precision, recall, and F1: 88.5, 89, 88.2, 88.6% is generated. Compare to benchmarking datasets validates the effectiveness of our method in sentiment analysis. The results indicate that the proposed method can achieve better performance than the existing methods.

4 Conclusion

This paper proposed a sentiment analysis method for windowed word vector and classifiers fusion in the decision layer. This approach can make full use of the combined dimensional information to represent the characteristics of time series in sentiment. The experimental results clearly show that the precision rate is improved by approximately 4% compared with the baseline. It is more effective to use combined information in word vector in short-text sentiment analysis task.

As for future work, it is worth to detect the product review's sentiment is negative or positive, by reducing the dimension after windowed word vector. This approach can eliminate dimension sparse, and it may bring a more accurate sentiment classification result.

Acknowledgements. The paper is supported by the National Natural Science Foundation (61673108), China Postdoctoral Science Foundation (2016M601695), and Jiangsu University Natural Science Research Project (17KJB510018).

References

1. Medhat, W., Hassan, A., Korashy, H.: Sentiment analysis algorithms and applications: a survey. *Ain Shams Eng. J.* **5**(4), 1093–1113 (2014)
2. <https://wordnet.princeton.edu>
3. Keshavarz, H., Abadeh, M.S.: ALGA: adaptive lexicon learning using genetic algorithm for sentiment analysis of microblogs. *Knowl. Syst.* **122**, 1–16 (2017)
4. Chen, T., Xu, R., He, Y., et al.: Learning user and product distributed representations using a sequence model for sentiment analysis. *IEEE Comput. Intell. Mag.* **11**(3), 34–44 (2016)
5. Manek, A.S., Shenoy, P.D., Mohan, M.C., et al.: Aspect term extraction for sentiment analysis in large movie reviews using gini index feature selection method and SVM classifier. *World Wide Web-internet Web Inf. Syst.* **20**(2), 135–154 (2017)
6. Rout, J.K., Choo, K.K.R., Dash, A.K., et al.: A model for sentiment and emotion analysis of unstructured social media text. *Electron. Commer. Res.* 1–19 (2017)
7. Abbasi, A., Chen, H., Salem, A.: Sentiment analysis in multiple languages. *ACM Trans. Inf. Syst.* **26**(3), 1–34 (2008)
8. Salton, G., Wong, A., Yang, C.S.: On the specification of term values in automatic indexing. *J. Doc.* (1973)
9. Mikolov, T., Chen, K., Corrado, G., et al.: Efficient estimation of word representations in vector space. *Comput. Sci.* (2013)
10. Chen, T., Xu, R., He, Y., et al.: Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. *Expert Syst. Appl.* **72**, 221–230 (2017)
11. <https://code.google.com/p/word2vec/>
12. <https://www.oschina.net/p/jieba>
13. <http://www.searchforum.org.cn/tansongbo/senti-corpus.jsp>
14. Vapnik, V.: *The nature of statistical learning theory*. Springer, New York, Inc. pp. 988–999 (1995)



FFnet: Residual Block-Based Convolutional Neural Network for Crowd Counting

Fei Lei, Qinyu Zhang^(✉), Peng Zhao, Dongqiang Chen, Xiu Chen, and Xiao Han

Communications Engineering Research Center, Harbin Institute of Technology (Shenzhen), Guangdong, China
leifei547@163.com, zqy@hit.edu.cn

Abstract. Due to the nonuniform scale variations and severe occlusion, most current state-of-the-art approaches use multicolumn CNN architectures with different receptive fields to tackle these obstacles. We design a single-column network to verify the necessity of multicolumn network, and we find that under similar number of parameters and size of receptive field, single network is able to perform as well as multicolumn network. Following that, we propose a single-column network called FFnet based on residual block. FFnet is a fully convolutional network and easy to train. We perform extensive experiments on Shanghaitech dataset and the UCF_CC_50 dataset, and the results show that our method achieves a better performance than Switch-CNN with nearly half number of parameters, and a closing performance to the state-of-the-art model CP-CNN with almost one-tenth parameters.

Keywords: Crowd counting · ResNet · Density map estimation · Multi column · Receptive field

1 Introduction

Crowd counting aims to count the number of people in various density situations. On some occasions such as large sporting events or political demonstrations, accurate crowd counting provides an effective guidance for the arrangement planning and space design. Meanwhile, the method developed for crowd counting can also be extended to other counting tasks in other fields such as cell counting [1], vehicle counting [2, 3], livestock management, etc. Today, the task of crowd counting benefits from ubiquitous monitoring devices and the rapid development of computer vision algorithms. Like other computer vision tasks, crowd counting also confronts the extreme nonuniform distribution number of people and severe occlusions.

Motivated by the success of Convolutional Neural Networks (CNN) methods in computer vision tasks, many CNN-based methods have been developed

© Springer Nature Singapore Pte Ltd. 2020

Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 517,
https://doi.org/10.1007/978-981-13-6508-9_23

to address the problem of crowd counting. One of the most representatives is multicolumn convolutional neural network (MCNN) [4]. Different columns use different kernel sizes to capture the characteristics of crowd density at different scales. The follow-up works such as Switch-CNN [5], CP-CNN [6] keep this idea, achieving state-of-the-art results. The key point of using multiple CNN is that convolutional layers with different kernel size owns different receptive fields. However, the Inception net [7] revealed that large-kernel-size convolutional layers can be replaced by stack of convolutional layers with small kernel size. Following the above idea, this paper designs a single-column CNN network which has similar model complexity and receptive fields with MCNN to verify effectiveness of multicolumn architecture. It turns out that multicolumn architecture is not necessary. Then, this paper proposes a novel single-column network based on residual block called FFnet. The experimental results show FFnet achieves a closing performance compared with state-of-the-art methods with fewer parameters.

2 Related Work

There has been a long history of scholars trying to achieve better performance with several kinds of crowd counting algorithms. In the past decade, some traditional methods such as detection-based approaches, regression-based approaches, and density estimation-based approaches take charge for most of the circumstances.

2.1 Detection-Based and Regression-Based Approaches

In Wu's work [8], he presented an approach that could detect multiple people in a video stream. With part detectors, it has been proved to work with some people partially occluded. Ge et al. [9] designed an algorithm that can learn the shapes from videos. The framework they present can search targets with flexible shapes, able to detect and count people in crowded scene.

For the reason that detection-based methods work not very well when the crowds are extremely dense or background noise exists, regression approaches are used to increase algorithms stability. Chan et al. [10] designed a Bayesian Poisson regression approach to solve crowd counting problems from low-level features, deriving a approximation to the predictive distribution of the model. Since problems such as perspective, occlusion, and clutter make crowd counting problem too complex to be done, Idrees et al. [11] brought a new approach. In his algorithm, information from multiple sources is conflated. Their algorithm proved a better result on extremely dense crowd than other regression-based approaches.

2.2 CNN-Based Approach

To mitigate the influences from occlusion, the above methods choose to avoid considering some spatial information. But in density estimation-based

approaches, such information is included. Recently, CNN-based methods have been achieving much higher performance in all kinds of computer vision tasks than the old ones. Therefore, researchers are including CNN methods to solve crowd counting problems. Zhang et al. [12] designed an approach using CNN for crowd counting. By using a data-driven approach, they managed to solve problems of people occluded. Also, a dataset including various kinds of crowd was used. Zhang et al. [4] developed a method with Multicolumn Convolutional Neural Network (MCNN) to describe crowd density, and the model is tested with a new dataset. Sindagi et al. [6] created an end-to-end cascaded network of CNNs for crowd counting and density estimation, achieving a lower count error and better quality density maps. Last year, Sam et al. [5] proposed a model using Switching-CNN, so that the variation of crowd density is leveraged. By using such a network, crowd scene can be mapped to its density, solving problems of extreme crowding and high similarity.

3 Proposed Method

In this section, ground truth generation and data augmentation will be introduced first, and the following network training and evaluation are both based on them. The experiment on single-column network compared with MCNN will be shown in Sect. 3.2. Then, a deep network single-column network called FFnet will be discussed in the Sect. 3.3. The training details will be shown as well.

3.1 Data Preprocessing

The annotated datasets usually consist of two parts: the image and the coordinate of each head in the image. If there is a head at pixel x_i , it can be represented by the delta function $\delta(x - x_i)$, hence the whole image can be represented by the following function:

$$H(x) = \sum_i^N \delta(x - x_i) \quad (1)$$

the ground truth density map can be encoded by blurring each head annotation using a Gaussian kernel which is normalized to 1. More precisely, the density map is

$$F = \sum_{i=1}^N \delta(x - x_i) * G(x) \quad (2)$$

since the network may use the pooling layer, the output size of network may be inconsistent with the original shape. Therefore, down-sample (by bilinear interpolation) might be used when getting loss between predicting density map and ground truth.

Each image is randomly cropped to nine patches with 1/4 size of the original one, and all the patches are used to train the model. The multi-patches training ensures that the network adapts both high-density scene and low-density scene.

3.2 Comparison Between Single Network and MCNN

The network proposed by MCNN [4] consists of three branches with different convolutional kernel size. The maximal receptive field among three branches is 50, and the number of total parameters is 133.4k. In contrast, we propose a single-column network that is nearly the same as MCNN in terms of the number of parameters and the size of receptive field (see Table 1). The single-column network also uses two layers of max pooling, so the spatial resolution is reduced by 1/4 for each image. Similarly with MCNN, the single-column network is full convolution network too, which means there is no limitation on input image size. We initialize the model using Gaussian distribution with standard deviation of 0.01. The loss function is defined as follows:

$$L = \frac{1}{2N} \sum_{i=1}^N \|F_i - F_i^{gt}\|_2^2 \quad (3)$$

The comparison evaluated on Shanghaitech between our model and MCNN dataset part B is shown in Table 1. Under the approximate number of parameters and similar receptive field, our network outperforms the MCNN, reducing MAE by 3.7%, which indicates that multicolumn architecture is actually unnecessary for crowd counting problem.

Table 1. Comparison of single-column network and MCNN. The convolution layers are denoted as “conv(kernel size)-(number of filters)”, “/2” refers to stride of 2 and for pooling layer and convolution layer. The number of parameters, receptive field, and their behaviors on SHB are shown as well

Architecture	MCNN			Single column
	First column	Second column	Third column	
	conv9-16	conv7-20	conv5-24	conv3-32 conv3-32
	Max pool, /2			
	conv7-32	conv5-40	conv3-48	conv3-64 conv3-64
	Max pool, /2			
	conv7-16 conv7-8	conv5-20 conv5-10	conv3-24 conv3-12	conv3-64 conv3-32 conv3-16 conv3-8
Receptive field	50	36	24	49
Parameters	60,336	347,940	25,128	126,432
	133,404			
MAE	26.4			25.4
MSE	41.3			41.5

3.3 FFnet Architecture

Inspired by the above results, we propose a simple but deep single-column network called FFnet based on the residual block [13]. We remove the fully connected layers thus FFnet is also fully convolutional network. Batch normalization [14] layers are not included in the network. Similarly, the output of FFnet is $1/4$ of the input image size, which means the down-sample still need to be done at the training state. What is more, the reducing of output size is not fully caused by max pooling. Instead, it is replaced by a convolution layer with a

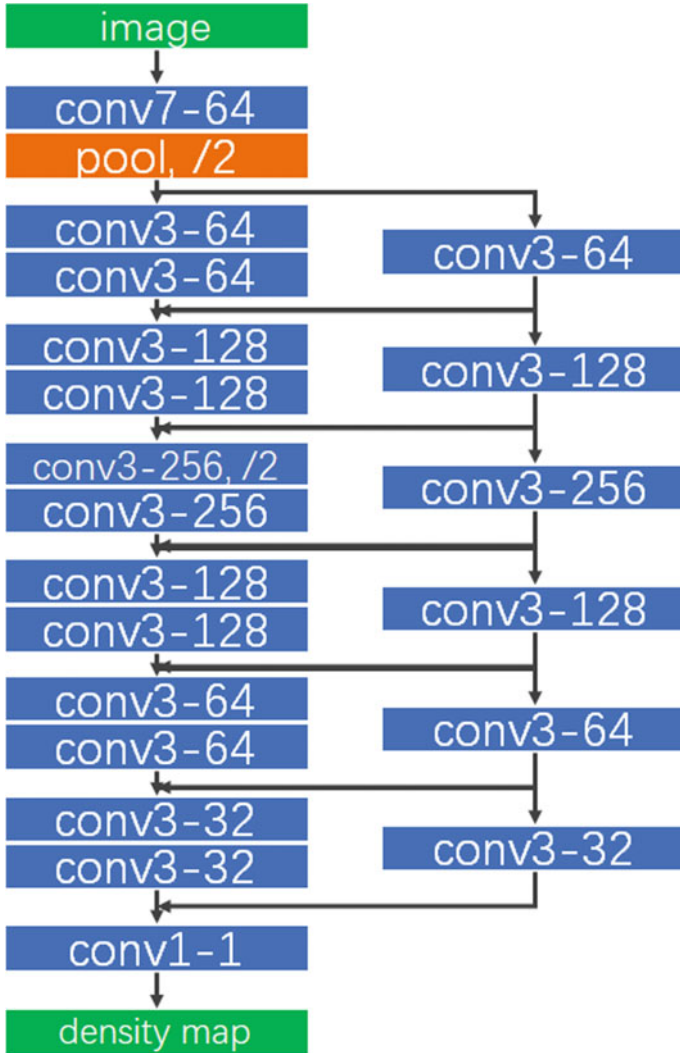


Fig. 1. Architecture of FFnet, Each convolution layer which is not shown in this figure follows an activation layer

stride of 2. Data augmentation and optimization method are the same with Sect. 3.2. The result of different datasets will be shown in the next section (Fig. 1).

4 Experimental Results

In this section, we evaluate the FFnet in two public datasets: ShanghaiTech [4] and UCF_CC_50 [11]. Mean Absolute Error (MAE) and Mean Squared Error (MSE) are used to evaluate the performance of model. They are defined as

$$MAE = \frac{1}{N} \sum_{i=1}^N |F_i - F_i^{gt}| \quad (4)$$

$$MSE = \sqrt{\frac{1}{N} \sum_{i=1}^N |F_i - F_i^{gt}|^2} \quad (5)$$

where N is the number of images and F_i is the prediction of i th density map and F_i^{gt} is ground truth density map corresponding to the i th sample. MAE measures the accuracy of model and MSE measures the stability of model (Table 2).

Table 2. Comparison of the number of parameters between several state-of-the-art models. M refers to million. Our proposed FFnet is 17% of Switching-CNN and 9% of CP-CNN

Method	Switching-CNN	CP-CNN	FFnet
Params	15.1 M	29.1 M	2.6 M

4.1 ShanghaiTech Dataset

The ShanghaiTech dataset was introduced by [4], and it consists of two parts: Part_A and part_B. There are 482 images with different resolution which is divided into a training dataset of 300 images and a test dataset of 182 test set, which comes from internet in Part_A. Part_B contains 716 images mainly taken

Table 3. Comparison results of different methods on Shanghaitech dataset

Method	Part_A		Part_B	
	MAE	MSE	MAE	MSE
Zhang [12]	181.8	277.7	32.0	49.8
MCNN [4]	110.2	173.5	26.4	41.3
Switching-CNN [5]	90.4	135.0	21.6	33.4
CP-CNN [6]	73.6	122.2	20.1	32.7
<i>FFnet (ours)</i>	90.3	122.2	20.3	32.7

from the busy streets in Shanghai with the same resolution of 768×1024 , and also Part_B is divided into a training dataset of 400 images and a test dataset of 316 images. The whole dataset have annotated 1198 images with a total of 330,165 people annotated. Our model is compared with some typical CNN-based network of crowd counting issue (shown in Table 3). Most of these work use MCNN as their backbone, and combine the multi-task learning.

Table 4. Comparison results of different methods on UCF_CC_50 dataset

Method	MAE	MSE
Idrees [11]	419.5	541.6
Zhang [12]	467.0	498.5
MCNN [4]	377.6	509.1
Switching-CNN [5]	318.1	439.2
CP-CNN [6]	295.8	320.9
<i>FFnet (ours)</i>	310.4	367.2

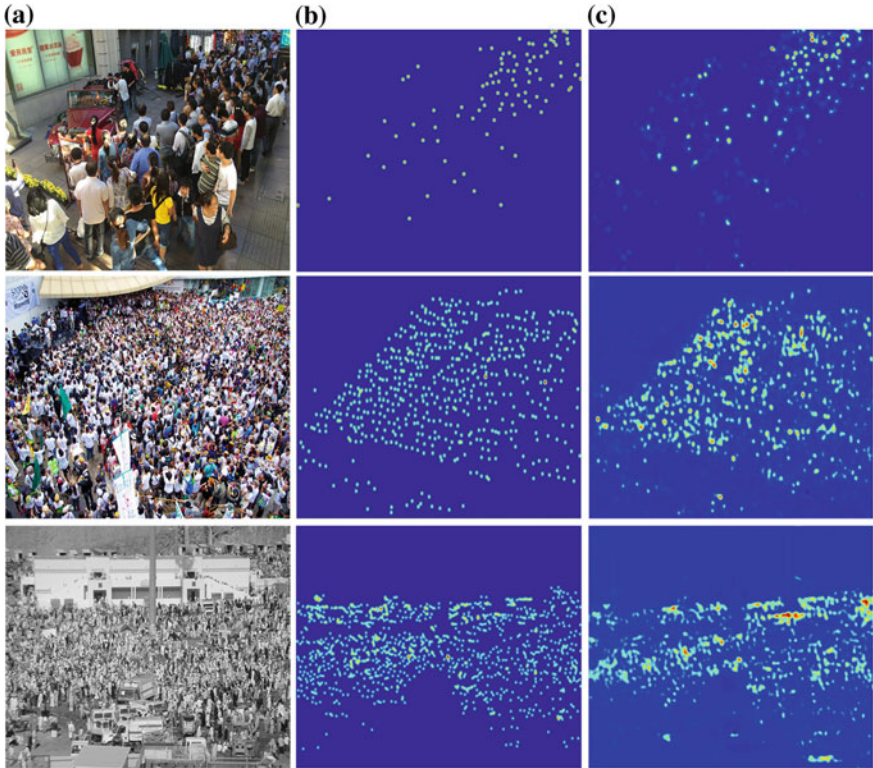


Fig. 2. Density map estimation results of FFnet. **a** Original. **b** Ground truth. **c** Prediction of density map. The first row is from Shanghaitech Part_A, the second row is from Shanghaitech_B, and the last row is from UCF_CC_50

4.2 UCF_CC_50 Dataset

The UCF_CC_50 is a extremely challenging dataset introduced by [11]. There are only 50 images with different resolution. The distribution of the number of people is so imbalanced that the head counts range from 94 to 4543 per image. There are 63974 individuals in total annotated in the UCF_CC_50 dataset. Fivefold cross validation is performed following the standard setting in [11]. It can be observed from Table 4 that our approach, though using fewer parameters, achieves close MAE and MSE with the approach mentioned above (Fig. 2).

4.3 Conclusion

In this paper, we show that multicolumn networks with different receptive fields are actually unnecessary for crowd counting issue. Then we propose a novel single-column network FFnet based on residual block. FFnet is a simple, easy-trained, and end-to-end network. Compared with networks applying multicolumn architecture as backbone such as Switching-CNN and CP-CNN, our FFnet uses one-sixth number of parameters to achieve a better performance than Switching-CNN and close behaviors with one-tenth parameters as CP-CNN.

Acknowledgement. This work was supported in part by the National Sciences Foundation for the Distinguished Young Scholars of China under Grant 61525103, and the Shenzhen Fundamental Research Project under Grant JCYJ20150930150304185.

References

1. Walach, E., Wolf, L.: Learning to Count with CNN Boosting (2016)
2. Lempitsky, V.S., Zisserman, A.: Learning to count objects in images. In: International Conference on Neural Information Processing Systems (2010)
3. Ooro-Rubio, D., Lpez-Sastre, R.J.: Towards Perspective-Free Object Counting with Deep Learning, pp. 615–629 (2016)
4. Zhang, Y., et al.: Single-image crowd counting via multi-column convolutional neural network. In: IEEE Conference on Computer Vision and Pattern Recognition (2016)
5. Sam, D.B., et al.: Switching convolutional neural network for crowd counting. *Comput. Vis. Pattern Recognit.* (2017)
6. Sindagi, V.A., Patel, V.M.: Generating high-quality crowd density maps using contextual pyramid CNNs. In: IEEE International Conference on Computer Vision (2017)
7. Szegedy, C., et al.: Rethinking the inception architecture for computer vision. *Comput. Sci.* 2818–2826 (2015)
8. Wu, B., Nevatia, R.: Detection and tracking of multiple, partially occluded humans by bayesian combination of edgelet based part detectors. *Int. J. Comput. Vis.* **75**(2), 247–266 (2007)
9. Ge, W., Collins, R.T.: Marked point processes for crowd counting. In: IEEE Conference on Computer Vision and Pattern Recognition, 2009. CVPR (2009)
10. Chan, A.B., Vasconcelos, N.: Bayesian Poisson regression for crowd counting. In: IEEE International Conference on Computer Vision (2010)

11. Idrees, H., et al.: Multi-source multi-scale counting in extremely dense crowd images. In: IEEE Conference on Computer Vision and Pattern Recognition (2013)
12. Zhang, C., et al.: Cross-scene crowd counting via deep convolutional neural networks. In: IEEE Conference on Computer Vision and Pattern Recognition (2015)
13. He, K., et al.: Deep Residual Learning for Image Recognition, pp. 770–778 (2015)
14. Ioffe, S., Szegedy, C.: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, pp. 448–456 (2015)



Deep Learning-Based V2V Channel Estimations Using VNETs

Qi Song^(✉), Tian Lan, Xuanxuan Tian, and Tingting Zhang

Shenzhen Graduate School, Communication Engineering Research Center,
Harbin Institute of Technology, Shenzhen 518055, China
song_sqi@163.com

Abstract. The development of cooperative intelligent transportation systems brings new challenges to wireless communication technologies, where the channel estimation becomes more and more important. In this paper, a novel data-driven channel estimation method based on deep learning framework is adopted. Based on the feedforward neural network, the VNET neural network based on the convolutional neural network is proposed. The simulations and practical measurements are also provided to verify the performance advantages. The results show the achieved performance advantages of the proposed VNET-based method, which is shown to be an effective solution.

Keywords: Channel estimation · Neural network · OFDM · Deep learning · CNN

1 Introduction

With the advancement of urbanization, traffic has become an increasingly serious problem in urban areas. In order to effectively deal with these issues, the network among vehicles came into being [1]. With the help of the Internet of Vehicles, on the one hand, smart driving can be achieved more quickly and safely, and the number of car accidents caused by human operations can be effectively reduced. On the other hand, it can be unified to prevent congestion and save resources [2, 3]. However, due to the high-speed motion and high real-time requirements of vehicles, a more severe test has been put forward for wireless communications. These challenges are everywhere. Among the most difficult to solve is the estimation of the wireless channel. In the car networking environment, the hypothesis of the generalized stationary channel is no longer valid. Traditional pilot-based channel estimation methods such as LS and MMSE have been unable to compete in a real multivariable environment. At this time, an effective channel estimation method is urgently needed to promote the further development of car networking. In recent years, the deep learning method has made great progress in various fields such as image processing, speech recognition, and machine translation.

Similarly, deep learning methods have long been introduced in the traditional communication field and have achieved good results [4–6].

Based on the performance of the feedforward neural network in wireless channel estimation, this paper proposes the application of convolutional neural networks to deal with the problem of wireless channel estimation in high-speed moving vehicle scenarios for the first time, and validates the performance of the algorithm in simulation and measured data. It further proves the great potential of deep learning methods in wireless channel estimation. The following organizational structure of the article is as follows. The second chapter introduces the basic content of deep learning, mainly for the feedforward neural network model; the third chapter focuses on the VNET convolutional network model proposed in this paper, and introduces the wireless vehicle communications system model used; the fourth chapter describes in detail the results of the model based on simulation data and measured data; the fifth chapter briefly summarizes the full text.

2 Deep Learning Basics

Neural network is a kind of network structure based on neurons. According to the different ways of connections between neurons, it can be converted into different network models [7]. Unit neurons can be illustrated in Fig. 1. The output of the unit neurons \mathbf{z} can be mathematically expressed as

$$\mathbf{z} = \sigma \left(\sum_{i=1}^I x_i w_i + b \right) \tag{1}$$

where I stands for the number of inputs, w_i denotes the weights and b is the bias. Function σ is called activation function, which can be the Sigmoid function or the Relu function [8].

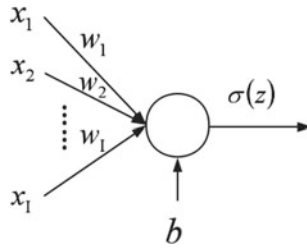


Fig. 1. One hidden layer neuron network architecture

One of the most classic network structures in neural networks is feedforward neural networks which are also known as fully connected network [9]. In the fully connected network, every two neurons have the weights to connect except the same layer neurons as shown in Fig. 2.

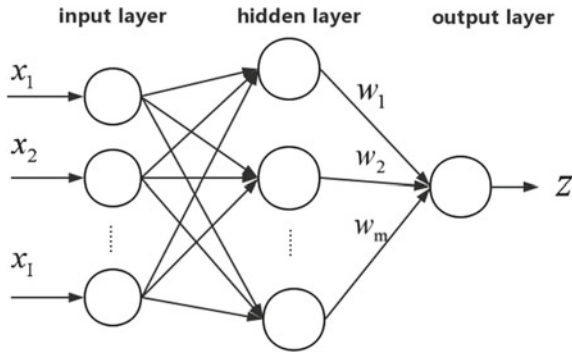


Fig. 2. Fully connected neuron network architecture

3 Learning-Based V2V Channel Estimation Method

3.1 System Model

The wireless vehicle communications system model used in this paper is shown in Fig. 3. The OFDM waveform is used in this paper, which can be mathematically expressed as

$$x(t) = e^{j2\pi f_c t} \sum_{m=0}^{N_s-1} \sum_{n=0}^{N_c-1} a_{n,m} e^{j2\pi n \Delta f t} \text{rect}\left(\frac{t - T_{sym}}{T_{sym}}\right) \tag{2}$$

where f_c denotes the carrier frequency, N_s stands for the number of OFDM symbols, N_c is the number of subcarriers, $a_{n,m}$ is the transmitted symbol, $T_{sym} = T + T_g$ is the total OFDM symbol duration composed of an elementary symbol duration T and a cyclic prefix(CP) duration T_g , $\Delta f = 1/T$ is the subcarrier separation, $\text{rect}(t/T_{sym}) = 1$ for $0 \leq t \leq T_{sym}$ and zero elsewhere. The baseband OFDM signal model is the same as the traditional one.

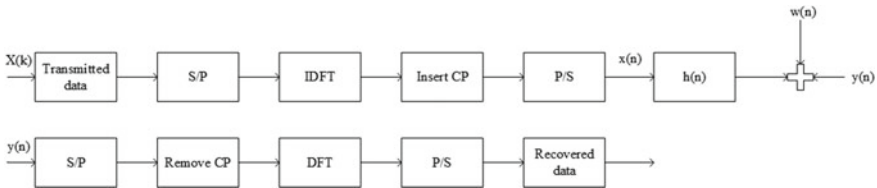


Fig. 3. The system model

3.2 Feedforward Neural Network

For typical applications of vehicle communications, the input of the deep feed-forward neural network is the received communication signal, and the output is the original transmission sequence after demodulation [10]. The entire network has five layers, and the number of neurons in each layer is 128, 55, 38, 25, and 64 in order, where the activation function used is the Sigmoid function except the last layer, and the function used in the last layer is the Sign function. In the process of designing the feedforward neural network, the main considerations are the form of the cost function, the choice of the optimization algorithm, the regularization method, and the generation of the test set.

The form of the cost function Since the cost function formed by BER is discrete, so in the actual training process, we choose an indirect indicator, that is, the cross-entropy of the ideal output value and the predicted value as the cost function of the entire neural network, mathematically expressed as

$$C = -\frac{1}{N} \sum_x [y_- \ln y + (1 - y_-) \ln(1 - y)] \tag{3}$$

where N denotes the number of data in the dataset, y_- stands for the ideal output value, and y is the neural network predicted output value.

The choice of the optimization algorithm The main neural network optimization algorithms are gradient descent methods based on backpropagation [11]. Because the cost functions to be optimized in the neural network are almost all non-convex functions, these optimization algorithms can only obtain a local optimal solution, which means that the performance of the algorithm depends heavily on the selection of initial value weights. To solve this problem, we add the pretraining process before the formal training as shown in Fig. 4, and completes the selection of initial value weights through an automatic encoder, effectively improving the performance of the optimization algorithm [12].

The regularization method Because of the high complexity of the neural network model, the overfitting phenomenon tends to occur during the training

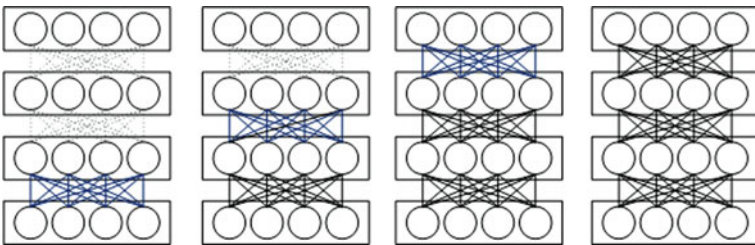


Fig. 4. The pretrain process

process, which reduces the generalization ability of the model and fails to achieve the expected results. One way to effectively reduce the overfitting is regularization, also known as normalization. The regularized form chosen in this paper is L2 regularization (weight decay). The idea of L2 normalization is to add an extra item to the cost function. This item is called the normalization item. The following is the normalized cross-entropy:

$$C = -\frac{1}{N} \sum_{x_j} [y_{-j} \ln y_j^L + (1 - y_{-j}) \ln(1 - y_j^L)] + \frac{\lambda}{2N} \sum_w w^2 \quad (4)$$

where the first term is the conventional cross-entropy expression and the second item is the sum of squared weights. Then, use a factor $\lambda/2N$ to quantize the adjustment, and $\lambda > 0$ can be called the normalization parameter, and N is the size of the training set.

The generation of the test set Due to the existence of overfitting, it is impossible to use training error as a standard of generalization ability. In practice, we can evaluate the generalization errors of the network model through experimental tests and make further choices. At this time, a “validation set” will be introduced to test the network model’s ability to predict new samples, and then “validation error” will be used as an approximation of the generalized error. The assumption is that the validation sample is also derived from an independent and identically distributed sample from the true distribution of the sample. The validation set should be mutually exclusive with the training set. Under normal circumstances, we only have one data set containing m samples, which need to be trained and validated. We need to perform appropriate processing to generate training sets and validation sets. The selection method used in this paper is cross validation.

3.3 VNET Model

In a fully connected network, we do not premeditate the property of the input data itself. Meanwhile, the numbers of the parameters of weights and biases in the network tend to be spanning huge which can make the training process difficult to converge. Therefore, the convolution neural network is more suitable for wireless channel estimation problem in the vehicle communication scene. Based on the understanding of the communication system and the outstanding performance of the feedforward neural network in wireless channel estimation, in order to improve the performance of wireless channel estimation in the car networking environment, we designed a convolutional neural network for recovering

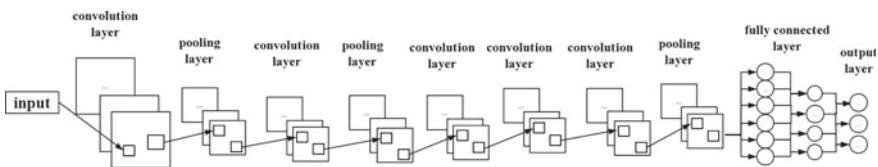


Fig. 5. The architecture of VNET

transmitted information, namely VNET. The structure of the VNET is shown in Fig. 5. For VNET, only some nodes are connected between two adjacent layers. In order to show the dimensions of each layer of neurons, the nodes of each layer of the convolution layer are usually organized into a two-dimensional matrix.

There are two important structures in VNET, namely convolution layer and the pooling layer. Convolution layers, also known as filters, can transform a sub-node matrix on the current level of neural networks into a single unitary node matrix on the next level of neural networks as illustrated in Fig. 6. After the filter operation, the first sub-node matrix becomes a unit node matrix with a value of $\sum_{i=1}^9 x_i w_i$. Next, as shown in Fig. 7, a series of unit node matrices is

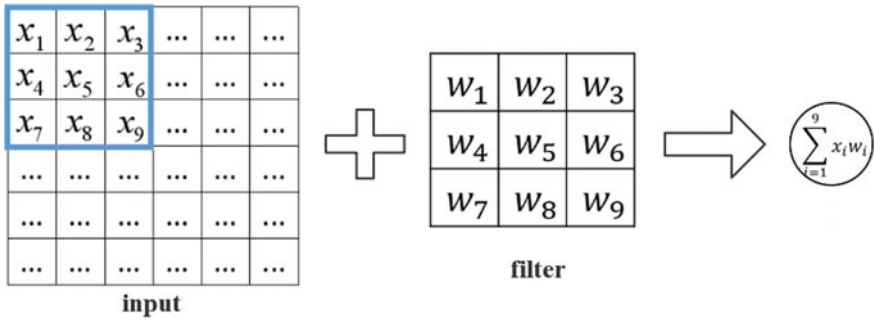


Fig. 6. The convolution process

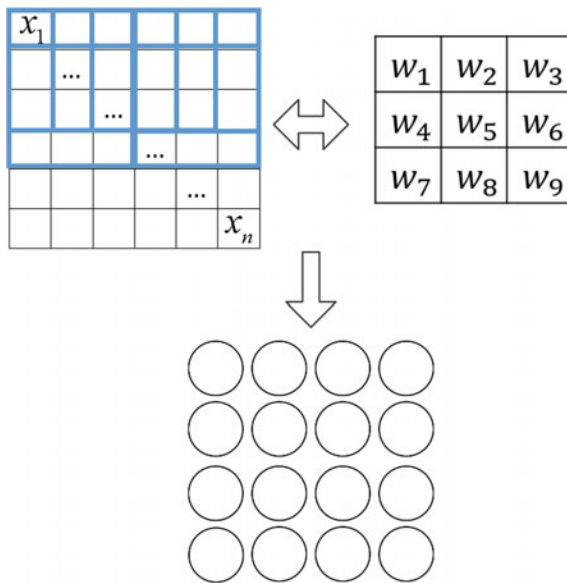


Fig. 7. The moving process

obtained from the range of the movable sub-node matrix, and finally get a brand new two-dimensional matrix.

Pooling layers can be very effective in reducing the size of the matrix, reducing the parameters in the final fully connected layer. Using a pooling layer can both speed up calculations and prevent overfitting. Convolution layer forward propagation is done by moving the filter, but the pooling layer directly selects the maximum or average of the sub-node matrix elements to complete the move with no complicated calculation. In this article, the maximum pooling is used.

3.4 Model Training

In the model training process, we regard the inter-vehicle communication environment as a black box. Whether it is feedforward neural network or VNET, their generalization ability depends heavily on the data. In this paper, the training data and test data are from not only simulations, but also practical measurements. In the feedforward neural network, the input data is the received signal vertical arrangement of the formation of a one-dimensional vector, the output is the recovered transmission symbols; for the VNET, the input data is also received signal, but need to be transformed into a square matrix form, the output is still the recovered transmission symbols.

In order to ensure the generalization ability of the model, we deal with the data set accordingly. First, the data set is divided into 100 similar-sized mutually exclusive subsets, each subset as much as possible to maintain the consistency of data distribution, that is, are obtained through hierarchical sampling. Then, each time regards 99 subsets of the union as a training set, the remaining subset as a test set; so we can get 100 sets of training and test set, which can be 100 training and testing process, the final return is the average of 100 test results.

In the testing process, the performance of the model is evaluated using the classic bit error rate of communication metrics.

4 Results Analysis

4.1 Simulation Data Section

Based on the particular application scenario of the OFDM wireless communication system and vehicle-mounted communications, the simulation process of generating training data was performed under the DSRC protocol. The specific parameters are as follows: the carrier frequency is 5.9 GHz, the bandwidth is 10 MHz, the number of subcarriers is 64, and the communication distance is about 400 m. After the training data is generated, the feedforward neural network and the VNET model are respectively applied to training, and the results are compared with the traditional LS channel estimation method, as shown in Fig. 8.

It can be seen that the channel estimation method based on deep learning is obviously better than the traditional channel estimation method. At the same time, the VNET model proposed in this paper is also superior to the feedforward neural network in performance.

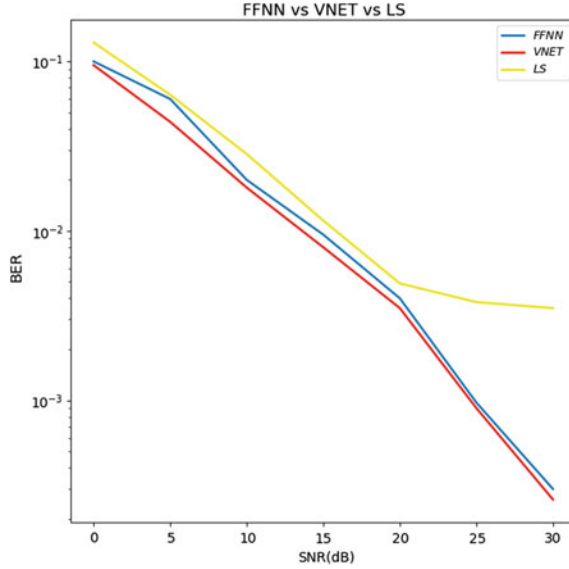


Fig. 8. The performance comparison among FFNN, VNET, LS

4.2 Measured Data Section

We were fortunate to have an opportunity to actually collect data, totaling 100,000 samples. We applied the feedforward neural network and VNET model to the measured data. After a difficult training process, we finally got a bit error rate of nearly 0.005. The VNET model is better than the feedforward neural network in the measured data regardless of the training process or the generalization ability. Because it is measured data, the results cannot be compared with the traditional methods, but the resulting performance has far exceeded our expected.

5 Conclusion

This paper proposes a new data-driven channel estimation method for the special scenarios of vehicle-mounted wireless communication. It can effectively improve the BER performance when the conventional pilot-based wireless channel estimation method cannot cope with it well. Based on the feedforward neural network, the VNET model based on the convolutional neural network was innovatively proposed and verified on the simulation data and the measured data, which confirmed the effectiveness of the proposed algorithm. Limited by the difficulty of acquiring datasets, future research will focus on the construction of datasets in specific scenarios to further enhance the generalization capabilities of the VNET model.

References

1. Bimeyer, N., St ubing, H., Schoch, E., Gotz, S., Stolz, J.P., Lonc, B.: A generic public key infrastructure for securing Carto-X communication. In: 18th World Congress on Intelligent Transport Systems (2011)
2. David, C.M., Caroline, C.L., Benjamin, D., Douglas, Y.: Driver distraction and advanced vehicle assistive systems (ADAS): investigating effects on driver behavior. *Adv. Human Aspects Transp.* **484**, 1015–1022 (2016)
3. Erica, D., Omar, A.S., Azhar, S., Haider, K., Carpenter, D.O.: Road traffic injury as a major public health issue in the Kingdom of Saudi Arabia: a review. *Front Public Health* **4**, 215 (2016). <https://doi.org/10.3389/fpubh.2016.00215>
4. X. Wang, L. Gao, S. Mao and S. Pandey, 2017. CSI-based fingerprinting for indoor localization: A deep learning approach, *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 763–776, Jan. 2017
5. OShea, T.J., Hoydis, J.: An introduction to machine learning communications systems. *CoRR* (2017). [arXiv:1702.00832](https://arxiv.org/abs/1702.00832)
6. Jiang, C., Zhang, H., Ren, Y., Han, Z., Chen, K.-C., Hanzo, L.: Machine learning paradigms for next-generation wireless networks. *IEEE Wirel. Commun.* **24**(2), 98–105 (2017)
7. Goodfellow, I., Bengio, Y., Courville, A.: Deep learning. MIT Press (2016). <http://www.deeplearningbook.org>
8. Cybenko, G.: Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst.* **2**(4), 303–314 (1989)
9. Schmidhuber, J.: Deep learning in neural networks: an overview. *Neural Netw.* **61**, 85–117 (2015)
10. Ye, H., Li, G.Y.: Power of deep learning for channel estimation and signal detection in OFDM systems. *IEEE Wirel. Commun. Lett.* **7**(1) (2018)
11. Papa, G., Clemencon, S., Bellet, A.: SGD algorithms based on incomplete U-statistics: large-scale minimization of empirical risk. *Neural Inf. Process. Syst.* 1027–1035 (2015)
12. Vincent, P., et al.: Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. *J. Mach. Learn. Res.* 3371–3408 (2010)



Research on Optimization of Evacuation Path Based on Fuzzy Control and Intelligent Agent Technology—With Open Block Emergency Evacuation as Example

Fu Erkang¹, Zhang Lingfei^{2(✉)}, Li Xinyun³, and Zhang Yujia¹

¹ College of Landscape Architecture, Sichuan Agricultural University, Room 605, Teaching Building #5, No. 211 Huimin Rd., Wenjiang District, Chengdu 611130, Sichuan Province, China

² Faculty of Architecture and Urban Planning, Science of Urban and Rural Planning, Chongqing University, No. 174 Shazheng Road, Shaping District, Chongqing 400030, China

+86 18628325690351813657@qq.com

³ College of Landscape Architecture, Sichuan Agricultural University, Room 606, Teaching Building #5, No. 211 Huimin Rd., Wenjiang District, Chengdu 611130, Sichuan Province, China

Abstract. Evacuation path has an important influence on emergency evacuation efficiency at open space, but the evacuation path based on the shortest path is not the best choice. Such environmental variables as coverage effect of 5G signal is an important basis for intelligent guiding signs to dynamically guide the stream of people for evacuation, with bigger intervention effect with a choice of evacuation path of crowd, influencing the science of planning of evacuation path. In the paper, Yulin open block of Chengdu City is taken as the object of empirical study, and layout of guiding signs is adjusted using fuzzy control and intelligent agent technology to improve distribution of the stream of evacuated people in the space, screen out the evacuation path with high pass rate, and low congestion rate. Research result shows that the effect of the evacuation signs at the existing congestion point is not obvious at all. Fuzzy control can optimize spatial distribution of the stream of evacuated people by adjusting and controlling the degree of familiarity of evacuated crowd with different exits.

Keywords: Open block · Intelligent agent technologies · Fuzzy control · Evacuation simulation

1 Introduction

Reasonable evacuation sign guidance and control can effectively increase evacuation efficiency of space, help people to quickly recognize the direction and location of evacuation [1], speed up the evacuation of evacuees, reduce congestion [2–4].

However, in practice, many evacuation signs have poor diversion effect, unscientific location of crowd diversion, and incomplete coverage of wireless signals such as 4G, resulting in the overall evacuation efficiency has not improved with the increase of guidance signs. Existing studies mainly planned paths from the perspective of the shortest path but neglect the impact of environmental variables on human behavior. As a result, the shortest path is often not the best path. Such paths have the problems of space congestion and low evacuation efficiency, which make it difficult to effectively guide practice. Therefore, in this paper, evacuation signs are used as the intervention variables of crowd path selection. Fuzzy control and intelligent agent technology are used to simulate the intervention of evacuation signs on crowd migration path, so as to improve the scientific of evacuation path.

Evacuation sign is spatial evacuation environmental information. Its transmission efficiency is influenced in two aspects including visibility of signs and readability of contents. Visibility refers to whether the object can see and recognize evacuation sign, including visible range, light condition, location of evacuation, readability of contents refers to whether the object can accurately understand the direction information transmitted by signs, including the color, graph, text of evacuation signs. Evacuation sign visibility and content readability can produce certain fuzziness when they interact, for example, how to judge if the information is transmitted accurately in the case where visibility is strong but readability is weak. So, fuzzy control and intelligent agent methods are adaptive when they are used distribution and adjustment of evacuated human flow. On this basis, in this paper, with Yulin open block Chengdu as an example, by using fuzzy control and intelligent agent technology to adjust and control the variable value of the familiarity with exits, the influence of adjustment and control of guiding signs on evacuation path is researched.

2 Research Method

2.1 Intelligent Agent Evacuation Model

Crowd evacuation model includes macro model and micro model, macro model mainly uses hydromechanics concept to explain and predict human flow, including the famous queen model, random model, and transition matrix model. With the development of computer technology, evacuation model enters micro phased, putting the focus of simulation on the individual, accurately predict the location and speed of every person at every moment, including cellular automation model and social force model which begin to consider the behavior characteristics of the crowd as evacuation variable. Intelligent agent has functions as follows: (1) sensing environment information and the cooperation information of other users, (2) with its own knowledge base and rule base, can constant evolve by learning such knowledge, (3) makes decision on its own

according to sensed information, (4) with response capability, etc. On the basis of fully analyzing the detail of behavior of people during evacuation and the influencing factors, by using the methods of mapping the evacuated people as the agent, mapping individual characteristics of evacuated people as the status attribute of the agent, mapping various behaviors spontaneously taken, or taken due to motivation of external events by the evacuated as the rule of the agent, we establish a person evacuation model at public place which comprehensively considers the influence of external factors such as competitiveness and judging ability of evacuated people, herd behavior between evacuated people during evacuation, as well as the effect of guidance, occurrence of fire, passage clogging, random event, etc. on evacuated people.

2.2 Fuzzy Control

In order to explore the judgment result of the crowd toward the information of evacuation sign, in the research fuzzy control model is introduced. Fuzzy control is a method using the knowledge of fuzzy set theory, fuzzy linguistic variable, and fuzzy logical reasoning to simulate fuzzy thinking of people, which can simulate inaccurate, nonlinear information processing capability of human brain to understand objective world [5], is widely used in the field of image identification. In the research two fuzzy factors are described with language, conditional statements (rule) corresponding to different fuzziness are set, output control result is judged by fuzzy control, where fuzzy control value adopts the variation value of exit familiarity.

3 Research on General Situation of Object and Distribution of Crowd

3.1 General Situation of Research Object

Located in the northeast of Wuhou District Chengdu City, Yulin open block is a traditional open block of Chengdu forming earlier, with the characteristics of early form of “small block”. Because it was built earlier, the street is narrow, inside livelihood roads and traffic roads interweave, under emergency situation the path for internal residents to excavate outward is not clear. According to field survey, there are 8 evacuation signs, 6 target evacuation exits in Yulin block currently, forming internal evacuation path network (as shown in Fig. 1).

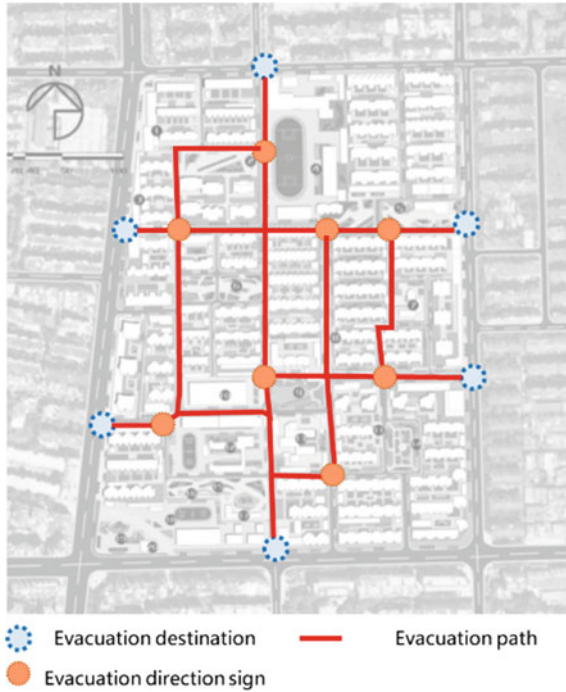


Fig. 1. YULIN open block evacuation situation

3.2 General Situation of Distribution of Population

In the form of questionnaire and observation at a fixed point based on zone, a survey on the spatial distribution of population and exit familiarity of Yulin open block is conducted. Data collection work was conducted from February to March 2018, totally select four working days and four weekend days (Fig. 2).

In order to accurate observation of population distribution of the block, this research on distribution adopts observation method, dividing the west street into four subregions to observe in turn in weekends with fine whether, ensuring every round of observation to cover the whole block. The time frames for observation are 10:00–12:00, 12:00–14:00, 14:00–16:00, 17:00–20:00, respectively, recording the data on actors in the same subregion in the same time frame every time, taking the highest value in different time frames as the data on population in the round, adding the number of people of 4 subregions and averaging the sum to obtain the number of population. Through six rounds of observation and statistics, the number of evacuated people at west street under emergency situation is 2,850.



Fig. 2. Yulin open block people distribution

4 Guiding Sign Adjustment and Control Based on Fuzzy Control and Intelligent Agent Technology

4.1 Input/Output Variable and Rules of Fuzzy Control

Input two variables of “content visibility of sign” and “content readability of sign”, with every variable classified as four fuzzy classes: “very good”, “good”, “poor”, “very poor”, being represented with NB, NS, PS and PB respectively. Assume content visibility of sign to be V whose quantified field is {4, 3, 2, 1}, assume content readability of sign is P whose quantified field is {4, 3, 2, 1}. Output variable is the increment of exit familiarity E, being represented with “very high”, “high”, “common”, “low”, “very low”, with quantified field of {40, 30, 20, 10, 0}. The increment of output exit familiarity can be represented by rule matrix as shown in Table 1.

Table 1. Rules of fuzzy control

V	E			
	NB	NS	PS	PB
NB	NB	MB	MB	PS
NS	MB	MB	MB	PS
PS	MB	MB	PS	PB
PB	PS	PS	PB	PB

4.2 Fuzzy Control Model Used for Controlling Exit Familiarity

According to subordinating degree function of input/output variables and fuzzy control rule, content visibility of sign is represented with x_1 , content readability of sign is represented with x_2 , increment of exit familiarity is represented with y , then the mathematic model of exit familiarity increase control system after evacuation signs are added at the intersection can be represented with the following equation.

$$y = \frac{\sum_{l=1}^M y^{-l} \cdot \mu A_1^l(x_1) \cdot \mu A_2^l(x_2)}{\sum_{l=1}^M y^{-l} \cdot \mu A_1^l(x_1) \cdot \mu A_2^l(x_2)}$$

The equation is the mathematic model of sign visibility x_1 and sign readability x_2 on variation value of exit familiarity, through MATLAB simulation software, the three-dimensional graph of its corresponding evacuation variation value is shown in Fig. 3.

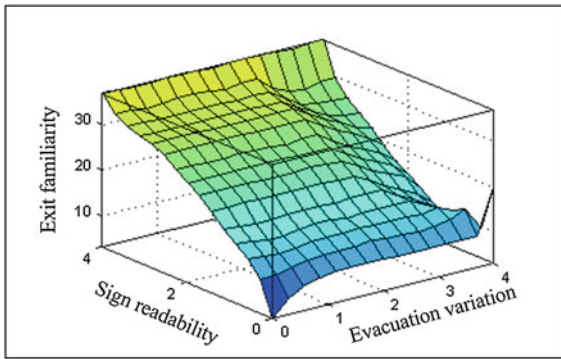


Fig. 3. 模糊控制数值函数图

4.3 Dynamic Precise Evacuation Model Based on Intelligent Agent Technology

Choice of evacuation path is the key to people evacuation model. Assume that in virtual planar space of public place, the conversion locus of Agent grid location status constitutes its evacuation path, with attraction probability of grid as the main basis of agent selecting evacuation path, then the basic rule of agent behavior of evacuated people can be established.

Definition 1 Grid attraction probability when Agent itself adopts the shortest path is

$$P_{(i,j)} = a_{(i,j)} \times P_{dis(i,j)}$$

where $P_{dis(i,j)}$ represents the location attraction probability of grid $N(i, j)$, $a(i, j)$ is the adjustment coefficient of attraction probability $P(i, j)$ decided by grid attribute, when a

$(i, j) = 1$, it indicates that grid $N(i, j)$ is usable; when grid is occupied by other Agent, wall or barrier, then $a(i, j) = 0$, indicating that the grid cannot be chosen. Location attraction probability $P_{dis}(i, j)$ is determined according to the distance of the path from every grid to evacuation exit, with computational formula as follows:

$$P_{dis(i,j)} = \frac{[\max(D_{(i,j)}) - D_{(i,j)}]}{[\max(D_{(i,j)}) - \min(D_{(i,j)})]}$$

where, $D(i, j)$ is the distance of grid $N(i, j)$ to the closest exit along the shortest path, $\max(D(i, j))$ and $\min(D(i, j))$ represent the maximum value and minimum value of all grids to the closest exit along the shortest path. It can be found from the equation that the closer to evacuation exit the grid is, the bigger its location attraction probability; on the contrary, the farther from evacuation exit a grid is, the smaller its location attraction probability.

Definition 2 Grid attraction probability when Agent itself adopts herd behavior completely is

$$P_{(i,j)} = a_{(i,j)} \times P_{dir(i,j)}$$

where, $a(i, j)$ is the adjustment coefficient of attraction probability $P(i, j)$ decided by grid attribute. $P_{dir}(i, j)$ represents direction attraction probability of grid $N(i, j)$, which can be obtained by the following equation:

$$P_{dir(i,j)} = \frac{N_{dir(k)}}{\sum_{k=1}^8 N_{dir(k)}}$$

The equation indicates that the more the people moving in a direction, the bigger the attraction of the direction.

5 Evacuation Simulation and Evacuation Path Analysis

5.1 Influence of Evacuation Sign Optimization on Crowd Evacuation Path

According to the calculation based on intelligent agent formula, adjust the location of existing eight evacuation signs and conduct evacuation simulation. It is known from comparison of evacuation efficiency that the evacuation velocity of the open block after improvement is higher than that of the scenario before improvement, overall evacuation time decreasing from previous 458 to 389 s; at the same time the crowd evacuation efficiency of the unused idle path before improvement increases dramatically, indicating that those emergency exits often neglected under guidance of evacuation signs are made full use of (see Figs. 4 and 5).

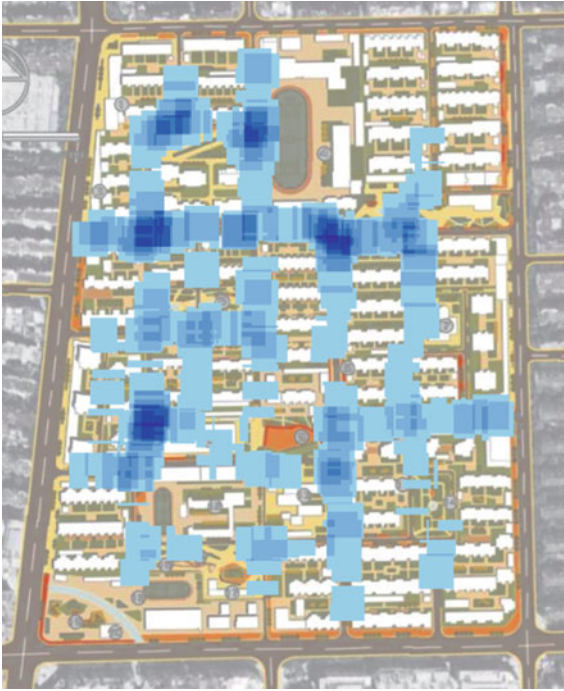


Fig. 4. Evacuation simulation before improvement



Fig. 5. Evacuation simulation after improvement

Input fuzzy set through fuzzy control toolbox of MATLAB simulation software, add the conditional statement rule of Table 1, calculate the degree of variation of evacuation familiarity of every exit, (see Table 2), after improvement the familiarity and perception of people with exit change obviously, on one hand, the familiarity of people with original exit is homogenized dramatically, avoiding the situation of large amount of people flocking to a few exits, reducing the possibility of congestion and stampede during evacuation; on the other hand, people perceive five newly added evacuation destinations, further distributing human flow to more evacuation exits, increasing path choice of crowd evacuation.

Table 2. Comparison of familiarity with evacuation exit before and after adjustment and control

Evacuation target exit	1	2	3	4	5	6	7	8	9	10	11
Familiarity with exit before improvement	75	15	12	13	60	65	0	0	0	0	0
Variation value of familiarity with exit	-27.5	+36.6	+36.6	+36.6	-27.5	-27.5	18.5	21.5	15.5	10.5	14.5
Familiarity with exit after improvement	47.5	51.6	48.6	49.6	32.5	37.5	18.5	21.5	15.5	10.5	14.5

Note Familiarity with exit before improvement was obtained according to questionnaire

5.2 Evacuation Path

From comparison of crowd evacuation distribution diagram before and after crowd evacuation (Figs. 4 and 5) it can be found that after adjustment crowd evacuation path changed to some extent, more fully using the access roads of open block for evacuation, increasing evacuation path by 6, increasing evacuation exit selected by people by 5. After adjustment the area covered by evacuation network increases by 24%, increasing the safety and reliability of evacuation for whole Yulin community (Fig. 6).



Fig. 6. Sketch of evacuation path after optimization

6 Discussion and Conclusion

Through simulation and comparison before and after simulation and verification, it is found that the evacuation efficiency, congested area and exit utilization of Yulin open block can be improved by adding evacuation signs, and the following can be concluded.

6.1 Fuzzy Control Can Optimize Reasonability of Spatial Distribution of Evacuated Human Flow

In the evacuation case of Yulin open block existing “main passage—access road”, “main passage—emergency exit” two-level guiding signs, “access road—emergency exit” are absent, the crowd lost in access road will be attracted to flock to main road due to herd effect, some closer secondary exits are ignored by the crowd. After improvement based on fuzzy control, part of people are guided to the exits closer to access roads, relieving the pressure on main evacuation paths, playing a role of separating some people from the main road exit with higher human traffic.

6.2 The Effect of Evacuation Sign at Congestion Point Is Not Obvious

Existing research has confirmed that in the same evacuation environment, there is an inflection point for the number of effective evacuation signs, to set up evacuation signs more than a certain number cannot change the overall evacuation effect [6], so it is needed to reasonably distribute the location for evacuation signs in the block. It is found in the research that original evacuation sign location and the location congested overlap highly, but the cause of congestion at congestion point is the convergence of human flow from upstream several adjacent human flows, so evacuation signs set up at the upstream of evacuation point play more obvious role than evacuation sign at congestion point. Evacuation simulation can be used in planning for various historical blocks as the technologic support of judging the location of congestion point and direction of evacuated human flow.

6.3 Limitation of Research and Outlook for Future

In this research optimization of distribution of evacuated people using fuzzy control in the case achieved a good result, but there are still some imperfect parts, during simulation people's judgment to path not only includes the variable of evacuation sign controlled this time, psychological variables including group psychology, relatives effect also influence people's choice for path, this research did not include these factors in fuzzy model considering operability of data acquisition. In the future, fuzzy control model can be perfected by combining 3D simulation technology and VR visual technology to quantify psychologic factors, so as to help further optimize the visual presentation of evacuation sign, improve safety management level of the outdoor public place.

References

1. Baoqing, G.: Brief introduction of fire emergency light system. *Build. Electric.* **23**(z1), 91–95 (2004)
2. Benthorn, L., Frantzich, H.: Fire alarm in a public building: how do people evaluate information and choose an evacuation exit? *Fire Mater.* **23**(6), 311–315 (1999)
3. Proulx, G., Kyle, B., Creak, J.: Effectiveness of a photoluminescent wayguidance system. *Fire Technol.* **36**(4), 236–248 (2000)
4. Shields, T.J., Boyce, K.E.: A study of evacuation from large retail stores. *Fire Saf. J.* **35**(1), 25–49 (2000)
5. Passino, K.M., Yurkovich, S.: *Fuzzy Control*. Tsinghua University Press (2001)
6. Lijuan, D.: *Research on Evacuation Effect of Emergency Signs*. Beijing University of Chemical Technology (2010)



Feature Engineering of Click-through-rate Prediction for Advertising

Jie Ren^(✉), Jian Zhang, and Jing Liang

University of Electronic Science and Technology of China, Chengdu, China
2356670185@qq.com

Abstract. We present the problem of click-through-rate (CTR) for search advertising in ALiMaMa, which displays user information, item information, shop information and trade results. Traditionally, people use logistic regression (LR) to predict it. However, because of the lack of learning ability and the sparse feature matrix, the prediction results are always not so satisfying. In this paper, we mainly propose some feature engineering methods based on gradient boosting decision tree (GBDT) and Bayesian smoothing to obtain a wonderful feature, which has more useful information and is not so sparse. Also, we use xgboost (XGB) instead of LR as our prediction model. The proposed methods are evaluated using offline experiments and the experiment results prove that the log loss drop near 5% after using these feature engineering methods and XGB. Obviously, it is an excellent performance.

Keywords: CTR · Feature engineering · GBDT · Bayesian smoothing · XGBoost

1 Introduction

CTR is a term commonly used in Internet advertising, referring to the click to reach rate of online advertisement, i.e. the actual number of clicked times of an advertisement dividing by the amount of displayed times of this advertisement.

CTR is very important in e-commerce and web-scale applications, but the extremely unbalanced samples make it difficult to be predicted. Although many machine learning (ML) algorithms such as deep neural network (DNN), GBDT and LR are used to deal with this type of problem, the performance will just improve a little if there is no good feature engineering. Therefore, feature engineering plays a vital role in CTR prediction.

Feature engineering can transform raw data into features which has a better performance in prediction models. So, wonderful feature engineering means an improved model accuracy on unseen data [1]. The quality of the features determines the upper limit of a model. Hence, how to obtain available features by feature engineering becomes significant.

Using GBDT to process features may be a helpful way [2]. Also, Google provided another online learning algorithm called FTRL-Proximal (FTRL) [3, 4]. It is like stochastic gradient descent, but a much sparser model. Apart from these models, the end-to-end model [5, 6] such as DNN, especially convolutional neural network (CNN), is also used to automatically process features. What is more, manually crafted combinatorial features [7] are still the most useful in the industry.

In this paper, we divide CTR prediction into two steps. First, we need to use feature engineering methods to preprocess raw data. In Sect. 2, we introduce how to make artificial features at first. After having artificial features, we use GBDT to automatically separate continuous values into different levels, and merge these levels into existed features. Then, because there are so many 'id' features which obviously have some hidden information but make the feature matrix massive and sparse, we proposed to replace them with the results of Bayesian smoothing. Second, having finished feature engineering, there are n samples and m features for each sample. In Sect. 3, we briefly introduce XGB which is famous recently, and then we send a $n \times m$ matrix into this prediction model. To prove that our feature engineering is useful, in Sect. 4, we repeatedly calculate log loss by adding the methods in Sect. 2 one by one, and the experiment results are evident. Finally, in Sect. 5, conclusions are presented.

2 Feature Engineering

2.1 Artificial Features

Artificial features are necessary due to the fact that raw features may be not suitable for predicting.

First, one-hot encoder is necessary to make some discrete features whose digital form is meaningless be meaningful. For example, a feature such as 'student_id' has three possible values '01', '02', '03'. One-hot encoder divides this feature into three features 'student_id.01', 'student_id.02', 'student_id.03' and each feature only has two values 0 and 1. This step is important, but it absolutely makes the feature matrix be very massive and sparse.

Second, using standardization $Y = (X - \mu)/\sigma$ to disperse close data is also useful to make the performance be better.

Third, analyse the relationship between features and labels, then divide the features into different levels and merge the level features to raw data.

Fourth, adding some cross features by multiplying or adding two features together sometimes shows wonderful performance. However, one thing that designers should pay attention to is that improper cross features have a bad influence on your model.

2.2 New Features Designed by GBDT

The features designed by people must be subjective. An advanced method is using GBDT to train the training data and the outputs of leaf nodes are the new features. The principle of GBDT adopts the linear combination of decision

trees. The boosting tree model can be represented as the additive model of decision trees

$$f_M(x) = \sum_{m=1}^M T(x; \Theta_m) \tag{1}$$

where $T(x; \Theta_m)$ means decision trees; Θ is the parameter of decision trees; M is the number of trees.

GBDT adopts forward stagewise algorithm. First, confirm initial boosting tree $f_0(x) = 0$. Later, the m th step model can be written as

$$f_m(x) = f_{m-1}(x) + T(x; \Theta_m) \tag{2}$$

where $f_{m-1}(x)$ is the current model. The parameters of next decision tree is confirmed by using empirical risk minimization:

$$\hat{\Theta}_m = \arg \min_{\Theta_m} \sum_{i=1}^M L(y_i, f_{m-1}(x_i) + T(x_i; \Theta_m)) \tag{3}$$

where $L(y, f(x))$ is loss function. When we use square error as loss function,

$$\begin{aligned} L(y, f_{m-1}(x) + T(x; \Theta_m)) &= [y - f_{m-1}(x) - T(x; \Theta_m)]^2 \\ &= [r - T(x; \Theta_m)]^2 \end{aligned} \tag{4}$$

where $r = y - f_{m-1}(x)$ is residual, and the purpose of boosting tree is to make residual as small as possible.

GBDT uses an approximate method of the steepest descent method. The key is to use the negative gradient of the loss function in the current model as the approximate value of the residual in the boosting tree algorithm. This value can be written as

$$r_{mi} = - \left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)} \right]_{f(x)=f_{m-1}(x)} \tag{5}$$

where m is the m th tree and i is the i th sample. Later, use r_{mi} to fit a tree.

Now, the question becomes how to use GBDT to train new features? The idea is that use existed features to train a GBDT model. Then, when we have new data, we can put these data into the trees of this GBDT model and the values of leaf nodes are new features. Finally, new features merge with existed features as the features used to train the final model.

2.3 Bayesian Smoothing

Another difficult problem is how to use hidden information such as ‘id’. Bayesian smoothing is a very useful method.

Bayesian smoothing comes from probabilistic machine learning. There are two assumptions. First, all advertisements have their own ctr, and these ctr obey Beta distributions. Second, for a specific advertisement, given its impression count and its ctr, the clicked times obey Binomial Distribution (I, ctr).

$$p(r|\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} r^{\alpha-1} (1-r)^{\beta-1} \quad (6)$$

$$P(C|I, r) \propto r^C (1-r)^{I-C} \quad (7)$$

where r is sample ctr and $r \sim \text{Beta}(\alpha, \beta)$, C is sample clicks and $C \sim \text{Binomial}(I, r)$, and $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$ is Gamma distribution.

Assume that there are N advertisements, the likelihood function can be written as follows:

$$\begin{aligned} P(C_1, C_2, \dots, C_N | I_1, I_2, \dots, I_N, \alpha, \beta) &= \prod_{i=1}^N P(C_i | I_i, \alpha, \beta) = \prod_{i=1}^N \int_{r_i} P(C_i, r_i | I_i, \alpha, \beta) dr_i \\ &= \prod_{i=1}^N \int_{r_i} P(C_i | I_i, r_i) p(r_i | \alpha, \beta) dr_i \propto \prod_{i=1}^N \int_{r_i} r_i^{C_i} (1-r_i)^{I_i-C_i} r_i^{\alpha-1} (1-r_i)^{\beta-1} dr_i \\ &= \prod_{i=1}^N \frac{\Gamma(\alpha + \beta)}{\Gamma(I_i + \alpha + \beta)} \frac{\Gamma(C_i + \alpha)}{\Gamma(\alpha)} \frac{\Gamma(I_i - C_i + \beta)}{\Gamma(\beta)} \end{aligned} \quad (8)$$

Then, the gradient of log likelihood function can be written as follows:

$$\begin{aligned} &\frac{\partial \log P(C_1, C_2, \dots, C_N | I_1, I_2, \dots, I_N, \alpha, \beta)}{\partial \alpha} \\ &= \sum_{i=1}^N [\Psi(\alpha + \beta) - \Psi(I_i + \alpha + \beta) + \Psi(C_i + \alpha) - \Psi(\alpha)] \\ &\frac{\partial \log P(C_1, C_2, \dots, C_N | I_1, I_2, \dots, I_N, \alpha, \beta)}{\partial \beta} \\ &= \sum_{i=1}^N [\Psi(\alpha + \beta) - \Psi(I_i + \alpha + \beta) + \Psi(I_i - C_i + \beta) - \Psi(\beta)] \end{aligned} \quad (9)$$

where $\Psi(x) = \frac{d}{dx} \ln \Gamma(x)$ is digamma function. The iteration optimization function can be written as follows:

$$\begin{aligned} \alpha^{new} &= \alpha \frac{\sum_{i=1}^N [\Psi(C_i + \alpha) - \Psi(\alpha)]}{\sum_{i=1}^N [\Psi(I_i + C_i + \alpha) - \Psi(\alpha + \beta)]} \\ \beta^{new} &= \beta \frac{\sum_{i=1}^N [\Psi(I_i - C_i + \beta) - \Psi(\beta)]}{\sum_{i=1}^N [\Psi(I_i + \alpha + \beta) - \Psi(\alpha + \beta)]} \end{aligned} \quad (10)$$

After calculating the value of α and β and the prior probability of each “id” C/I , the posterior probability of the “id” can be calculated by using the formula $r = (C + \alpha)/(I + \alpha + \beta)$.

According to experiments, when I is large enough, the correction by using Bayesian smoothing has just a little influence on ctr. However, when I is small, the correction is very useful to improve ctr [8].

Finally, after having Bayesian smoothing feature, in order to avoid to have an overfitting model, we should divide it into different levels and use the levels

as a new feature. Even, we can just use the levels as our new features and delete the original ‘id’ features to make the matrix not so sparse.

3 Prediction Model

3.1 Loss Function

In CTR, the common loss function is cross entropy:

$$l = -\frac{1}{N} \sum_{i=1}^N (y_i \log(p_i) + (1 - y_i) \log(1 - p_i)) \tag{11}$$

where N is the number of test samples, y_i is actual label of i th sample and p_i is the predicted rate of i th sample. Absolutely, the smaller log loss means a better performance.

3.2 Xgboost

GBDT is a wonderful algorithm to predict CTR. However, when we use GBDT to train n th tree, we need the residual of $n - 1$ th tree. Therefore, GBDT is difficult to be parallel computing. XGB is an improvement of GBDT, which can be parallel computing. Also, XGB has more developments than GBDT.

The objective function can be written as

$$Obj^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \tag{12}$$

where l is the loss function, and $\Omega(f_t)$ is the regularization which includes $l1$ and $l2$. Then, use Taylor expansion to approximate this objective function. Taylor expansion can be written as

$$f(x + \Delta x) \simeq f(x) + f'(x)\Delta x + \frac{1}{2}f''(x)\Delta x^2 \tag{13}$$

Then, we can define that as follows:

$$\begin{aligned} g_i &= \partial_{\hat{y}_{(t-1)}} l(y_i, \hat{y}_i^{t-1}) \\ h_i &= \partial_{\hat{y}_{(t-1)}}^2 l(y_i, \hat{y}_i^{t-1}) \end{aligned} \tag{14}$$

So, the objective function can be written as

$$Obj^{(t)} \simeq \sum_{i=1}^n \left[l(y_i, \hat{y}_i^{t-1}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \tag{15}$$

If we split the tree into structure part $q(x)$ and weight part w_n , we can obtain $f_t(x) = w_{q(x)}$. Also, the regularization can be written as

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (16)$$

where γ is the number of leafs and λ is the penalty coefficient of l_2 regularization. Finally, the new objective function is

$$\begin{aligned} Obj^t &\simeq \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i w_{q(x_i)}^2 \right] + \Omega(f_t) \\ &= \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i w_{q(x_i)}^2 \right] + \gamma T + \lambda \frac{1}{2} \sum_{j=1}^T w_j^2 \\ &= \sum_{j=1}^T \left[\left(\sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] \end{aligned} \quad (17)$$

where $I_j = \{i | q(x_i) = j\}$, and the function can be simplified as

$$Obj^{(t)} = \sum_{j=1}^T [G_j w_j + \frac{1}{2} (H_j + \lambda) w_j^2] + \gamma T \quad (18)$$

According to $\frac{\partial Obj^{(t)}}{\partial w_j} = 0$, $w_j^* = -\frac{G_j}{H_j + \lambda}$ is obtained. Then,

$$Obj = -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma T \quad (19)$$

where Obj represents the maximum reduction of the objective function when we assign a tree structure, which is called structure score. The smaller this score is, the better the tree structure is. Later, we use greedy algorithm to try to split existed leafs.

$$Gain = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right] - \gamma \quad (20)$$

where the first to third part in bracket means the score of left subtree, right subtree, whole tree without split, respectively. γ is the cost of complexity by adding a new leaf node.

Finally, enumerate all possible scenarios to get the result. Therefore, the parallelism of XGB is feature parallelism.

4 Experiment Results

Finally, we use two experiments to prove the efficiency of our methods. First, we add different feature engineering, respectively, and then we train them by using our prediction model to obtain their log loss. Second, we use different prediction models to show that our model is useful.

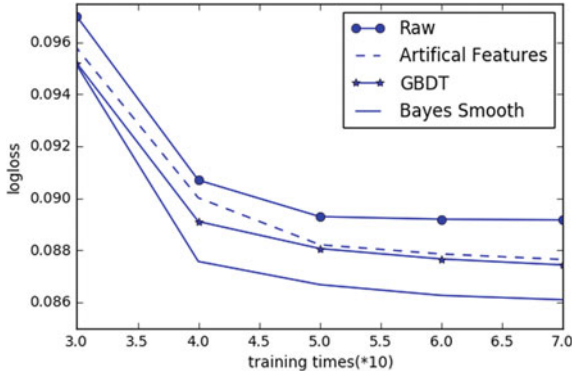


Fig. 1. The performance by adding new features

Figure 1 shows the training performances by continually adding different feature engineering methods. We can easily find that the performance increases dramatically after doing feature engineering.

Figure 2 is the testing data performance. It demonstrates that dnn which is an end-to-end memory network has the ability to select features to some extent. This is the reason why it has a wonderful performance than traditional machine learning algorithms. However, dnn is not suitable for CTR because there is a high dimension feature matrix after one-hot encoder. It may make the network too time consuming to be trained. Also, we can find that if a good feature engineering is done, the prediction performance will increase dramatically.

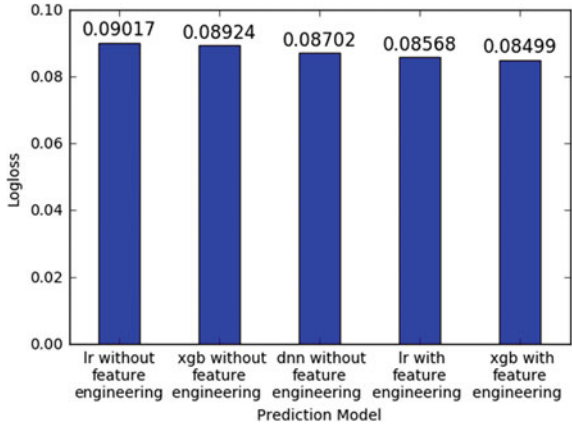


Fig. 2. The performance of different model

5 Conclusions

In conclusion, we can easily find that cross features have the best performance in artificial features. They make log loss reduce from about 0.0089 to near 0.0088. Also, even if GBDT do just a few contributions to prediction result, it certainly reduces log loss. Amazingly, log loss drops from 0.088 to 0.085 by using Bayes smooth. What is more, we can find that even if LR will obtain a better performance than DNN, if effective feature engineering is processed. Finally, if we use our feature engineering methods and prediction model, the log loss will reduce 5% than that with LR and no feature engineering.

Acknowledgements. This work was supported by the National Natural Science Foundation of China (61671138, 61731006), and was partly supported by the 111 Project No. B17008.

References

1. Discover Feature Engineering, How to Engineer Features and How to Get Good at It. <https://machinelearningmastery.com/discover-feature-engineering-how-to-engineer-features-and-how-to-get-good-at-it/s>
2. Bowers, S., et al.: Practical lessons from predicting clicks on ads at Facebook. In: Eighth International Workshop on Data Mining for Online Advertising. ACM, pp. 1–9 (2014)
3. He, X., Chua, T.S.: Neural factorization machines for sparse predictive analytics. In: The, International ACM SIGIR Conference. ACM, pp. 355–364 (2017)
4. McMahan, H.B., Holt, G., Sculley, D., et al.: Ad click prediction: a view from the trenches. In: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, pp. 1222–1230 (2013)
5. Sukhbaatar, S., Szlam, A., Weston, J., et al.: End-to-end memory networks. *Comput. Sci.* (2015)
6. Cheng, H.T., Koc, L., Harmsen, J., et al.: Wide and deep learning for recommender systems, 7–10 (2016)
7. Shan, Y., Hoens, T.R., Jiao, J., et al.: Deep crossing: web-scale modeling without manually crafted combinatorial features. In: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, pp. 255–262 (2016)
8. Wang, X., Li, W., Cui, Y., et al.: Click-through rate estimation for rare events in online advertising. *Online Multimed. Adv. Tech. Technol.* (2011)



Application of Fuzzy C-Means Algorithm in Complex Background Image Segmentation of Forensic Science

Zhuang Chen, ChunYu Li^(✉), ZhanQing Jiang, and Yongqiang Zhao

Forensic image technology direction, People's Public Security, University of China (PPSUC), No. 1, South Muxidi Lane, Beijing Xicheng District, People's Republic of China
lichunyu@ppsuc.edu.cn

Abstract. In the field of forensic science, image segmentation is required as a basic and significant stage in forensic image analysis. It is very important to segment the stamp impression image with a complex background precisely. This paper puts forward a feasible and efficient approach for complex background stamp impression image segmentation based on Fuzzy C-Means (FCM) algorithm. The fuzzy feature of forensic image can be handled efficiently using Fuzzy C-Means (FCM) algorithm in the forensic science field. The results of the experiments demonstrate the validity and accuracy of Fuzzy C-Means (FCM) algorithm.

Keywords: Image segmentation · Fuzzy C-Means · Forensic science · Stamp impression · Complex background

1 Introduction

The stamp impression, one of the most commonly examined objects in the field of forensic science, is the reflective image of structural morphology of seal surface of the stamp on the receiving body. In our daily life, stamp impressions are sealed on all kinds of contracts and bills to prove their validity. However, the cases of forging seals occur frequently because people are driven by interests. Because of the limit of other external conditions, such as the complex background, it is impossible to detect the case by using traditional inspection instruments. Based on the principle and method of image segmentation, segmenting the specific parts of stamp impressions with a complex background and extracting relevant features is of great significance on the analysis and verification of subsequent stamp impression.

As an important part of artificial intelligence and the critical link of image processing, image segmentation has been extensively applied to the fields of transportation, pattern recognition, etc. With the purpose of dividing an image into several regions sharing certain uniformity, image segmentation is the process of clustering all similar pixels in an image, so as to extract people's interested objects from complex scenes for further analysis and processing. Based on laboratory work, the commonly

employed methods of image segmentation include image threshold value segmentation method, image edge detecting method, method based on neural networks, etc.

Fuzzy C-Means (FCM) algorithm as an unsupervised fuzzy clustering technique has been widely applied in the field of image segmentation since the fundamentals of this theory proposed by J. C. Dunn was established. FCM algorithm is an approach of flexible partition, which has proper local convergence. As an improvement of the Hard C-Means (HCM) algorithm, its image segmentation effect is superior to the results of the HCM algorithm segmentation. This algorithm is based on the key idea: it makes objects divided into several clusters, making a cluster of objects in connection with high similarities and different clusters with the minimum similarity.

In forensic science, there are only a few researches focusing on the image segmentation of stamp impression under complex background. On the basis of Fuzzy C-Means clustering algorithm, this paper segments the stamp impression images with complex background. Meanwhile, it provides a method for analyzing and verifying the stamp impressions in forensic science, which plays high realistic and innovative significance to the development of forensic document and image examination technologies.

2 Methodology

There are various kinds of clustering analysis algorithms, among which FCM algorithm is one of the widely used and successful algorithms. We use FCM algorithm in this paper.

2.1 Principle of FCM Algorithm

According to the membership function, FCM algorithm determines the degree to which each pixel belongs to a cluster. Using the initialization method to determine several initial clustering centers, this algorithm is an iterative optimization which continuously adjusts and optimizes the clustering centers to achieve the minimal variance within clusters. The basic idea is as follows.

The iterative optimization's objective function is defined as

$$J_f(t) = \sum_{j=1}^C \sum_{i=1}^N [U_j(x_i)]^{\rho} (d_{ij})^2 \quad (1)$$

where the following constraints:

$$\begin{aligned} U_j(x_i) &\in [0,1], \forall j, i \\ \sum_{j=1}^C [U_j(x_i)] &= 1, \forall i \end{aligned} \quad (2)$$

where $X = \{x_1, x_2, \dots, x_N\} \in R^{N \times q}$ denotes a data set consisting of N samples of dimension q to be divided into C classes. x_i represents features data is a q -dimensional vector; $\rho (\rho > 1)$ is a constant representing the weighted index that controls fuzziness of the resulting partition. The value of $\rho (\rho > 1)$ is usually set to $\rho = 2$.

Using the Euclidean norm, the distance metric d_{ij} means the distance between sample x_i and a cluster center Z_j . The Euclidean distance $d_{ij} = \|x_i - Z_j\|$. $Z = \{z_1, z_2, \dots, z_C\} \in R^{N \times q}$ denotes the set of C cluster centers.

The membership functions and cluster centers are updated by the following expressions:

$$\begin{aligned}
 U_j(x_i) &= \frac{\left(\frac{1}{\|x_i - Z_j\|^2}\right)^{1/(\rho-1)}}{\sum_{j=1}^C \left(\left(\frac{1}{\|x_i - Z_k\|^2}\right)^{1/(\rho-1)}\right)} \rho \\
 &= \frac{1}{\sum_{j=1}^C \left(\frac{\|x_i - Z_j\|}{\|x_i - Z_k\|}\right)^{2/(\rho-1)}}
 \end{aligned} \tag{3}$$

and

$$Z_j = \frac{\sum_{i=1}^N [U_j(x_i)]^\rho x_i}{\sum_{i=1}^N [U_j(x_i)]^\rho} \tag{4}$$

2.2 Steps of FCM Algorithm

The implementation steps of FCM algorithm is described as follows:

- Step 1: Initialize the input parameters C , ρ ; Select $\varepsilon > 0$, that denotes a threshold of stopping the iteration; Set maximum permission iterative number to a specific value.
- Step 2: Initialize the input membership function $U_j(x_i)$.
- Step 3: The center for each cluster Z_j is calculated with reference to $U_j(x_i)$.
- Step 4: Calculate the Euclidean distance d_{ij} for every pixel in each cluster with reference to Z_j .
- Step 5: The objective function $J_f(t)$ and $J_f(t+1)$ is calculated for each cluster.
- Step 6: If $|J_f(t+1) - J_f(t)| < \varepsilon$, the clustering stops; otherwise $t = t + 1$ and it goes back step 2.

3 Experiment Results and Discussions

For testing the effect of FCM algorithm to complex background image segmentation of forensic science. The experiment of FCM algorithm was analyzed using stamp impression image with a complex background.

The FCM algorithm was implemented under the software MATLAB 2016a environment and experiment sample, a stamp impression image with a complex background, obtained from the case of forging seals was taken as a test image. Change the default parameter value, and set some parameters to a specific value. The weighted index ρ ($\rho > 1$) that controls fuzziness of the resulting partition was fixed as two. The

parameter $\epsilon (\epsilon > 0)$ denoting a threshold of stopping the iteration was fixed as $+1e-5$. Maximum permission iterative number is fixed as 100.



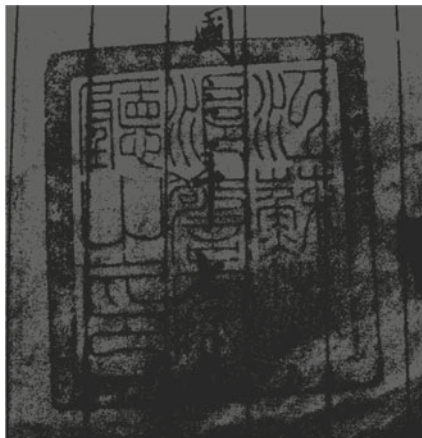
A



A1



A2



A3

In order to test the reliability of FCM algorithm, the number of clusters was varied. In all experiments, we set the different number of clusters $C = 2$, $C = 3$, and $C = 4$. The results are shown as follows.

In the results, picture (A) portrays the original stamp impression image. The pictures (A1)–(A3) depict the different results of stamp impression image segmentation with different number of clusters: for picture (A1), the number of clusters were fixed as 2; for picture (A2), the number of clusters were fixed as 3; for picture (A1), the number of clusters were fixed as 4.

As shown in the resulting graph, the setting of different cluster quantities generates various results of image segmentation are different. When the number of clusters is two, the stamp impression presents poor edge sharpness, so that it cannot be distinguished from the complex background. When the number of clusters is three, image segmentation reaches an ideal result, the outline of stamp impression is relatively clear and it can be properly distinguished from the complex background, which meets the requirements for image segmentation of stamp impression. When the number of clusters is four, the stamp impression can be distinguished from the complex background, but the segmentation of details is undesirable.

Therefore, with respect to the stamp impression under complex background in forensic science, FCM algorithm helps obtain a satisfactory effect in segmenting forensic image and meet relevant requirements.

4 Conclusion

The results of the experiment show that Fuzzy C-Means algorithm has better pixel clustering for the stamp impressions under a complex background with prominent effect of image segmentation. Experiments have proved the feasibility of this method in image segmentation of forensic science. In addition, it can achieve the ideal effect for most of the stamp impressions under a complex background, which gives convenience for the inspection and identification of stamp impressions in forensic science.

Acknowledgements. This work was supported by the National Key Research and Development Plan (2017YFC0822004) and by the Key Project of Basic Scientific Research Service Fee of People's Public Security University of China (2018JKF220). The authors would like to thank the editorial team and reviewers for supporting this paper.

Heartfelt thanks are also given for the comments and contributions of reviewers and members of the editorial team.

References

1. Pal, N.R.: A review on image segmentation techniques. *Pattern Recognit.* **26**(9), 1277–1294 (1993)
2. Shantaiya, S., Verma, K., Mehta, K.K.: Multiple object clustering using fcm and k-means algorithms. *Int. J. Comput. Vis. Robot.* **6**(4), 331 (2016)
3. APAKang, J.Y., Gong, C.L., Zhang, W.J.: Fingerprint image segmentation using modified fuzzy c-means algorithm (report). *J. Biomed. Sci. Eng.* **2**(8), 656–660 (2009)

4. Shi, S.P., Yang, X.U., Qian, H.G., Che, X.U.: Issues related to examination of stamp impressions: identification of stamp impressions forged by high simulation techniques. *Chin. J. Forensic Sci.* (2008)
5. Zhen, D., Zhongshan, H., Jingyu, Y., Zhenmin, T., Yongge, W.: A quick fcm algorithm for gray images segmentation. *Pattern Recognit. Artif. Intel.* (1997)
6. Pal, N.R., Bezdek, J.C.: On cluster validity for the fuzzy c-means model. *IEEE Trans. Fuzzy Syst.* **3**(3), 370–379 (2002)
7. Yu, J., Yang, M.: A study on a generalized fcm. *Lect. Notes Comput. Sci.* **2639**, 390–393 (2003)
8. Boss, R.S.C., Thangavel, K., Daniel, D.A.P.: Mammogram image segmentation using fuzzy clustering. In: *International Conference on Pattern Recognition, Informatics and Medical Engineering*, vol. 02, pp. 290–295. IEEE (2012)



Steady-State Performance Analysis of Quaternion-Valued Least Mean Square Adaptive Algorithm

Sen Li¹(✉), Fengzhi Liu¹, Bin Lin¹, Rongxi He¹, and Xiaomei Zhu²

¹ Department of Information Science and Technology, Dalian Maritime University, Dalian, China

{listen, fzliu, binlin, hrx}@dlmu.edu.cn

² College of Computer Science and Technology, Nanjing Tech University, Nanjing, China
njiczxm@njtech.edu.cn

Abstract. The quaternion-valued least mean squares (QvLMS) adaptive algorithm has been proved valid for the adaptive filtering in quaternion domain. However, there have been few researches on its performance. This paper firstly deduces the energy conservation relation in quaternion domain and then analyzes the steady-state performance of QvLMS algorithm in stationary and non-stationary environment by using the quaternion energy conservation relation. The relevant expressions are deduced and then the step-size range which can guarantee the algorithm convergence is obtained. Simulation results demonstrate the rationality of the analysis.

Keywords: Steady-state performance · Quaternion adaptive filter · Quaternion energy conservation relation

1 Introduction

The quaternion-valued least mean squares (QvLMS) adaptive filtering algorithm has been successfully employed in the statistical signal processing [1]. It is an extension of the LMS algorithm to quaternion-valued signals. Although QvLMS algorithm has been shown desirable properties for different applications, such as wind profile prediction, channel equalization and beamforming [2–4], it is still lack of research on its performance, especially the steady-state performance. In this paper, we firstly extend the energy conservation relation proposed by Sayed [5] to quaternion domain, and then analyze the steady-state performance of QvLMS algorithm in stationary and non-stationary environment by using the quaternion energy conservation relation. The relevant expressions are deduced and then the step-size range which can guarantee the algorithm convergence is obtained. Simulation results demonstrate the rationality of the analysis.

2 Quaternion Energy Conservation Relation

The quaternion is an extension of the real and complex. It includes a real part $\Re[\cdot]$ and an imaginary part $\Im[\cdot]$ which contains three imaginary components, that a quaternion variable $q \in \mathbb{H}$ (where \mathbb{H} denotes quaternion domain) can be expressed as

$$\mathbf{q} = \Re[\mathbf{q}] + \Im[\mathbf{q}] \quad (1)$$

where $\Re[\mathbf{q}] = q_r$ is the real part and $\Im[\mathbf{q}] = q_i I + q_j J + q_k K$ is the imaginary part including three imaginary elements q_i, q_j, q_k , K, J and I are imaginary units that follow the rules of operation $IJ = -JI = K$, $JK = -KJ = I$, $KI = -IK = J$ and $I^2 = J^2 = K^2 = -1$. Multiplication of quaternion is noncommutative, that is for quaternions \mathbf{q}_1 and \mathbf{q}_2 : $\mathbf{q}_1 \mathbf{q}_2 \neq \mathbf{q}_2 \mathbf{q}_1$. The quaternion conjugate is defined as $\mathbf{q}^* = \Re[\mathbf{q}] - \Im[\mathbf{q}]$ and we have $(\mathbf{q}_1 \mathbf{q}_2)^* = \mathbf{q}_2^* \mathbf{q}_1^*$. $|\mathbf{q}| = \sqrt{\mathbf{q} \mathbf{q}^*} = \sqrt{q_r^2 + q_i^2 + q_j^2 + q_k^2}$ is defined as the quaternion modulus.

The update equation of the QvLMS adaptive algorithm is defined as [6]:

$$\mathbf{W}_{i+1} = \mathbf{W}_i + \mu \mathbf{e}(i) \mathbf{S}_i^* \quad (2)$$

where μ is the step size, $\mathbf{W}_i \in \mathbb{H}^L$ is the quaternion weights vector of length L at time instant i , $\mathbf{S}_i \in \mathbb{H}^L$ is the input quaternion signal vector, $\mathbf{e}(i) = \mathbf{d}(i) - \mathbf{W}_i^T \mathbf{S}_i$ is the quaternion error where $\mathbf{d}(i)$ is the reference quaternion signal which can be defined as $\mathbf{d}(i) = (\mathbf{W}^{opt})^T \mathbf{S}_i + \mathbf{n}(i)$, \mathbf{W}^{opt} is the optimal quaternion weights and $\mathbf{n}(i)$ is the quaternion noise.

Define the priori and posterior quaternion estimation error as $\mathbf{e}_a(i) = \tilde{\mathbf{W}}_i^T \mathbf{S}_i$ and $\mathbf{e}_p(i) = \tilde{\mathbf{W}}_{i+1}^T \mathbf{S}_i$, respectively, where $\tilde{\mathbf{W}}_i = \mathbf{W}^{opt} - \mathbf{W}_i$ denotes the quaternion weight error vector and the three errors $\mathbf{e}_a(i)$, $\mathbf{e}_p(i)$ and $\mathbf{e}(i)$ are related via

$$\mathbf{e}(i) = \mathbf{e}_a(i) + \mathbf{n}(i) \quad (3)$$

$$\mathbf{e}(i) = (\mathbf{e}_a(i) - \mathbf{e}_p(i)) / (\mu \|\mathbf{S}_i\|^2) \quad (4)$$

Use \mathbf{W}^{opt} to subtract the Eq. (2) from both sides and then substitute Eq. (4) into it, we can get

$$\tilde{\mathbf{W}}_{i+1}^T + \frac{\mathbf{e}_a(i) \mathbf{S}_i^*}{\|\mathbf{S}_i\|^2} = \tilde{\mathbf{W}}_i^T + \frac{\mathbf{e}_p(i) \mathbf{S}_i^*}{\|\mathbf{S}_i\|^2} \quad (5)$$

Evaluating energies of both sides of (5) and applying the statistical expectation operator on the equation, the quaternion energy conservation relation can be represented

$$E \left[\left\| \tilde{\mathbf{W}}_{i+1}^T \right\|^2 + \frac{|\mathbf{e}_a(i)|^2}{\|\mathbf{S}_i\|^2} \right] = E \left[\left\| \tilde{\mathbf{W}}_i^T \right\|^2 + \frac{|\mathbf{e}_p(i)|^2}{\|\mathbf{S}_i\|^2} \right] \quad (6)$$

The Eq. (6) shows the relationship between the quaternion priori and posterior estimation errors energies and the two successive time instants quaternion weight error vectors energies. It can be seen that (6) is the quaternion domain direct extension of the energy conservation relation for complex signals. But this extension only holds for the update equation defined in (2) and does not hold for the other version of the QvLMS algorithm which were discussed in [6] due to the noncommutative characteristic of the quaternion multiplication.

3 Steady-State Performance Analysis

The criterion that is widely used to analyze the steady-state performance of the adaptive filtering algorithm is excess mean square error (EMSE) which is defined as $\xi_{EMSE} = \lim_{i \rightarrow \infty} E \left[\|\mathbf{e}_a(i)\|^2 \right]$. To analyze the EMSE, we assume that the noise $\mathbf{n}(i)$ is an i.i.d. zero-mean noise with variance σ_n^2 and is independent of the input signal \mathbf{S}_i and the priori estimation error $\mathbf{e}_a(i)$. Substitute (4) into (6) and take the limitation of both sides, as $\lim_{i \rightarrow \infty} E \left[\left\| \tilde{\mathbf{W}}_{i+1}^T \right\|^2 \right] = \lim_{i \rightarrow \infty} E \left[\left\| \tilde{\mathbf{W}}_i^T \right\|^2 \right]$ when in steady-state, we can get the following formula

$$2 \lim_{i \rightarrow \infty} \left[E|\mathbf{e}_a(i)|^2 \right] = \mu \lim_{i \rightarrow \infty} \left[E[\|\mathbf{S}_i\|^2 |\mathbf{e}_a(i)|^2] \right] + \mu \sigma_n^2 E \left[\|\mathbf{S}_i\|^2 \right] \quad (7)$$

Define $\text{Tr}(\mathbf{R}_s) = E(\|\mathbf{S}_i\|^2)$ and further assume $\|\mathbf{S}_i\|^2$ is statistically independent of $|\mathbf{e}_a(i)|^2$. We can obtain

$$\xi_{EMSE} = \frac{\mu \sigma_n^2 \text{Tr}(\mathbf{R}_s)}{2 - \mu \text{Tr}(\mathbf{R}_s)} \quad (8)$$

Under a non-stationary environment, the variation of the optimal weights follows a random-walk mode $\mathbf{W}_{i+1}^{opt} = \mathbf{W}_i^{opt} + \mathbf{V}_i$, where \mathbf{V}_i is an i.i.d. quaternion vector with quaternion positive-definite autocorrelation matrix $\mathbf{Q} = E(\mathbf{V}_i \mathbf{V}_i^H)$ and is independent of \mathbf{S}_i and $\mathbf{n}(i)$. The posterior estimation error can be redefined as $\mathbf{e}_p(i) = (\tilde{\mathbf{W}}_{i+1}^T - \mathbf{V}_i) \mathbf{S}_i$. Using a method similar to the above, we can get the EMSE of the QvLMS algorithm in a non-stationary environment is

$$\xi_{EMSE} = \frac{\mu\sigma_n^2\text{Tr}(\mathbf{R}_s) + \mu^{-1}\text{Tr}(\mathbf{Q})}{2 - \mu\text{Tr}(\mathbf{R}_s)} \quad (9)$$

Comparing the Eq. (9) which is the ξ_{EMSE} in a non-stationary environment and the Eq. (8) which is the ξ_{EMSE} in stationary environment, we can see that there is an additional term $\mu^{-1}\text{Tr}(\mathbf{Q})$ which reflecting the influence of the non-stationary character of the weight vector to the filter steady-state performance. As the definition of the EMSE indicated, the denominator of (8) and (9) must be positive to guarantee EMSE positive, therefore, we can get a approximately estimate range of the value of step-size μ for QvLMS algorithm as

$$0 < \mu < \mu_{\max} = \frac{2}{\text{Tr}(\mathbf{R}_s)} \quad (10)$$

It also can be observed that there exists an optimum step size to minimize the ξ_{EMSE} in a non-stationary environment which can be deduced as

$$\mu^{opt} = \sqrt{\frac{\text{Tr}(\mathbf{Q})}{\sigma_n^2\text{Tr}(\mathbf{R}_s)} + \frac{(\text{Tr}(\mathbf{Q}))^2}{4\sigma_n^4}} - \frac{\text{Tr}(\mathbf{Q})}{2\sigma_n^2} \quad (11)$$

4 Simulations

In this work, the length of the QvLMS adaptive filter is 15 and the optimal quaternion weights \mathbf{W}^{opt} is a random quaternion weight vector. The four elements q_r, q_i, q_j, q_k of the input signal and noise is zero-mean Gaussian sequence with variance 1 and 0.0025, respectively. So the covariance matrix of the signal is $\mathbf{R}_s = 4\mathbf{I}$ where \mathbf{I} defined as the identity matrix and the variance of the noise is $\sigma_n^2 = 0.01$. Figure 1 shows the steady-state EMSE against the step-size μ in the stationary environment. The theoretical EMSEs come from the expression of (8). The experimental EMSEs were obtained from the average of 200 independent simulation runs which the individual EMSE is the ensemble-average of 4000 iterations after the QvLMS algorithm convergences. It can be seen that the EMSE increases with the step size μ and the experimental results match the theoretical results reasonably well.

Set $\mathbf{Q} = \sigma_v^2\mathbf{I}$ and $\sigma_v^2 = 10^{-8}$, Fig. 2 demonstrates the experimental and theoretical EMSEs under non-stationary environment which are very consistent. In this case, there is no proportional relationship between EMSE and the step size μ , but there is an optimal step size μ^{opt} that makes EMSE minimum which is very close to the theoretical value 4.9256×10^{-4} that is calculated by the formula (11).

To confirm the step-size scope of (10), we define the probability of divergence as [7]

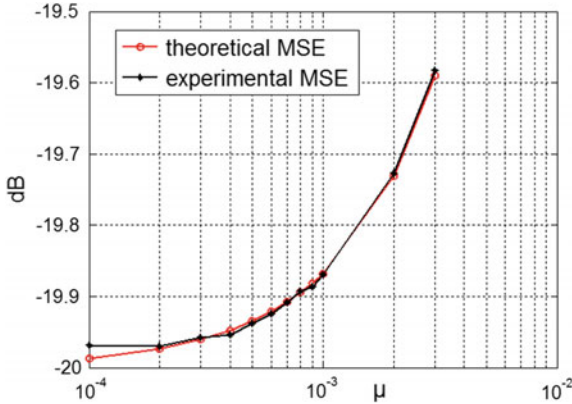


Fig. 1. Theoretical and experimental EMSE in stationary environment.

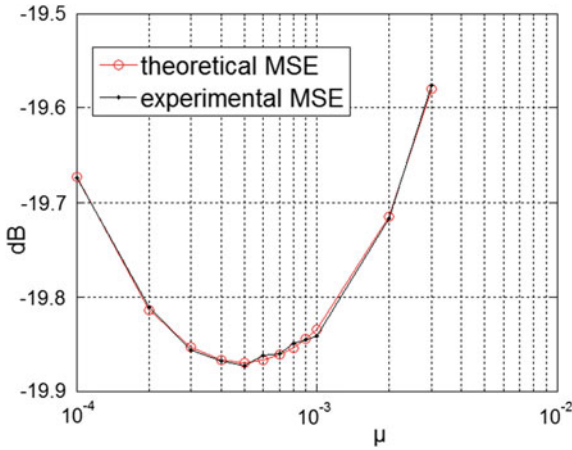


Fig. 2. Theoretical and experimental EMSE in non-stationary environment.

$$P_d = \frac{N_{di}}{N_{exp}} \tag{12}$$

where N_{exp} is the number of Monte Carlo simulation experiments which are starting from the same initial condition and N_{di} is the number of curves diverging in the N_{exp} simulation experiments. The curve is considered as diverged if $\|W_{N_{it}}\| \geq 10^3$ after N_{it} iterations. Figure 3 shows the estimated P_d changing with the step-size $\mu = \delta\mu_{max}$ with $\delta \in [0.1, 1.2]$ in a stationary environment, and $N_{exp} = 100, N_{it} = 1000$. We can find that, to ensure the convergence of QvLMS algorithm, the value of the step-size μ is supposed to be chosen less than $0.9\mu_{max}$.

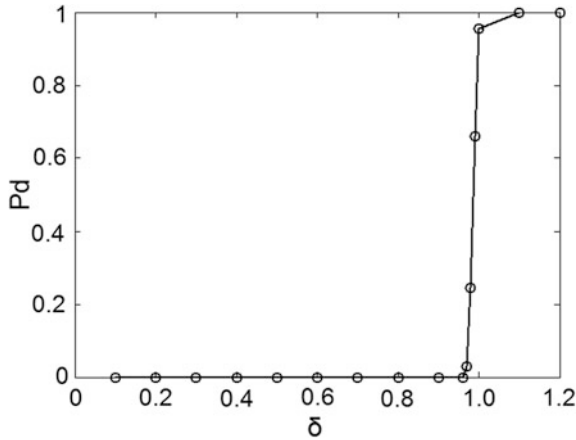


Fig. 3. Probability of divergence against step-size $\mu = \delta\mu_{\max}$.

5 Conclusion

In this paper, an effective method to analyze the steady-state performance of QvLMS adaptive algorithm had been introduced in both stationary and non-stationary environments. The quaternion energy conservation relation was derived and used to deduce the relevant EMSE expressions which had the similar forms as the complex LMS adaptive algorithm. In the meanwhile, the rough interval of step-sizes to ensure the convergence of QvLMS algorithm was obtained. Finally, the simulation results have been provided to support the analysis.

Acknowledgements. This work was supported in part by the National Natural Science Foundation of China under Grants 61301228, 61501223 and the Fundamental Research Funds for the Central Universities under Grant 3132016331 and 3132016318.

References

1. Took, C.C., Mandic, D.P.: The quaternion LMS algorithm for adaptive filtering of hypercomplex processes. *IEEE Trans. Signal Process.* **57**(4), 1316–1327 (2009)
2. Jiang, M.D., Liu, W., Li, Y., Zhang, X.R.: A general quaternion-valued gradient operator and its applications to computational fluid dynamics and adaptive beamforming. In: 19th International Conference on Digital Signal Processing, Hong Kong, China, vol. 8, pp. 821–826 (2014)
3. Jiang, M.D.: Quaternion-valued adaptive signal processing and its applications to adaptive beamforming and wind profile prediction. Ph.D. thesis, University of Sheffield (2017)
4. Liu, W.: Channel equalization and beamforming for quaternion-valued wireless communication systems. *J. Frankl. Inst.* **354**(18), 8721–8733 (2017)

5. Sayed, A.H.: Fundamentals of adaptive filtering. Wiley, New York (2003)
6. Barthélemy, Q., Larue, A., Mars, J.I.: About QLMS derivations. *IEEE Signal Process. Lett.* **21**(2), 240–243 (2014)
7. Li, S., Song, L.M., Qiu, TSh: Steady-state and tracking analysis of fractional lower-order constant modulus algorithm. *Circuits Syst. Signal Process.* **30**, 1275–1288 (2011)



Research on Concept-Drifting Data Stream Based on Fuzzy Integral Ensemble Classifier System

Baoju Zhang^(✉), Yidi Chen, and Lei Xue

Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
wdxzyzbj@163.com

Abstract. With the arrival of the era of big data, a large amount of data stream generates in the real world. However, the existence of concept drift has brought great challenges to data stream classification. Therefore, this paper proposed an ensemble classifier system based on fuzzy integral to solve the above problem. And after the experimental evaluation, we can approve the proposed algorithm outperforms other algorithms in terms of classification performance and the ability to adapt to new concepts efficiently.

Keywords: Data stream · Concept drift · Fuzzy integral · Ensemble classification

1 Introduction

As a new data form, data stream is different from static data, which has dynamic variability. As time goes on, the context content it contains also changes. And, this feature is defined as concept drift. This characteristic not only requires that learning algorithms on the data stream should have a high classification accuracy, but also the ability to processing and decision-making adaptively to dynamic environment. Therefore, in recent years, lots of effective learning algorithms have been proposed for detecting the concept drift in data streams.

From the perspective of the error variance between concepts, in [1], it predefined a threshold based on Bernoulli data distribution to detect concept drift more clearly, and this method was named Drift Detection Method (DDM). Reference [2] used entropy to judge whether drift occurs, and this method calculated the used entropy value to evaluate the difference of the data distribution between training sets. Reference [3] combined information degree and detection methods, and established a classifier model. On the other hand, the Accuracy-Weighted Ensemble in [4] is a general framework for mining concept drift using weighted ensemble classifiers. Reference [5] integrated the idea of ensemble classification into the flow field creatively, and proposed the SEA algorithm. Reference [6] proposed an online ensemble classifier, which can shorten training time of the existing ensemble classifiers. These methods have good effects to adapt to concept drift, however, they need a long time period in the dynamic process.

From the research above, the idea of integrating multiple base classifiers has been proved to be effective in processing the drift of concepts. The key goal is to discover the occurrence position and time point timely, and to minimize the fluctuation of classification accuracy. Considering this, the paper focuses on analyzing the concept drift's influence on classifier, exploring the concept change based on degree of drift, and combining this approach with fuzzy integral ensemble learning algorithm, so as to build an ensemble classifier which is able to adapt to the concept drift.

The other parts of the article are conceived as below: the second part mainly puts forward a quantitative index to measure whether concept drift occurs, and expounds the basic theory of ensemble classifier constructed using fuzzy integral. In the third part, we summary the implementation process of the new model presented in this paper, and demonstration of experimental results employing real and simulated datasets are shown in the next part. The last part summarizes the full text and looks forward to further research directions.

2 Theory

2.1 Detection Method Based on Degree of Drift

To detect concept drift, in this paper, we proposed a judgment indicator of degree of drift and introduce the Euclidean distance to achieve the goal.

For each of the current data block, we calculate the Euclidean distances of all instances in previous chunk, and compare their values, and the instance with minimum value is been regarded as its nearest neighbor.

$$d_E(x_1, x_2) = \sqrt{\sum_{a=1}^m d_a^2 \left(\frac{x_{1a} - x_{2a}}{\text{range}_a} \right)} \quad (1)$$

where x_1 is the instance in current chunk, and x_2 is the nearest neighbor of all instances in previous data chunk, m represents the number of instance attributes. For the numerical attribute a , range_a is the largest difference of the corresponding attribute, it can be normalized by the width of the value.

Then, we compare the labels of x_i and its nearest neighbor, if they have the same label, $\text{dis}(i) = 0$, otherwise, $\text{dis}(i) = 1$.

$$\text{dis}(i) = \begin{cases} 1 & \text{if } x_i \text{ and its nearest neighbor haven different label} \\ 0 & \text{if } x_i \text{ and its nearest neighbor have the same label} \end{cases} \quad (2)$$

From this, D_E is proposed to describe the degree of concept drift, and it can be given as

$$D_E = \frac{\sum_{i=1}^n d_E(i) * \text{dis}(i)}{\sum_{i=1}^n d_E(i)} \quad (3)$$

where $d_E(i)$ means the minimum value, mentioned in Eq. (1), $dis(i)$ is drawn from Eq. (2), n is the size of data chunk. We can easily understand that the significant increase of D_E indicates that the high likelihood of the occurrence of the concept drift.

2.2 An Ensemble Algorithm Based on Fuzzy Integral

2.2.1 Fuzzy Measure and Choquet Integral

The fuzzy measurement and integration theory, as a fuzzy measure that describes the degree of importance of a generalized framework, it can better integrate basic information and explore the relationship between fuzzy measurement and fuzzy integration. It has great practical significance in classification technology, comprehensive evaluation, and other issues.

In this paper, we choose g_λ fuzzy measure, one kind popular method, as the major learning tool. We call a fuzzy measure μ on $F(X)$ a g_λ fuzzy measure, if there exists $\lambda \in [-1, \infty)$ such that make Eq. (4) holds:

$$\mu(X_1 \cap X_2) = \mu(X_1) + \mu(X_2) + \lambda\mu(X_1)\mu(X_2) \tag{4}$$

where, $X_1, X_2 \in F(X)$. The parameter λ is determined by

$$\lambda + 1 = \prod_{i=1}^n (1 + \lambda\mu_i) \tag{5}$$

Assuming X is a finite set, we record fuzzy measure in the subset as $\mu = \mu(\{X_i\})$. In multiple classifiers learning system, μ_i indicates the degree of each base classifier's function in classification. The value of μ_i can be given as

$$\mu_i = \frac{P_i}{\sum_{j=1}^L P_j} \text{dsum} \tag{6}$$

Here, dsum is a constant parameter, L represents the number of classifiers. Hence, using Eq. (5) to calculate the value of λ , and the g_λ measure is determined.

In a variety of fuzzy integral methods, this paper selects excellent Choquet fuzzy integral. For X , μ is regarded as a fuzzy measure of $h(X)$, then resort the set element by their values, denoted as $x_1^*, x_2^*, \dots, x_n^*$, to maintain $h(x_1^*) \leq h(x_2^*) \leq \dots \leq h(x_n^*)$. Consequently, Choquet fuzzy integral can be simplified using the Eq. (7):

$$(C) \int h d\mu = \sum_{i=1}^n [h(x_i^*) - h(x_{i-1}^*)] \mu(A_i) \tag{7}$$

where $A_i = \{x_1^*, \dots, x_n^*\}$, specify $h(x_0) = 0$. In this definition, the increase of function h or μ will cause the increase of integral.

2.2.2 The Ensemble System Based on Choquet Integral

As for the proposed ensemble classifier model, the g_λ measure and base classifier output is taken as the input of the Choquet integral. The construction process of the specific model is as follows.

A total of L classifiers are used hypothetically, and is denoted as E_1, E_2, \dots, E_L . By learning instance X , we get an m -dimensional vector from each classifier's output. In this paper, the matrix can be used to describe classification results of all classifiers on current instance X , this matrix is called the decision profile:

$$DP(X) = \begin{bmatrix} d_{1,1} & d_{1,2} & \dots & d_{1,k} & \dots & d_{1,m} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ d_{j,1} & d_{j,2} & \dots & d_{j,k} & \dots & d_{j,m} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ d_{L,1} & d_{L,2} & \dots & d_{L,k} & \dots & d_{L,m} \end{bmatrix} \quad (8)$$

where the j th line indicates the output vector of the j th classifier and the k th column indicates the possibility of current instance X belonging to the k th class according to each classifier.

We use the k th column of $DP(X)$ as the function $h(x)$ of Choquet integral, then calculate its value, and send the corresponding class to output $C_X = \arg(\max_{1 \leq i \leq L} \{E_i\})$ according to the Maximum principle.

The integration process of classifiers based on Choquet integral is shown in Fig. 1.

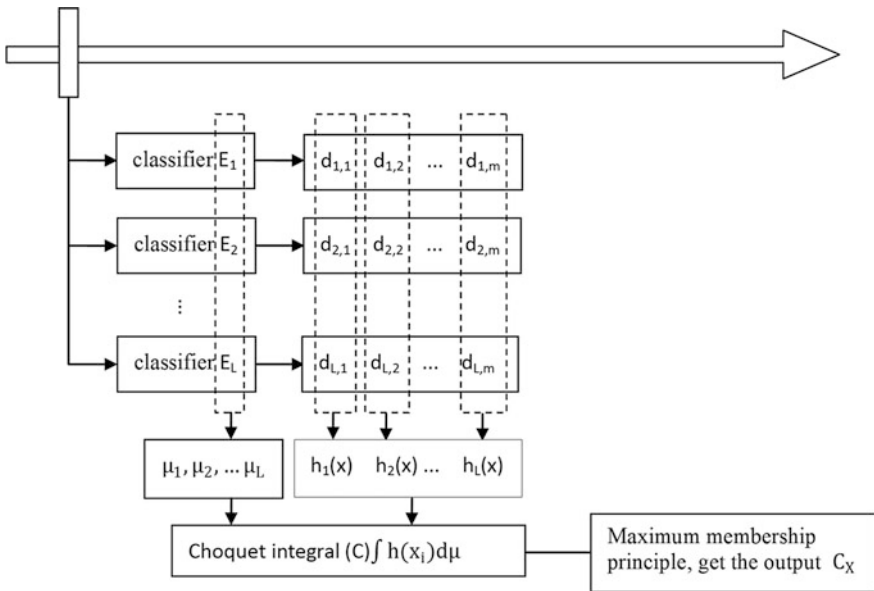


Fig. 1. The framework of ensemble system based on Choquet integral

3 Fuzzy Integral Ensemble System Adapted to Concept Drift

Since testing the change of concepts in time, and maximizing the prediction accuracy are the main goal in data stream classification algorithm, the method in our paper proposes an ensemble classification system based on Choquet integral.

The construction of the ensemble algorithm for concept drift is as follows (Fig. 2).

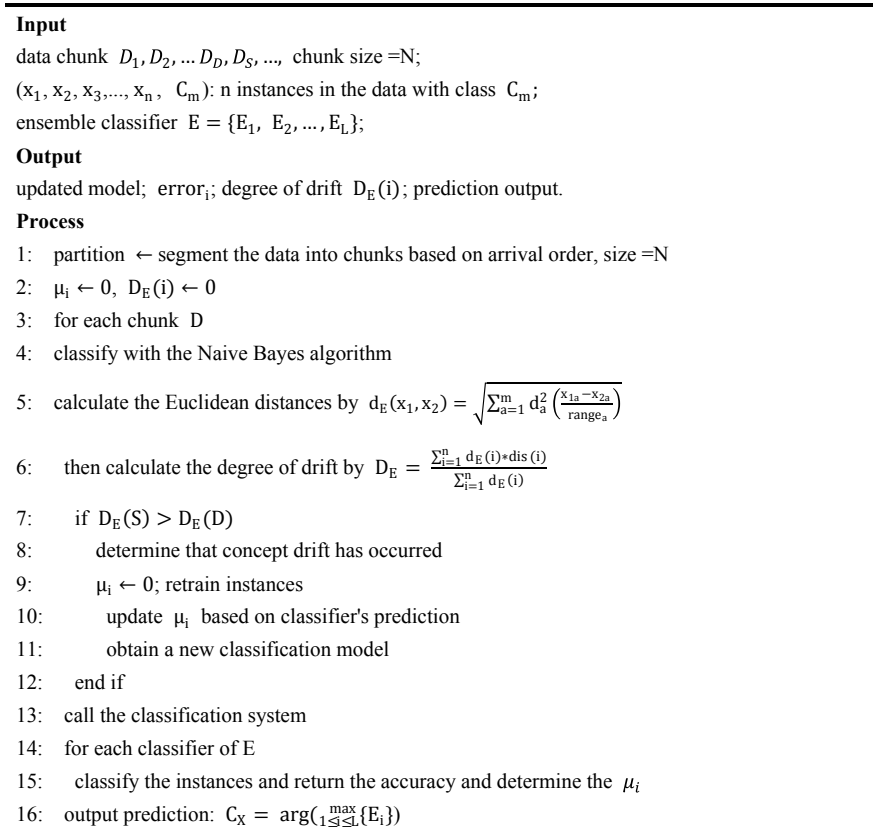


Fig. 2. The construction of the ensemble algorithm for concept drift

4 Experimental Analysis

4.1 Data Sets

- (1) UCI datasets (Table 1).
- (2) SEA dataset: The structure of the SEA dataset is (f_1, f_2, \dots, C) , where f_1, f_2 is the attributes related to classification, C is a category attribute containing two category labels. When $f_1 + f_2 \leq \theta$, classifying the instance in category “1”, else

Table 1. Dataset description of UCI library

Dataset	Instance	Attribute	Class
Iris	150	4	3
Hypothyroid	3772	29	4

category “2”. To simulate the drift of concepts, we set threshold θ to 9.5, 9, 8, 7, respectively. Randomly generating 10 k instances, contains four concepts and three drifts, the noise rate is 10%.

- (3) Electricity grid dataset: To verify the wide availability of the algorithm, we apply the real dataset about electricity in the experiment, and the dataset is downloaded from the Official website of Swedish power grids. After our arrangement, it includes 2190 instances and covers the electricity consumption of 3 years in Swedish power grids, and it has 12 attributes and records every hour of days. Those recorders involve two classes, day time and night time.

4.2 Ensemble Classifier Algorithms on UCI Datasets

Table 2 shows the results of different ensemble classification algorithms using the three datasets in the UCI library. In terms of the overall classification accuracy, the generation model has better performance, and ensures its effectiveness.

Table 2. Accuracy of fuzzy integral ensemble classifier and other ensemble classifiers

Ensemble classifier	Iris (%)	Hypothyroid (%)
Vote	96.67	94.41
Bagging	94	95.28
Proposed	96.67	94.54

4.3 Analysis of Classification Accuracy

In this part, three other algorithms, the Naïve Bayes algorithm, the Accuracy-Weighted Ensemble (AWE) algorithm and the Ozabag, are the comparative methods to assess the capability of new generation model.

Figure 3 shows the test result on SEA dataset. For all algorithms, the accuracy will decrease significantly as the occurrence of concept drift. Secondly, at the early training phase, comparing with the stable period, the accuracy rate will at a lower level.

For the electricity grid dataset, in Fig. 4, there is a clear sign of improvement on accuracy of the proposed algorithm. But the curve of proposed algorithm is not smooth enough, that’s because the partition of dataset into chunks lead to a smaller size of data chunk, therefore cause the fluctuation of accuracy even though there is just one instance which can make a difference.

Table 3 records the accuracy of algorithms on different datasets. In terms of the accuracy value, we can see that the AWE algorithm is superior to processing the data

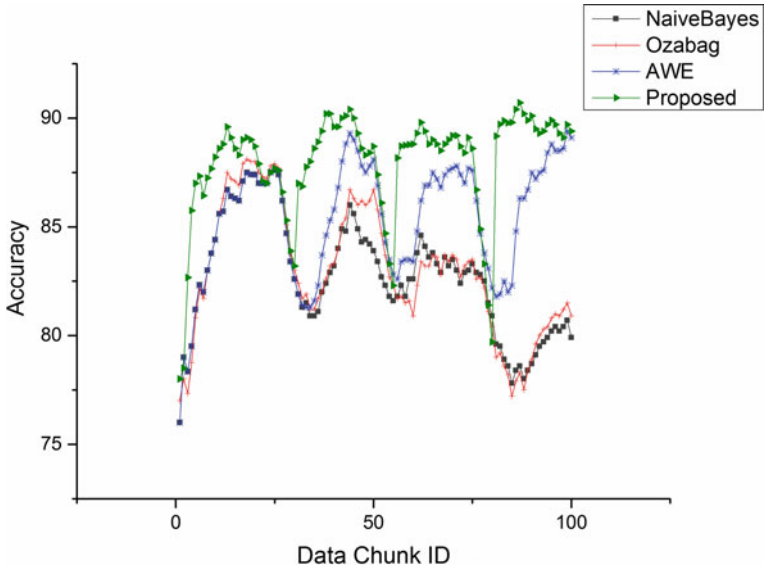


Fig. 3. The predictive accuracy on SEA dataset

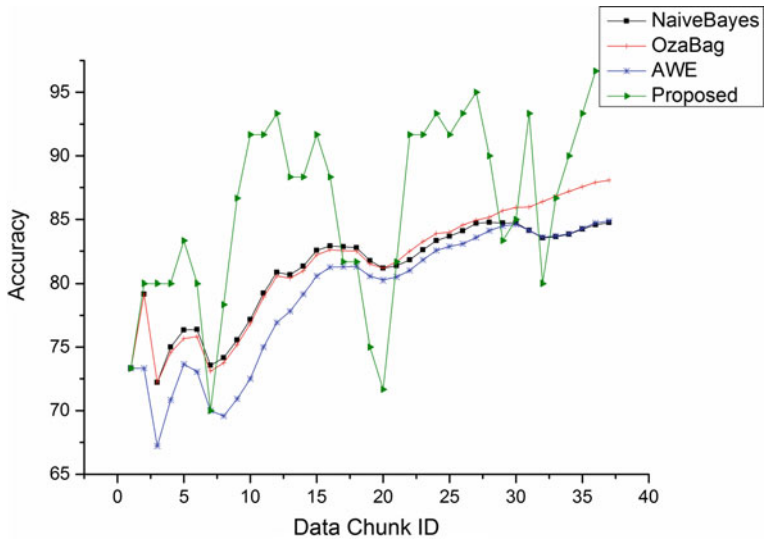


Fig. 4. The predictive accuracy on Swedish electricity grid dataset

streams with noise. For the real dataset about electricity, the advantages of the use of the proposed algorithm in the accuracy performance are more obvious. On the whole, the proposed algorithm achieves the desired effect in performance.

Table 3. The comparison of classification accuracy on experimental datasets

Dataset	Naïve Bayes (%)	AWE (%)	Ozabag (%)	Proposed (%)
SEA	82.72	85.39	82.98	87.94
Electricity grid	80.90	79.09	81.48	85.88

5 Conclusion

This paper proposed a judgment indicator of degree of drift to improve reliability of the detection method, meanwhile applying a new kind of algorithm based on fuzzy integral adaptively updating the model to get the corresponding classification result. According to the experimental results, results have verified the proposed algorithm can effectively improve the classification accuracy while also exerting a good demonstration effect on the solution of the concept drift problem. The change can be detected in time, and quickly adapt to the new environment and maintain a high classification accuracy. In addition, when the occurrence of concept drift has been identified, there exists a short period that the classification accuracy is at a low value, it's still challenging and necessary to improve in our future work.

Acknowledgements.. This paper is supported by the Natural Science Foundation of China (61271411), Natural Youth Science Foundation of China (61501326). It is also supported by Tianjin Research Program of Application Foundation and Advanced Technology (15JCZDJC31500) and Tianjin Science Foundation (16JCYBJC16500).

References

1. Gama, J., Medas, P., Castillo, G., Rodrigues, P.: Learning with drift detection. In: Brazilian Symposium on Artificial Intelligence, pp. 286–295. Springer Berlin Heidelberg (2004)
2. Peter, V., Abraham, B.: Entropy-based concept shift detection. In: International Conference on Data Mining, pp. 1113–1118. IEEE (2006)
3. Nishida, K., Shimada, S., Ishikawa, S., Yamauchi, K.: Detecting sudden concept drift with knowledge of human behavior. In: IEEE International Conference on Systems, Man and Cybernetics, pp. 3261–3267. IEEE (2008)
4. Wang, H., Yu P.S., Jiawei, H.: Mining concept-drifting data streams using ensemble classifiers. In: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 226–235. ACM (2003)
5. Street, W.N., Kim, Y.S.: A streaming ensemble algorithm (SEA) for large-scale classification. In: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 377–382. ACM (2001)
6. Chu, F., Wang, Y., Carlo, Z.: An adaptive learning approach for noisy data streams. In: IEEE International Conference on Data Mining, pp. 351–354. IEEE Computer Society (2004)



An Improved Speech Synthesis Algorithm with Post filter Parameters Based on Deep Neural Network

Shunjie Dong¹, Chunyang Li², and Hong Zhang¹(✉)

¹ School of Microelectronics, Xidian University, Xi'an, China
{15029017541, smezhang}@163.com

² School of Computer Science and Technology, Xidian University, Xi'an, China
lichunyang_1@outlook.com

Abstract. Statistical parameters speech synthesis typically relies on context-dependent Hidden Markov Model (HMM) that is based on decision tree clustering. However, the shortcomings of clustering decision tree, restricted to a feature rigid subdivision model space, results in smooth speech parameters generated from HMM. In this paper, Deep Neural Network (DNN) is put forward to replace clustering decision tree, and we propose a post filter-parameter-based speech synthesis improvement algorithm. This method enhances the formant region of synthesized speech spectrum by selecting the most optimized filter parameter according to the flatness of spectrum. The experimental results show that DNN effectively can modify the deficiency of two smooth parameters. Furthermore, the improved post filter algorithm increases the naturalness of synthesized speech.

Keywords: HTS · DNN · Post filter · HMM · Naturalness

1 Introduction

In recent years, Deep Neural Network (DNN) which won an unprecedented success in the field of speech recognition has significantly improved speech recognition [1]. DNN-HMM has become one of the most advanced acoustic models in the field of speech recognition. Compared with the Gaussian Markov Model (GMM-HMM) based on hybrid Gaussian model, DNN-based speech recognition method achieves long span, high dimension, and strong correlation of feature input. Meanwhile, this method can find the highly nonlinear mapping of input and output characteristics through DNN structure, and train the model parameters by using error backpropagation algorithm and cost function of the gradient [2].

Speech synthesis is a technique that translates text into sound. HMM-based statistical parametric speech synthesis has become one of the most popular technologies for its high fluency and flexibility. However, the HMM synthesized speech also causes problems, such as low naturalness, serious decline in sound quality, etc. Several approaches have been proposed to solve these problems, which have two main directions: one is to improve the statistical parameter model and the other is to optimize vocoding.

Due to the success of DNN in speech recognition, some DNN-based statistical parameter speech synthesis methods have been tried to improve the performance of speech synthesis. Zen et al. listed the limitations of HMM models, e.g., decision tree-based contextual state clustering, and suggested using DNN to overcome these limitations [3]. In the case of a large training corpus, DNN outperforms GMM-HMM with similar parameters in synthesizing higher quality speech. Qian et al. examined various aspects of DNN-HMM model with a medium-sized corpus, which makes the statistical parametric model easier to train [2]. Yoshimura et al. proposed the application of the Mel-cepstrum post filter in linear predictive coding to the method of speech synthesis [4]. Takamichi et al. introduced a post filter based on a modulation spectrum [5]. Ling et al. adopted an LSP-based post filter [6], which enhances the formant by the relationship between LSP and the formant. But due to the fixed post filter parameters in the existing algorithm, it cannot match voice signals with different degree of distortion.

This paper further investigates an improved post filter on DNN-based speech synthesis. Formant is first enhanced by the post filter with different parameters, which are based on the different distortion degree of speech synthesized by DNN. And then, it is compared with natural voice, and the optimal solution of post filter parameter can be determined by Barker spectrum distortion measure. Additionally, the exponential relationship between optimal post filter parameter and speech distortion is concluded. In the process of synthesizing speech, optimal post filter parameters are calculated in accordance with the distortion of current speech, and the enhancement of formant region is realized ultimately. Verified by the experimental results, this method effectively improves the natural degree of synthetic speech.

2 System Architecture

2.1 Deep Natural Network

DNN is a feedforward artificial neural network, having enough hidden layers between its input and output. For each hidden unit j , a nonlinear activation function $f(x)$ is used to map all inputs from x_j in the next layer to a scalar state y_j which is subsequently delivered to the previous layer:

$$y_j = f(x_j)$$

$$x_j = b_j + \sum_i y_i w_{ij}$$

where b_j represents the bias of unit j , i is the unit of the next layer, w_{ij} is the weight of connection between unit j and unit i in the layer below, while the sigmoid function is selected as the activation function:

$$f(x_j) = \frac{1}{1 + e^{-x_j}}$$

The hyperbolic tangent function is used as the mapping between input and output:

$$f(x_j) = \frac{e^{x_j} - e^{-x_j}}{e^{x_j} + e^{-x_j}}$$

Since a hyperbolic tangent function is a regulated version of the sigmoid function, the two activation functions own the same modeling capability. The output range of a sigmoid function is (0, 1), which contributes to the expression of sparse algebra, but it results in the asymmetry of activation value. On the other hand, the output range of a hyperbolic tangent function is (-1, 1), so it is symmetrical, which will help training [7].

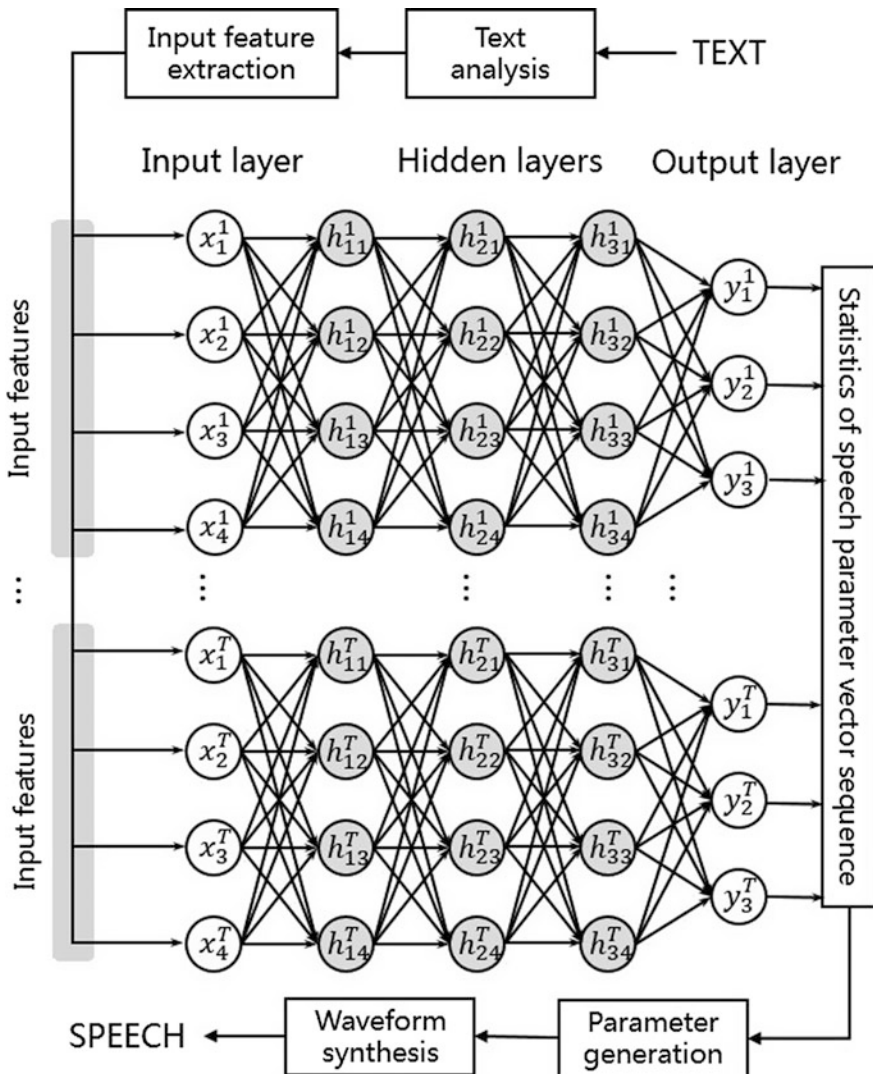


Fig. 1. A speech synthesis framework based on DNN

All weights and biases are initialized before model training [8] and trained with an optimized cost function. The function uses the backpropagation algorithm to measure the discrepancy between predicted output and target vector [9]. Given a training set $\{(x^{(1)}, y^{(1)}), \dots, (x^{(T)}, y^{(T)})\}$, the minimized cost function is defined as

$$C = \frac{1}{2T} \sum_{t=1}^T \|f(x^{(t)}) - y^{(t)}\|^2$$

To prevent over-fitting, the weight decay term is added to the right of the equation [10] and the model is trained by using batch gradient descent algorithm:

$$(W^l, b^l) \leftarrow (W^l, b^l) + \varepsilon \frac{\partial c}{\partial (W^l, b^l)} \quad 0 \leq l < L$$

where ε represents a preset learning rate.

Figure 1 illustrates a speech synthesis framework based on DNN, including both training and synthesis. In the process of synthesis, the input text is fed into a trained DNN after being analyzed to generate acoustic features.

In order to produce relatively smooth parameter trajectories, the speech parameter generation will be constrained by dynamic features, where predicted features of DNN are used as average vectors and the global variances of training data are used as covariance matrix to maximize the probability.

2.2 Filter Parameter

To match the parameters of post filter with the distortion of speech spectrum adaptively, it is necessary to set a parameter to characterize the distortion. Compared with zero-crossing rate, gradient index, spectral dynamics, spectral flatness, is selected to better describe the degree of speech distortion. Spectral flatness is defined as

$$\zeta_n = \frac{\exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln P_n(\omega) d\omega\right)}{\frac{1}{2\pi} \int_{-\pi}^{\pi} P_n(\omega) d\omega}$$

where $P(\omega)$ represents signal power spectral density and n is the number of the frame.

Considering the differences in pronunciation duration and rate, natural voice and synthetic speech cannot be compared with each other directly. Accordingly, this paper adopts DTW to reduce the influence of the differences between pronunciation duration and rate, so that they can be eliminated.

Subjective evaluation is consistent with the actual feeling of voice quality. But it is complex, awkward and vulnerable to subjective feelings. In this paper, Barker Spectral Distortion (BSD) is used as an objective evaluation method to determine the optimal filter parameters. BSD, with good objective evaluation performance, is widely used in speech distortion evaluation. Only with respect to the spectral distortion perceived by the human ear, BSD has a better performance than other spectral coefficients of

distortion measure [7]. The Bark spectral discrepancy between natural voice and synthetic speech is defined as

$$BSD_n = \sum_{i=1}^{15} [\widehat{L}_n(i) - L_n(i)]^2$$

where $\widehat{L}_n(i)$ and $L_n(i)$ are, respectively, the i th loudness value of the n th Bark band of natural speech and synthesized speech, and n is the number of the frame.

Through the comparison between natural voice and synthesized speech which is processed by the post filter with different parameters, BSD discrepancy is calculated and a series of the optimal filter parameters with the smallest BSD discrepancy is obtained as a sequence of the spectral flatness of the speech. The correlation between the optimal filter parameters and the speech spectral flatness can be established by polynomial fitting [11] or exponential relationship [12]. This paper chooses the last one. The relationship is defined as

$$\beta = \alpha_1 + e^{\alpha_2 \zeta + \alpha_3}$$

where coefficient $\alpha_1, \alpha_2, \alpha_3$ can be gained by the LM method [13].

2.3 Post filter

When enhancing the formant of synthesized speech, the spectral flatness of synthesized speech needs to be calculated first, then the optimal filter parameters are obtained by fitting the relationship between filter parameters and spectral flatness, and finally, the synthesized speech is filtered by the improved post filter. The transfer function of post filter is defined as

$$H_S(z) = G_s D_\beta(z)$$

where G_s is the gain correction factor and $D_\beta(z)$ represents the formant enhancement filter.

Emphasizing formants by using a post filter enables the improvement in the naturalness of synthesized speech. In this paper, Mel-cepstral coefficients, considered as the spectral parameters in synthesis, are applied to the post filter. The spectral model with Mel-cepstral coefficients represented by $c(m)$ is expressed as [4]

$$D(z) = \exp \sum_{m=0}^M c(m) \bar{z}^{-m}$$

because

$$b(m) = \begin{cases} c(m) & m = 0 \\ c(m) - \alpha b(m+1) & 0 \leq m < M \end{cases}$$

so

$$D(z) = \exp \sum_{m=0}^M b(m) \phi_m(z)$$

where $\phi_m(z)$ represents as:

$$\phi_m(z) = \begin{cases} 1 & m = 0 \\ \frac{(1-\alpha^2)z^{-1}}{1-\alpha z^{-1}} \tilde{z}^{-(m-1)} & m \geq 1 \end{cases}$$

where

$$\tilde{z}^{-1} = \frac{z^{-1} - \alpha}{1 - \alpha z^{-1}}$$

where $c(m)$ is Mel-cepstral coefficient, M is dimensionality of the Mel-cepstral analysis, \tilde{z}^{-1} is Mel frequency deflection function, and α is a parameter to approximate the Mel scale. In order to avoid emphasizing the spectral overall trend, we set $c(1)$ to 1. And, β represents the filter parameter calculated from spectral flatness, when $\beta = 0$, the post filter is not applied, when $\beta > 0$, it is used to enhance the spectra. The post filter is expressed as

$$D_\beta(z) = \exp \sum_{m=1}^M \beta c(m) \tilde{z}^{-m}$$

The gain correction factor G_s allows that the energy of the output signals is equal to that of the input signals which avoids the changes of the speech signal energy in the filtering process [14]. We set $x(n)$ as the input of the emphasizing formant filter, $y(n)$ as the output and N is the frame length. The gain control factor is expressed as

$$G_s = \sqrt{\frac{\sum_{n=1}^N x^2(n)}{\sum_{n=1}^N y^2(n)}}$$

3 Experiment

3.1 Database

A database of two native US English speaker sampled at 16 kHz was used for the GMM-HMM experiment [15]. There were 1000 utterances (500 males and 500 females) in the database, among which 900 were applied as training data, and 100 were used for subjective and objective evaluation. Speech signals are windowed by a 25-ms window and shifted every 5-ms. Meanwhile, 24-order Line Spectral Pair (LSP) plus delta, which is transformed into static LSPs and their dynamic counterparts, and delta-

delta were extracted to represent the spectrum. Log F0 plus delta and delta-delta were used to model pitch.

3.2 DNN Training

In DNN-HMM model training, the input feature vector contains binary features for categorical linguistic contexts and numerical linguistic contexts. And, the output feature vector includes a voiced/unvoiced flag (a binary feature that indicates the voicing of the current frame), log F0, LSP, gain, their dynamic counterparts, and Mel-cepstral coefficients. DNN is set with 4 hidden layers and 1024 nodes for each layer. The deflection parameter of Mel-cepstral coefficients is 0.55. An exponential decay function is used to interpolate F0 in unvoiced speech regions. 80% of silence frames are removed from the training data to balance the training data and to reduce the computational cost [16]. The input and output features of DNN are normalized to zero mean and unified variance. At the same time, the backpropagation algorithm based on the small batch stochastic gradient descent algorithm is used to train the weight, the minimum and maximum output of DNN are normalized at (0.01, 0.09). The output of DNN is fed into the speech parameter generation algorithm (MLPG) to produce smooth characteristic parameters with dynamic feature constraints.

3.3 Filtering

The natural voice compared with the corresponding synthesized speech by using DTW to align them in the time domain, and the spectral flatness between each other was calculated every 20 frames [14]. The BSD discrepancy between natural voice and synthesized speech was gained through post filter with different parameters. The BSD discrepancies were superimposed every 20 frames, and minimized BSD discrepancy, which means optimal filter parameters, was received. Spectral flatness and the optimal filter parameters are fitted to obtain the relationship of them. Post filter parameters were determined according to the spectral flatness of synthesized speech. The post filter is used to optimize synthesized speech formant to obtain the final synthesized speech.

4 Results

4.1 Objective Evaluation

Figure 2 shows the logarithmic amplitude spectrums of natural voice, the synthesized speech of traditional post filters and improved post filters. As can be seen from the figure, the peak and valley of natural voice are easily distinguished, but the spectrum of the synthesized speech by using traditional post filters is relatively flat, especially reflecting in the expansion of the high-band formant. On the contrary, there are relatively clear formants in the spectrum of the synthesized speech by using improved post filters, accompanying with less speech distortion. However, it still lacks some details that can cause a certain degree of speech distortion (Fig. 3).

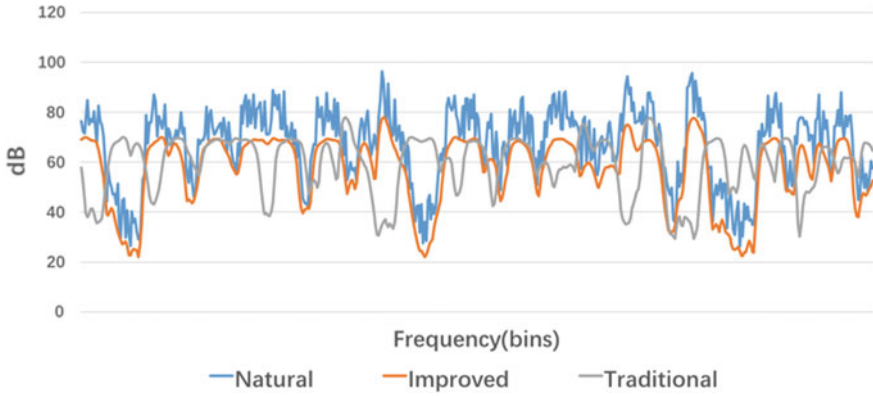


Fig. 2. Log magnitude spectrum for a frame from an utterance in the test set

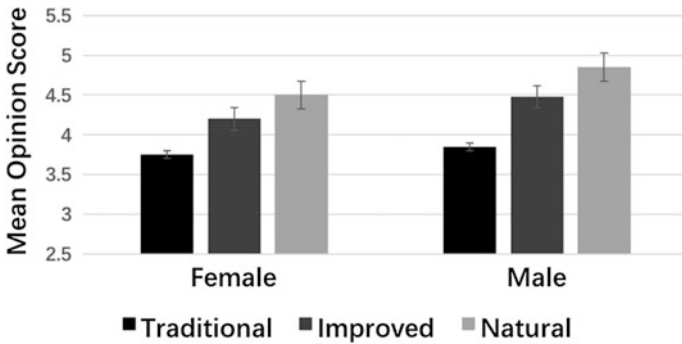
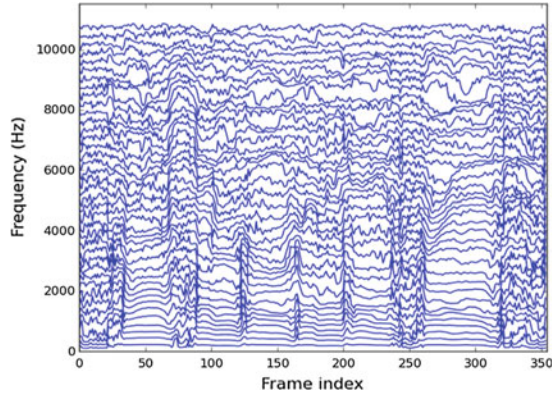


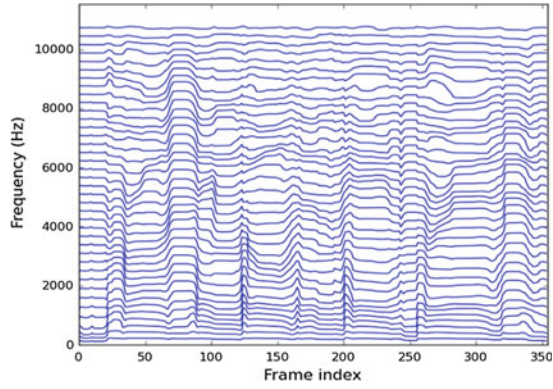
Fig. 3. Naturalness MOS results of synthesized speech (male and female) with traditional and improved post filter

Perceptual Evaluation of Speech Quality (PESQ) is an objective approach to assess Mean Opinion Score (MOS). This paper compares natural voice with 40 correspondent utterances that are dealt with the traditional and improved post filters. Figure 5 illustrates the mean of PESQ. It can be concluded that the improved post filter performs better than the traditional one.

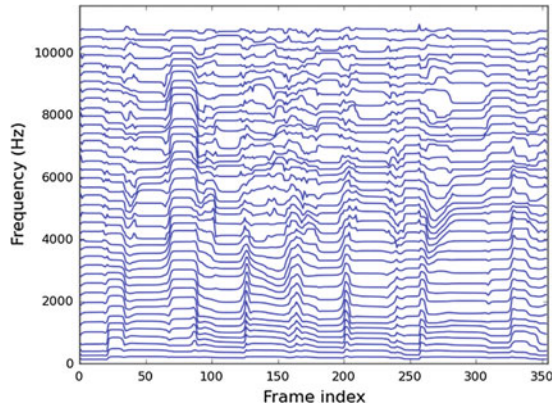
Figure 4 illustrates the trajectory of LSP coefficients of a randomly selected test utterance. In natural voice (a), remarkable details, especially in the high-frequency band, can be acquired in the process of observing the trajectory. By contrast, the trajectory of synthesized speech based on the traditional post filter is comparatively smooth (b). Compared with the former one, the level of fluctuation is augmented in the high order LSPs.



(a) Natural



(b) Traditional Postfilter



(c) Improved Postfilter

Fig. 4. LSP trajectory for an utterance in the test set, **b** and **c** synthesized with features generated by using aligned duration

4.2 Subjective Evaluation

Subjective experiments were conducted by using a contrast selection method. 30 synthesized sentences, respectively, selected from the traditional post filter and the improved one, was randomly played to 100 testers. Each tester is given three preferences, and they can only have one choice: the former is better, the latter is better or no difference (difference between natural voice and synthesized speech is difficult to perceive). After the test, the concluded statistic proportion, which represents the choice results of three preferences, can be gained. The preference scores are presented in Fig. 5 (male and female).



Fig. 5. The preference scores of synthesized speech (male and male) with traditional and improved post filter

It shows that the improved post filter can yield more satisfying synthesized speech than the traditional post filter. Its preference score (51%) is conspicuously higher than the traditional one (29%). However, there are still 20% testers holding the view that the difference between the two is negligible. Based on the above results, a conclusion can be drawn that the improved post filter generates synthesized speech with better sound quality and higher naturalness.

5 Conclusion

In this paper, a DNN-based improved speech synthesis with post filter is proposed to deal with the problem of smoothness and low naturalness existing in HMM-based speech synthesis. Currently, the speech synthesis based on DNN has outperformed the HMM-based speech synthesis. Therefore, in this paper, DNN is applied to the speech synthesis model. The synthesized speech deals with the post filter with different parameters, and after comparing with natural voice, exponential relationship can be obtained between the optimal filter parameter and the spectral flatness. After using the optimal parameter post filter to filter the test speech, the speech's formant was enhanced. Both objective and subjective evaluations showed that the method proposed in this paper effectively improved the smoothness and low naturalness of synthesized speech, which better matches the human's hearing. In future work, we will refine the DNN-based speech synthesis model to reduce computations and introduce more input features including more weak features such as emphasis tone.

Acknowledgements. This paper is supported by the URTP project of School of SME, the demonstration course project of Xidian University, and the Ministry of Education cooperation collaborative education project.

References

1. Dahl, G.E., Yu, D., Deng, L., et al.: Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition. *IEEE Trans. Audio Speech Lang. Process.* **20**(1), 30–42 (2012)
2. Qian, Y., Fan, Y., Hu, W., et al.: On the training aspects of deep neural network (DNN) for parametric TTS synthesis. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 3829–3833. IEEE (2014)
3. Ze, H., Senior, A., Schuster, M.: Statistical parametric speech synthesis using deep neural networks. In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 7962–7966. IEEE (2013)
4. Yoshimura, T., Tokuda, K., Masuko, T., et al.: Incorporating a mixed excitation model and postfilter into HMM-based text-to-speech synthesis. *Syst. Comput. Jpn.* **36**(12), 43–50 (2005)
5. Takamichi, S., Toda, T., Neubig, G., et al.: A postfilter to modify the modulation spectrum in HMM-based speech synthesis. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 290–294. IEEE (2014)
6. Ling, Z.H., Wu, Y.J., Wang, Y.P., et al.: USTC system for Blizzard challenge 2006 an improved HMM-based speech synthesis method. In: *Blizzard Challenge Workshop (2006)*
7. Deng, L.: *Analysis of Deep Learning*. Publishing House of Electronics Industry (2016)
8. Hinton, G.E., Osindero, S., Teh, Y.W.: A fast learning algorithm for deep belief nets. *Neural Comput.* **18**(7), 1527–1554 (2006)
9. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. *Cogn. Model.* **5**(3), 1 (1988)
10. Krogh, A., Hertz, J.A.: A simple weight decay can improve generalization. In: *NIPS*, vol. 4, pp. 950–957 (1991)
11. Grancharov, V., Samuelsson, J., Kleijn, W.B.: Distortion measures for vector quantization of noisy spectrum. In: *INTERSPEECH 2005 - Eurospeech, European Conference on Speech Communication and Technology, Lisbon, Portugal, September*, DBLP, pp. 3173–3176 (2005)
12. Grancharov, V., Plasberg, J.H., Samuelsson, J., et al.: Generalized postfilter for speech quality enhancement. *IEEE Trans. Audio Speech Lang. Process.* **16**(1), 57–64 (2008)
13. Koishida, K., Tokuda, K., Kobayashi, T., et al.: CELP coding based on mel-cepstral analysis. In: *International Conference on Acoustics, Speech, and Signal Processing*, vol.1, 33–36. IEEE (1995)
14. Ge, Y.K.: *Postfilter Parameter Adapted Speech Synthesis Modified Algorithm*. Advance publish house (2015)
15. Kominek, J., Black, A.W.: The CMU Arctic speech databases. In: *Fifth ISCA Workshop on Speech Synthesis* (2004)
16. Fan, Y., Qian, Y., Soong, F.K., et al.: Multi-speaker modeling and speaker adaptation for DNN-based TTS synthesis. In: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4475–4479. IEEE (2015)



Research on Machine Translation Model Based on Neural Network

Zhuoran Han^(✉) and Shenghong Li

Shanghai Jiao Tong University, Shanghai, China
{hzrtom, shli}@sjtu.edu.cn

Abstract. Machine Translation is an important part of Natural Language Processing. The model based on convolution neural network and attention mechanism (Fcnn model), which was proposed by Facebook in 2017, has been successful. We use this Fcnn model as the baseline model of this subject, and on the base of this model, we try to improve it by the combination of bytes pair encoding method, model ensemble method. In this issue, we use Bilingual Evaluation Understudy (BLEU) as a criterion to measure the quality of translation. After testing, these methods can improve the translation quality of the model. Finally, the overall translation quality increased from 0.28 of the baseline Fcnn model to 0.32.

Keywords: Machine translation · CNN · BLEU

1 Background

In recent years, with the rise of the deep learning method represented by the convolution neural network, recurrent neural network and the efficient computing means represented by the GPU, deep learning has begun to combine with many fields such as computer vision and Natural Language Processing, and the performance of the method is greatly exceeded. Machine Translation, as an important topic in Natural Language Processing, naturally attracts the interest of machine learning workers. The purpose of this study is to improve the existing problems on the basis of the existing domestic and foreign research results.

Before the convolution neural network entered the field of vision [1], the mainstream of Machine Translation was based on the rules and the statistical model. Rule-based approach tries to translate from the vocabulary, grammar and other aspects by artificial rules. This method is often too rigid. According to the method of statistics, translation is considered as a kind of probability problem. For the same sentence, many translations may be translated, but the probability of being selected as the final translation is different according to the quality of the translation. Statistical Machine Translation's representative model is Och's model, which was based on the maximum cross entropy [2]. In addition, as a common standard for evaluating the quality of translation, BLEU is the main evaluation standard used in this paper, and it can also be classified as a kind of statistics model-based method [3].

After the rise of neural network represented by deep learning, researchers have begun to combine Machine Translation with neural network. After the proposed word2vec method [4], the researchers can describe words or words with one-dimensional vector, which is easy to calculate and be rich in description ability, which paved the way for the application of neural network to the field of Natural Language Processing. After that, the encoder–decoder structure, as an end-to-end [5] learning framework, puts forward a fairly standard model building method for Machine Translation and other Natural Language Processing problems. On this basis, CNN, RNN, and other networks have been tried for Machine Translation, and have achieved some results. But an important problem is that when dealing with long sentences, neural network is hard to “remember” the information in the foregoing, which leads to the phenomenon of less translation. In terms of network architecture, the attention mechanism is proposed to solve this problem [6]; in the terms of network unit, the unit structure with “memory” represented by LSTM [7] is proposed, and the two methods have achieved significant performance enhancement.

The main research content of this topic is how to optimize the Machine Translation based on neural network to meet the practical application needs. In this paper, we regard the BLEU index as a measure of the quality of the translation, and the improvement of the translation effect is quantified to improve the BLEU evaluation index of the translation results. On the base of Fcnn model, we try to improve it by the combination of bytes pair encoding method and model ensemble method.

2 Fcnn Model

In this paper, we use the Fcnn model as the basic model, and optimize the existing problems of the basic model combined with a variety of optimization methods. Next, we will briefly review the model first.

When a translation of the original text $\mathbf{x} = (x_1, x_2, \dots, x_n)$ is input, the model first translates it into $\mathbf{s} = (s_1, s_2, \dots, s_n)$ through the source language word embedding, where s_i is the word vector representation of x_i , $s_i \in \mathbb{R}^d$, d is the word vector dimension. In this paper, we use word2vec as the word embedding method. To put it simply, word2vec is a way to express words in a vector way that computers can understand. The transformed \mathbf{s} is encoded as context vector \mathbf{c} by the encoder volume layers. In vector \mathbf{c} , the semantic information of the source text \mathbf{x} is included. Then, go into the decoder stage. At this stage, the attention mechanism is set up between each layer of volume. By calculating the vector \mathbf{c} and \mathbf{s} , the decoder can correctly assign the weight to each word and focus on the important semantic part. After the last layer of decoder is completed, the model obtains the word vector of the translation result to represent \mathbf{t} . Finally, through the word embedding operation of the target language, the \mathbf{t} is translated into readable language, and the translation ends. The above is the basic translation process of the Fcnn model. Next, we will focus on the two parts of the model, the operation of convolution neural network layer and the attention mechanism set up in decoder.

2.1 Convolutional Neural Network

For encoder and decoder networks, they have the same algorithm for calculating the intermediate state for a fixed length input. We mark the output of the l th layer of the encoder network as $h_l = (h_l^1, h_l^2, \dots, h_l^n)$, and the output of the l th layer of the decoder network as $j_l = (j_l^1, j_l^2, \dots, j_l^n)$. Each layer consists of a convolution function and a nonlinear activation function. For a convolution kernel with a width of k , each intermediate state h_l^j contains information about k elements in the input vector. Obviously, when multiple layers are stacked up, the number of primitive input elements that can be represented by the intermediate output state will increase.

The Fcnn model chooses GLU as activation function. We mark the output of the convolution function as $\mathbf{Y} = [\mathbf{A} \mathbf{B}]$. Then, we have Eq. (1):

$$V([\mathbf{A}\mathbf{B}]) = \mathbf{A} \otimes \sigma(\mathbf{B}) \quad (1)$$

Here, $\mathbf{A}, \mathbf{B} \in \mathbb{R}^d$, they are the input of the activation function; $V([\mathbf{A} \mathbf{B}])$ is the output of the activation function; \otimes is the dot product; $\sigma(\mathbf{B})$ is a control gate.

After the calculation of the last layer of the decoder, we could calculate the vector presentation of the final result. Finally, we use word embedding again to transform the vector into the target language.

2.2 Attention Mechanism

An input of the encoder is handled by the encoder network into a context vector. As for the calculation in decoder network, we have Eq. (2):

$$y_{i+1} = f_1(y_1, y_2, \dots, y_i, \mathbf{c}, \mathbf{h}_i) \quad (2)$$

Here, y_i is the i th word of the translation result, \mathbf{c} is the context vector, and \mathbf{h}_i is the mediate state of the decoder network. It can be seen that \mathbf{c} remains unchanged while i changes. This makes the output sequence of the output sequence lose contact with the output sequence. It can also be understood as that the translation model at this time can only focus on the part and lack the understanding of the integrity of semantics. In this case, the accuracy of decoder translation for long sentences will be significantly reduced.

Aimed at the intermediate state vector is invariable when the encoder–decoder model generating each element, a weight parameter is added to allocate the weight for the intermediate state vector. In this way, when the different parts of the input sequence are processed, the decoder can give up the weight of the larger part of the current element, thus improving the translation effect.

Figure 1 shows the equation of decoder network changes into Eq. (3) with the attention mechanism.

$$y_{i+1} = f_1(y_1, y_2, \dots, y_i, \mathbf{c}_i, \mathbf{h}_i) \quad (3)$$

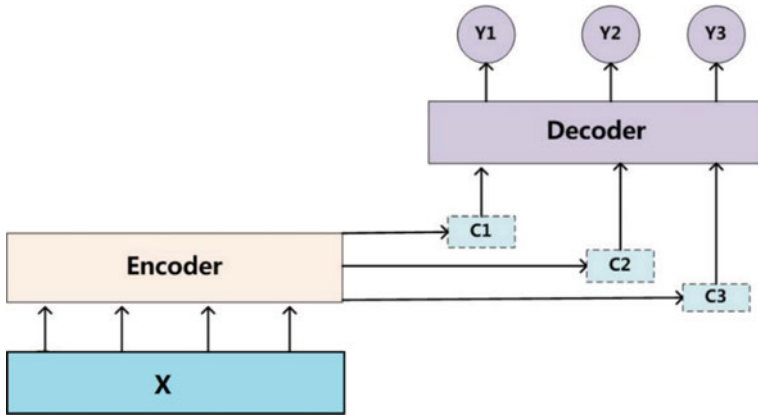


Fig. 1. Diagram of attention mechanism

c_i is calculated as Eq. (4)

$$c_i = \sum_{j=1}^{T_x} \alpha_{ij} h_j \tag{4}$$

h_j is the state of the last layer of the decoder network; α is a weight parameter; T_x is the number of layers.

3 Optimization Methods

The idea of the basic Fcnn model is unique, and has achieved good results in the English–German test dataset, but its effect in the English–Chinese dataset is not satisfactory. We think the main problems existing in the basic Fcnn model are: difficulty when dealing with low-frequency words, the phenomenon of gradient disappearance when the number of network layers increases, and the phenomenon of low training loss but not high BLEU value. In view of low-frequency words, we use bytes pair encoding method to optimize. In addition, the model ensemble method is an effective way to improve the performance of the model, so it is also used to optimize the basic model.

3.1 Bytes Pair Encoding

Bytes pair encoding can be regarded as a participle operation with granularity between characters and words. Frequently used words with high frequency are still in the form of words and phrases. Low-frequency words are segmented by sub-words. Finally, new words can be formed through stitching. For example, if word “solarsystem” is used in the BPE method, it will be divided into two words: solar and system. After using the BPE method, the translation of the rare words such as organization names and personal names and the translation of combinatorial words will be effectively improved.

First, the Chinese and English corpus is segmented separately. The processing of English corpus is simple, only according to the space participle between words, and

then the unification of the lowercase, the English word list vocab_en_1 is obtained. The Chinese corpus needs to be the first participle. Here, we use the word segmentation mode of Jieba word segmentation to divide, and then divide the result statistics to get the Chinese word list vocab_cn_1. In these two dictionaries, when generating, the word needs to be sorted according to the frequency of occurrence.

Next, we are going to make the BPE character dictionary. First, take the English corpus as an example. First, we initialize the BPE character dictionary in English letters, and express each word as a string and add a special terminator, which is used as “@ @”. The function of this special character is to reconvert the BPE character to the normal word after the translation. Next, we count the most frequent character pairs (“A”, “B”) and replace them with a new character “AB”. Repeat this merge until the number of merging operations reaches a preset parameter. This parameter is used to describe the granularity of BPE participle. Drawing on the previous work related to BPE and Natural Language Processing, we set the number of BPE characters to 32,000.

The main purpose of the BPE method is to improve the phenomenon of low-frequency words. However, the BPE method cannot completely avoid the emergence of <unknown> words. In fact, for several sources of <unknown> words, untrained or poorly trained low-frequency words, decoder-decoded vectors can correspond to no words, the source text has not appeared in the corpus of corpus, after BPE treatment, the first and third conditions can be improved. However, because the degree of BPE segmentation is limited by the number of BPE parameters, it is possible that some low-frequency words cannot be fully trained. The second case is generally due to inadequate training of the model. As for the third case, knowledge that has not been learned is still unavoidable for human beings, and Machine Translation is the same. After the combination of the basic Fcnn model and the BPE method, the BLEU of the translation result is worth upgrading.

3.2 Model Ensemble Method

Model ensemble, the basic idea is for the same task, training a variety of solution models, and then the results of multiple models are weighted average, so that multiple models can be made up and improve the quality of the final results. In this paper, we use the following methods to get different Machine Translation models.

- For the basic model of Fcnn model, we have saved multiple checkpoint files with different iterations and different network parameters in the training process, and different models are obtained.
- The tensor2tensor model of Google [8].
- Traditional NMT model based on neural network for Machine Translation RNN + LSTM [9].

The above model is in the decode stage, the process begins with the start character, and the output of the last layer of the decoder is obtained through the decoder, and the output is softmax to get the probability distribution of the length of the word list. This probability distribution is the probability of the word to be translated into individual words. Then, beam search is updated according to the probability distribution, and candidate list is updated, and then the next word is translated. Finally, after translating

the terminator, one of the candidates with the greatest probability in the list is chosen as the final translation result. Our ensemble method is in the decode process of each model. After calculating the probability distribution of softmax, the weighted average of the softmax vector is weighted, and the average result is unified as the softmax probability distribution of all the current time of the model, and then the translation process is continued according to the new average distribution.

4 Experiment

In Table 1, the base Fcnn model is trained to be convergent under different model parameters (including the number of encoder and decoder network layers, convolution kernel parameters), and the training loss value of the model and the BLEU index value of the model translation results. The 5 and 7 selection of the network layer and the 3 and 5 of the convolution kernel width are chosen. However, the optimal model parameters are generally different for different source language–target language pairs, which should be related to linguistic characteristics. In addition, because there is only one computer, the training time is more intense, so there are not many parameter combinations.

Table 1. Fcnn model’s test data

Encoder	Decoder	Width	Training loss	BLEU
5	5	3	2.791	27.87
7	7	3	3.880	26.52
5	5	5	2.722	28.13
7	7	5	3.874	26.78

According to the data in the table, the convolution kernel width 5 is a better choice than 3. On the network level, the 7 level of theory expression ability should be better than the 5 level, but training loss and BLEU values are all worse. This may be because the learning rate is not good enough, or the training time is not optimal, so that the model does not converge to the best state. In summary, we use the encoder–decoder network layer 5 layers, the convolution kernel width 5, BLEU value 28.13 as the basic model. Then we add BPE.

The number of BPE characters in Table 2 illustrates the particle size of participles. 32,000 of the best results here, we think the low BPE number makes low-frequency

Table 2. Test data of Fcnn model with BPE

Encoder	Decoder	BPE	Training loss	BLEU
5	5	16,000	8.9	26.2
5	5	32,000	2.9	28.7
5	5	48,000	12.6	22.8

words did not improve very well, and the high BPE number requires longer training time, perhaps more complex model structure. After selecting the BPE number of 32000, the BLEU value of the model increases.

Next, we add the model fusion method on the basis of the BPE method. The test results of the Fcnn + BPE + model fusion method are shown in Table 3.

Table 3. Test data of Fcnn model with BPE and model ensemble

Model	BLEU
Base model	28.7
Base model + tensor2tensor model	29.3
Base model + tensor2tensor model + NMT model	32.12

It can be found that the model fusion method has a significant effect on BLEU value enhancement, and the addition of each new model is improved. It is also proved that the previous model fusion theory, that is, different types of models or the same type of different parameter combination models, can be used to model fusion.

5 Conclusion

In this paper, we use Fcnn Machine Translation model as a basic model, and carry out a series of optimization for its shortcomings. We think the main problems existing in the basic model are: it is difficult to deal with low-frequency words, the phenomenon of gradient disappearance when the number of network layers increase, and the phenomenon of low training loss but low BLEU value. We use this Fcnn model as the baseline model of this subject, and on the base of this model, we try to improve it by the combination of bytes pair encoding method, model ensemble method. After testing, these methods can improve the translation quality of the model. Finally, the overall translation quality increased from 0.28 of the baseline Fcnn model to 0.32.

References

1. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. *Commun. ACM* **60**(2), 2012 (2012)
2. Och, F.J., Ney, H.: Discriminative training and maximum entropy models for statistical machine translation. In: *Meeting on Association for Computational Linguistics*. Association for Computational Linguistics, pp. 295–302 (2002)
3. Papineni, K., Roukos, S., Ward, T., et al.: BLEU: a method for automatic evaluation of machine translation. In: *Meeting on Association for Computational Linguistics*. Association for Computational Linguistics, pp. 311–318 (2002)
4. Mikolov, T., Chen, K., Corrado, G., et al.: Efficient estimation of word representations in vector space. *Comput. Sci.* (2013)
5. Sutskever, I., Vinyals, O., Le, Q.V.: Sequence to sequence learning with neural networks, vol. 4, pp. 3104–3112 (2014)

6. Bahdanau, D., Cho, K., Bengio, Y.: Neural machine translation by jointly learning to align and translate. *Comput. Sci.* (2014)
7. Graves, A.: Long short-term memory. *Supervised Sequence Labelling with Recurrent Neural Networks*, pp. 1735–1780. Springer, Berlin (2012)
8. Vaswani, A., Bengio, S., Brevdo, E., et al.: Tensor2Tensor for neural machine translation. *CoRR* (2018)
9. Luong, M.T., Brevdo, E., Zhao, R.: Neural machine translation (seq2seq) tutorial (2017). <https://github.com/tensorflow/nmt>



Multi-node Repair Based on GA_PSO with Fractional Regenerating Code Combined with Prior Replication

Niannian Wang^(✉), Ye Wang, Jia Yu, and Siyun Chen

Communication Engineering Research Center, Harbin Institute of Technology
(Shenzhen), Shenzhen, Guangdong, China

{wangniannian, chensiyun}@stu.hit.edu.cn, wangye.hitsz@qq.com,
yujia.hitsz@hotmail.com

Abstract. Erasure codes can improve the reliability of modern Distributed Storage Systems (DSS) by preventing data loss and nodes failure. Regenerating code is a class of erasure codes that allow for repairing of failed nodes. However, regenerating code increases the amount of the participating nodes and its coding parameters are difficult to determine. In addition, it has huge computational overhead and low repair efficiency that prohibit its applications. Hence, we first propose a fractional regenerating code combined with prior replication with uncoded repair. Simulation results show that it can reduce repair bandwidth and computational complexity by increasing the number of high prior nodes. Second, we formulate the problem of computing multiple failure repairs cost using the proposed code as a redundancy scheme. We model the problem as an Integer Linear Programming problem (ILP) and solve it by Genetic Algorithm-Particle Swarm Optimization (GA_PSO) algorithm. We present results of repairing bandwidth cost for our proposed algorithm in two scenarios to evaluate the effectiveness of the solution approaches. Simulation results demonstrate that GA_PSO can get smaller repair bandwidth cost than GA.

Keywords: DSS · Fractional regenerating code · Multi-node repair · GA_PSO

1 Introduction

With the arrival of the era of big data, global data traffic presents explosive growth, which puts forward higher requirements for the data storage capacity of the systems. Compared with traditional centralized storage systems, DSS has been widely used due to their low cost, large storage capacity, and strong scalability.

Redundancy must be introduced into DSS to improve reliability against node failures. Increasing redundancy through replication is simple and easy to implement but reduces disk utilization [1]. In 2007, Dimakis proposed regenerating code on the basis of network coding. Regenerating code can solve the problem of excessive recovery costs of traditional erasure correction schemes, however, it needs to increase the number of helping nodes and has high computational complexity [2]. To perform repair, the system downloads some data from a subset of the surviving nodes, called help nodes [3]. To save energy, distributed storage systems are required to periodically replace some of the dormant nodes. Therefore, the number of helping nodes cannot be too large [4]. Thus, in this paper, we first propose multi-node repair fractional regenerating code combined with prior replication with uncoded repair. It can reduce repair bandwidth cost and has zero computational complexity.

Achieving high reliability in distributed storage systems at a reasonable cost can be challenge because they are built from large numbers of ordinary devices which suffer frequent failures. Reed–Solomon (RS) codes have to reconstruct the entire file in order to repair a single node failure, resulting in large consumption of bandwidth and network resources [5]. Although regenerating code is optimal in terms of repair bandwidth, reduction of the imposed computational complexity remains a challenging task [6]. To get the reasonable repair bandwidth cost, we formulate the problem of computing multiple failure repairs cost using the proposed code as a redundancy scheme. We model the problem as an ILP problem. It is an NP-hard problem and we solve it by GA_PSO. Simulation results show that GA_PSO can get smaller repair bandwidth cost than GA.

2 System Model

2.1 Fractional Regenerating Code Combined with Prior Replication

Fractional regenerating code combined with prior replication is based on the traditional regenerating code, so it also meets the Maximum Distance Separable (MDS) characteristics. First, the original data of size M is divided into k data blocks. After encoded, the data blocks are distributed to n nodes and each storage node stores data size as $\alpha = \frac{M}{k}$. We taken k from any n nodes to reconstruct the original data. Then, the n nodes are divided into two different priorities according to the access frequency of the nodes. The high-priority nodes replication multiplier is p_2 , otherwise the replication multiple is p_1 , $p_1 * (n - m) + p_2 * m$ ($p_2 \geq p_1$). Where m is the number of nodes with high priority. When a node fails, a new node downloads data from a replica node that stores failure data and repair bandwidth is α . When all replica nodes fail, a new node downloads size β from d helper nodes respectively ($\beta \leq \alpha$) and repair bandwidth $\gamma = d * \beta$ ($\gamma \geq \alpha$). Where (n, m, p_1, p_2, k, d) are nonnegative integers, (α, β, γ) are nonnegative real numbers. In particular, when $m = 0$ and $p_1 = 1$, it is traditional regenerating code.

For arbitrary $\alpha \geq \alpha^*(n, k, d, \gamma)$, any node (n, k, d, γ) is feasible and linear network coding can be achieved [3]. The threshold function is shown in Eq. (1)

$$\alpha^*(n, k, d, \gamma) = \begin{cases} \frac{M}{k}, \gamma \in [f(0), +\infty) \\ \frac{M-g(i)\gamma}{k-i}, \gamma \in [f(i), f(i-1)) \end{cases} \quad (1)$$

α indicates the amount of data stored in each node, the new node connects d helper nodes and downloads data of size β to repair the failed node. Its recovery bandwidth $\gamma = d * \beta$, where $f(i)$ is defined as shown in Eq. (2) and $g(i)$ is shown in Eq. (3).

$$f(i) = \frac{2Md}{(2k-i-1)i+2k(d-k+1)} (i = 1, 2, \dots, k-1) \quad (2)$$

$$g(i) = \frac{(2d-2k+i+1)i}{2d} (i = 1, 2, \dots, k-1) \quad (3)$$

According to Eqs. (1)–(3) we can get:

As for an eight-tuple $(n, k, d, \alpha, \gamma, p_1, p_2, m)$, according to the minimum cut capacity lower bound in the information flow diagram, the following formula can be obtained. Where $\alpha^*(n, k, d, \gamma, m)$ is the threshold function of α [7]:

$$\alpha^*(n, k, d, \gamma, m) = \begin{cases} \frac{M}{k}, \gamma \in [p(m), +\infty) \\ \frac{M-g(i)*\gamma+g(m)*\gamma}{(k+m)-i}, \gamma \in [p(i), p(i-1)) \end{cases} \quad (4)$$

$$p(i) = \frac{2f(i)*md}{2md+f(i)*(2mi-m^2+m)}$$

where $f(i)$ is defined as shown in Eq. (2) and $g(i)$ is shown in Eq. (3).

2.2 Minimum Storage Regenerating (MSR) and Minimum Bandwidth Regenerating (MBR)

Fractional regenerating code combined with prior replication is similar to the regenerating code. That means there are also two special points of *MSR* and *MBR*. *MSR* is a point that can reach the lower bound of storage cost and has the smallest bandwidth cost. *MBR* is the point with minimal storage cost under the condition of optimal bandwidth cost. In the traditional regenerating code, the minimum storage cost equals to $\alpha = \frac{M}{k}$. Under this condition, the minimum bandwidth cost is $\gamma = f(0) = \frac{Md}{k(d-k+1)}$. We can get the minimum storage cost of the *MSR* as shown in Eq. (5).

$$(\alpha_{MSR}, \gamma_{MSR}) = \left(\frac{M}{k}, \frac{Md}{k(d-k+1)} \right) \quad (5)$$

The bandwidth cost γ is a decreasing function of the number of helper nodes d , which is the minimum value reached when connecting to $d = n - 1$ nodes [8], as shown in Eq. (6).

$$(\alpha_{MSR}, \gamma_{MSR}^{\min}) = \left(\frac{M}{k}, \frac{M}{k} \cdot \frac{n-1}{n-k} \right) \quad (6)$$

Another extremum point is the MBR that achieves the lowest bandwidth cost. In the traditional regenerating code, $\gamma \in [f(i), f(i - 1))$, and $f(i)$ is a decreasing function of i . So when i takes the maximum $i = k - 1$, $\gamma = f(k - 1)$ reach minimum value, the storage and bandwidth cost of the MBR code are as shown in Eq. (7).

$$(\alpha_{MBR}, \gamma_{MBR}) = \left(\frac{2Md}{2kd - k^2 + k}, \frac{2Md}{2kd - k^2 + k} \right) \tag{7}$$

According to Eqs. (5), (6), we can get MSR of the proposed code, minimal storage cost $\alpha_{MSR} = \frac{M}{k}$. Under this condition, the minimum bandwidth cost:

$$\gamma_{MSR} = p(m) = \frac{2f(m) * md}{2md + f(m) * (2m * m - m^2 + m)} = \frac{Md}{k * (d - k + m + 1)} \tag{8}$$

According to Eq. (7), we can get MSR of the proposed code. $\gamma \in [p(i), p(i - 1))$ and $p(i)$ is a decreasing function on i . So when i takes the maximum $i = k - m$ and $\gamma = p(k - m)$ get the minimum value:

$$\alpha_{MBR} = \gamma_{MBR} = \frac{2Md}{2kd - (k - m)^2 + k - m} \tag{9}$$

3 Problem Formulation and Algorithm

3.1 Numerical Example

Consider a distributed storage system with six nodes and encoded data with four distinct blocks are to be stored. As for fractional regenerating code combined with prior replication, we assume $m = 4$, $p_2 = 3$. It means that we repeated four distinct blocks for three times and each node can store two distinct blocks. Assuming different communication link costs, link transmission speeds, I/O, disk access, and network load can affect the communication cost between distinct nodes. A random data blocks assignment matrix D where rows correspond to nodes and columns correspond to blocks together with a given communication cost matrix cost, is shown in Eq. (10).

$$D = \begin{matrix} n1 \\ n2 \\ n3 \\ n4 \\ n5 \\ n6 \end{matrix} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}_{n * k} \quad Cost = \begin{matrix} n1 \\ n2 \\ n3 \\ n4 \\ n5 \\ n6 \end{matrix} \begin{bmatrix} 0 & 8 & 3 & 6 & 4 & 1 \\ 8 & 0 & 7 & 2 & 4 & 6 \\ 3 & 7 & 0 & 4 & 9 & 8 \\ 6 & 2 & 4 & 0 & 2 & 6 \\ 4 & 4 & 9 & 2 & 0 & 1 \\ 1 & 6 & 8 & 6 & 1 & 0 \end{bmatrix}_{n * n} \tag{10}$$

where $D_{n,k} = \begin{cases} 1, & \text{block.k} \in \text{node.n} \\ 0, & \text{otherwise} \end{cases}$, $Cost_{n1,n2} = 8$ indicates that the communication cost between node 1 and node 2 is 8. Node 1 is help node and node 2 is fail node.

For all possible two-node failure repair patterns, we first check where failure nodes have at least one block in common, because the order of repair affects the repair cost unless the common lost block has its replica available on two helper nodes. For example, when node 2 and node 3 fail at the same time, node 2 and node 3 have one block in common, so only node 5 can repair the common block. We repair failed node 2 through node 5 and 6, and repair failed node 3 by node 5 and 1. In this condition, we get the system repair cost (*SRCost*), $SRCost = C_{13} + C_{53} + C_{52} + C_{62} = 3 + 9 + 4 + 6 = 22$. If we first use node 5 repair node 2, then we use node 2 repair node 3. In this case, we can get $SRCost = C_{52} + C_{23} + C_{42} + C_{13} = 2 + 9 + 2 + 3 = 16$. The total repair cost of the system is to repair the cost of two failed nodes by selecting the optimal repair order.

3.2 Mathematical Formulation of Multiple Failure Recovery Cost

To select the node allocation scheme and get the optimized *SRCost*, we model it as the *ILP* [9]:

$$RCost_{x,y} = \min_{\substack{\alpha=1,2,\dots,N \\ \alpha \neq x, \alpha \neq y}} \sum_{n=1}^N \sum_{k=1}^K C_{\alpha,n} * D_{n,k} * D_{\alpha,k}, (y = x + 1, x \in 1, 2, \dots, N - 1) \tag{11}$$

$$\begin{aligned} \min & \sum_{y=x+1}^N \sum_{x=1}^{N-1} RCost_{x,y} \\ \text{s.t.} & \quad (1) : \sum_{n=1}^N D_{n,k} = p_2 \quad (k = 1, 2, \dots, K) \\ & \quad (2) : \sum_{k=1}^K D_{n,k} = d \quad (n = 1, 2, \dots, N) \\ & \quad (3) : \sum_{k=1}^K S_k \leq SC \quad (n = 1, 2, \dots, N) \end{aligned} \tag{12}$$

Assume that the DSS is for 2-node failure scenarios. $RCost_{x,y}$ is the minimum required cost for failed node x, y. $C_{\alpha,n}$ indicates the communication cost of the helper node and other nodes. The constraint conditions are depicted as follows:

- (1) set the limit on the total number of replicas every block on all nodes.
- (2) set the limit on the total number of blocks that can be stored on any given nodes.
- (3) set the limit on the total storage size every node to be below a given Storage Capacity (SC).

When the network node is large, the problem is NP-hard problem. So we use the GA_PSO to solve this problem.

3.3 Genetic Algorithm_Particle Swarm Optimization (GA_PSO) Algorithm

Using GA_PSO to optimize the function, it can enhance the global optimization ability, speed up the evolution of the algorithm and improve the convergence

accuracy of the algorithm. First, the data blocks are randomly allocated to the nodes as initial members. Each member has an initial speed and has a corresponding fitness value by calculating the fitness function. Then, speed and population are updated based on fitness values. Next, the selection, crossover, and mutation of the genetic algorithm is performed. Finally, the individual and the population are updated optimally until the convergence rule is satisfied. The flowchart of the GA_PSO algorithm is shown below (Fig. 1).

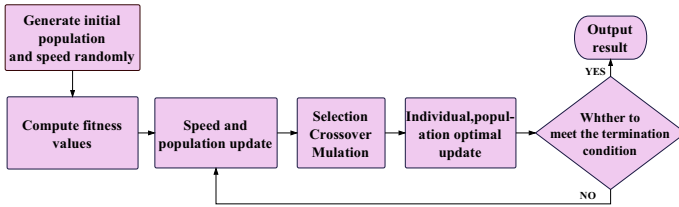


Fig. 1. Flow chart of GA_PSO

4 Simulation Results and Analysis

4.1 Fractional Regenerating Code Combined with Prior Replication

According to Eqs. (2)–(4), we can get stored-bandwidth cost trade-off curve. Figure 2 is the case of $M = 1, n = 10, k = 5, p_1 = p_2 = 1, m = 0, 1, 2$ and d takes a trade-off curve when it is 6 or 9.

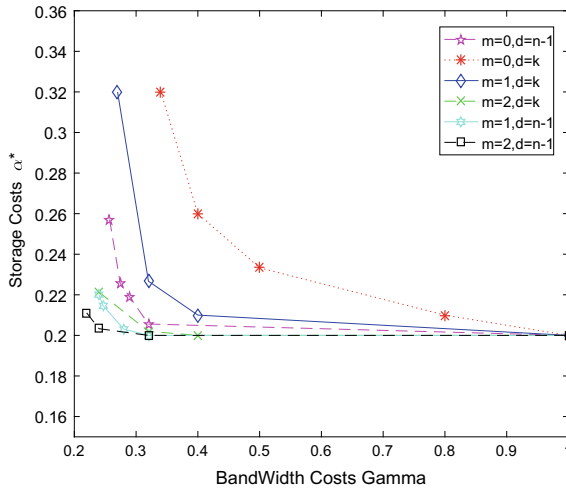


Fig. 2. Stored-bandwidth cost curve

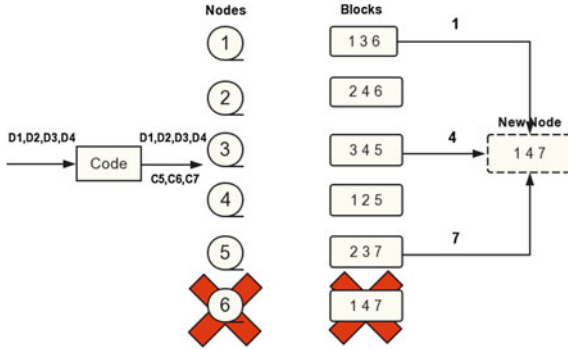


Fig. 3. (6, 4, 3) distributed storage system

In Fig. 2, it is the optimal curve that can be achieved by the traditional regenerating code when $m = 0, d = n - 1$. When the number of helper nodes d is equal, the bandwidth cost decreases as the number of high-priority nodes increases. So, we can decrease the bandwidth cost by increasing the number of high-priority nodes.

To illustrate much more, the operation of fractional regenerating code combined with prior replication, a (6, 2, 3) DSS is depicted in Fig. 3. The system has six nodes. A file of $M=4$ systematic packets ($D1, D2, D3, D4$) needs to be stored in the system. After a (7, 4) MDS code, we obtain three parity packets ($C5, C6, C7$). Systematic packets are accessed frequency, so we repeated three times. Parity packets have a repetition degree of 2. That is, $p_1=2, p_2=3$. When

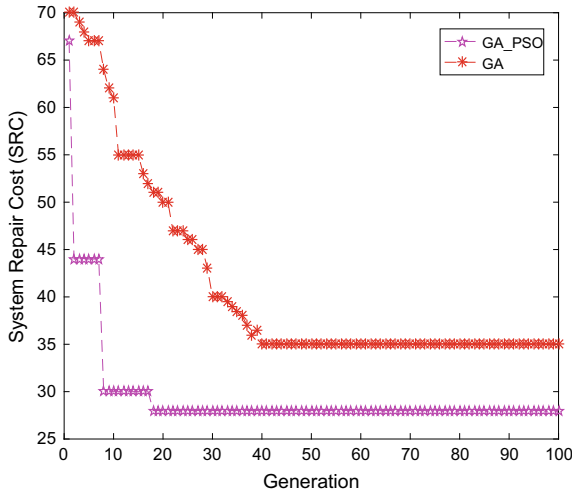


Fig. 4. SRC for $n = 50, k = 50, p_2 = 3$

node 6 fails, a data block is downloaded from node 1, node 3, and node 5 to help node 6 recovery. At this time, the number of helper nodes is 3, and the number of general regeneration code helper nodes is at least 4 for (6, 4) distributed storage system.

4.2 Multi-node Repair Cost Based on GA_PSO

The convergence curves in Fig. 4 are the optimized system recovery cost for all generations. They are for the case of a DSS with 50 storage nodes and 50 distinct data blocks are replicated three times. The data blocks are distributed on the 50 nodes in an optimized scheme for multiple failure recovery. The probability of crossover is 0.9 and the probability of mutation is 0.1. The example is for 2-node failure scenarios. The two curve decreases in every generation except for the case of mutation, this illustrates the fact that better solutions are being generated in newer population. In addition, from Fig. 4, we can get that GA_PSO tends to be more stable than GA and GA_PSO can get smaller repair bandwidth cost.

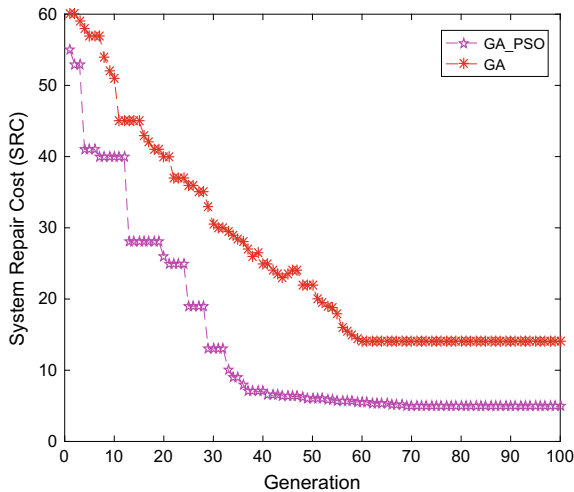


Fig. 5. SRC for n = 50, k = 50, p2 = 6

To avoid being stuck in a local optima , we increase the probability of mutation equals 0.3. At the same time, we increase p₂= 6. It means that we make the system can tolerate much more failures. We get the resulting curve in Fig. 5 by GA_PSO and the repair bandwidth cost is smaller than in Fig. 4.

5 Conclusions

In this paper, we propose fractional regenerating code combined with prior replication. It can effectively reduce the repair bandwidth and computational complexity. For the problem of computing multiple failure recovery cost with fractional regenerating code combined with prior replication, we solve it by GA_PSO.

Simulation results demonstrate that GA_PSO can get smaller repair bandwidth cost than GA.

Acknowledgments. This work has been supported in part by the National Natural Sciences Foundation of China (NSFC) under Grants 61501140, 61701136, and 61525103.

References

1. Gerami, M., Xiao, M., Skoglund, M.: Two-layer coding in distributed storage systems with partial node failure/repair. *IEEE Commun. Lett.* **21**(4), 726–729 (2017)
2. Dimakis, A.G., Godfrey, P.B., Wu, Y., Wainwright, M.J., Ramchandran, K.: Network coding for distributed storage systems. *IEEE Trans. Inf. Theory* **56**(9), 4539–4551 (2007)
3. Mahdavian, K., Mohajer, S., Khisti, A.: Product matrix MSR codes with bandwidth adaptive exact repair. *IEEE Trans. Inf. Theory* **64**(4), 3121–3135 (2018)
4. Li, J., Wang, X., Li, B.: Pipelined regeneration with regenerating codes for distributed storage systems, pp. 1–6 (2011)
5. Li, J., Li, B.: Beehive: erasure codes for fixing multiple failures in distributed storage systems, pp. 6–6 (2015)
6. Papailiopoulos, D.S., Luo, J., Dimakis, A.G., Huang, C., Li, J.: Simple regenerating codes: network coding for cloud storage. In: *Proceedings - IEEE INFOCOM* (2012)
7. Xu, L., Pavlo, A., Sengupta, S., Li, J., Ganger, G.R.: Reducing replication bandwidth for distributed document databases, pp. 222–235 (2015)
8. Hu, Y., Lee, P.P.C., Shum, K.W.: Analysis and construction of functional regenerating codes with uncoded repair for distributed storage systems, pp. 2355–2363 (2012)
9. Itani, M., Sharafeddine, S., Elkabani, I.: Dynamic multiple node failure recovery in distributed storage systems. *Ad Hoc Netw.* **72**, 1–13 (2017)



A Super-Resolution Reconstruction Algorithm Based on Learning Improvement

Han Gao, Xinwei Li, and Aiping Jiang^(✉)

Heilongjiang University, Harbin, China
gh19951230@163.com, 863301127@qq.com, jap62@126.com

Abstract. In view of the problem of edge blurring and slow reconstruction speed in the existing super-resolution reconstruction of learning-based images, this paper proposes an improvement to the original learning-based reconstruction algorithm and applies Markov random fields to image super-resolution. In the rate reconstruction, during the dictionary training phase, training image blocks are randomly selected, and the texture part of the image is learned and reconstructed. The bicubic interpolation method is used to enlarge the image structure part and color information, the final reconstructed image and the interpolation-amplified image are merged. That is the final result. The experimental results show that the peak signal-to-noise ratio (PSNR) is used to objectively evaluate the image reconstruction effect, and it is concluded that the algorithm of this paper reconstructs the image better, and the reconstruction time also has a certain increase.

Keywords: Learning-based · Super-resolution reconstruction · Image decomposition · Markov random field

1 Introduction

There are many ways to perceive the world. Images are one of the indispensable ways for humans to obtain outside information. The quality of images is closely related to the resolution of the images. With the rapid development of information technology, people's requirements for image resolution have also been greatly improved. However, due to the influence of multiple link factors such as imaging system and graphics acquisition, many image resolutions were not able to meet people's requirements. Therefore, we constantly explore ways to improve image resolution. Image super-resolution reconstruction is one of the effective methods. This technology has very good practicality, has been widely used in many areas of image processing, and has a wide application prospect. Therefore, research on this technology has very important practical significance.

Image super-resolution reconstruction is a reconstruction technique that uses low-resolution images to recover high-resolution images. Nowadays, there are two main types of reconstruction methods. One is a reconstruction-based image super-resolution reconstruction [1]. The other is learning-based image super-resolution reconstruction [2, 3]. The learning-based method is a new reconstruction method that has emerged in recent years. This method can increase the multiples of resolution more. Such methods mainly include Freeman [4] on the basis of low-order visual problems and propose an example-based reconstruction method. Yang [5] proposed reconstruction using sparse representation of image blocks. Tang [6] proposed a local learning method to reconstruct a single image.

This paper improves on the basis of Freeman [7] algorithms reconstruction. This method relies on the relationship between low-resolution images and their corresponding high-resolution images to estimate the high-frequency details are lost in low-resolution images. This paper randomly selected training. In the image block, the obtained block image is used as a training database in the training stage, and the low-resolution image block and the high-resolution image block are simulated in a Markov network. In the reconstruction stage, an interpolation method is used to enlarge the image structure part and the color information. The final reconstructed image and the interpolated magnified image are merged to obtain the final result.

2 Learning-Based Image Super-Resolution Reconstruction Algorithm Improvement

2.1 Preprocessing of Training Images

First, preprocessing of training images is performed. In this preprocessing step, a group of HR images is selected to form a high-resolution image set, and then all the HR images are, respectively, taken to be blurred and down-sampled. One-quarter of the original image can be called an input image. The same size is called the output image. The input image is interpolated to obtain a low-resolution image (LR) without high-frequency details. The following Fig. 1a is the result of the original HR downsampling and Fig. 1b is Fig. 1a LR image obtained after interpolation. Figure 1c is the original high-resolution image, Fig. 1d is the result of the band-pass filtering, that is the intermediate frequency information about the LR image to be reconstructed. Figure 1e is high-frequency information to be supplemented for the LR image.

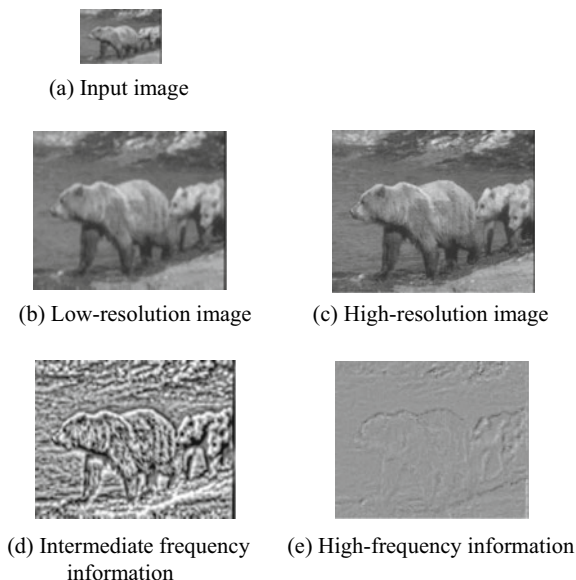


Fig. 1. Preprocessing of training images

2.2 Training Phase

At this stage, the intermediate and high-frequency images obtained by preprocessing are used to build a database, and the image blocks of the images in the database are randomly selected. During the process of generating the training database in blocks, overlapping sampling is needed to obtain training data. On the one hand, with overlapping sampling, so block-like blurring can be avoided, and on the other hand, the relationship between adjacent high-frequency blocks can be limited to obtain more effective data. As shown in Fig. 2a is part of the IF block, Fig. 2b is part of the high-frequency block.

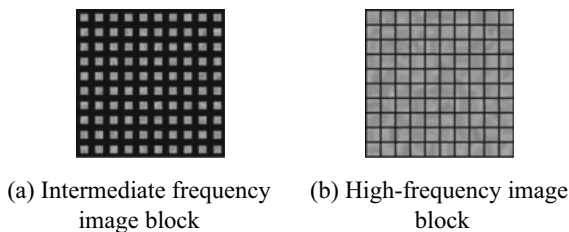


Fig. 2. Image blocks in training

In this paper, the Markov network will be used to simulate the relationship between image blocks to facilitate the clearest expression of the relationship between high-resolution images and low-resolution images.

2.3 Reconstruction Stage

2.3.1 Texture Image Reconstruction

First of all, the color and space need conversion. Because there is a possibility that the color information of the training set image and the color information of the LR image to be reconstructed may not be consistent, in order to reduce the influence of this factor, the gray information of the image should be reconstructed.

Second, the image is decomposed. After the YCbCr transformation is performed on the image to be reconstructed and trained, the gray component (the Y component) of the image is selected, and the total variation decomposition based on the L2 norm is performed to obtain the structure and texture information of the image. Then, the texture part in the high-resolution image block is selected to be used as a high-resolution image in the training set, and the texture part of the input image is interpolated to obtain a low-resolution image as the low-resolution image to be reconstructed in this paper. The following figures show the texture decomposition of HR and LR images. The texture and structure of HR and LR are obtained. Figure 3a is an HR image, Fig. 3b is an LR image, Fig. 3c, d are the image structure parts obtained by HR and LR decomposition, and Fig. 3e, f are the image texture parts obtained after decomposition.

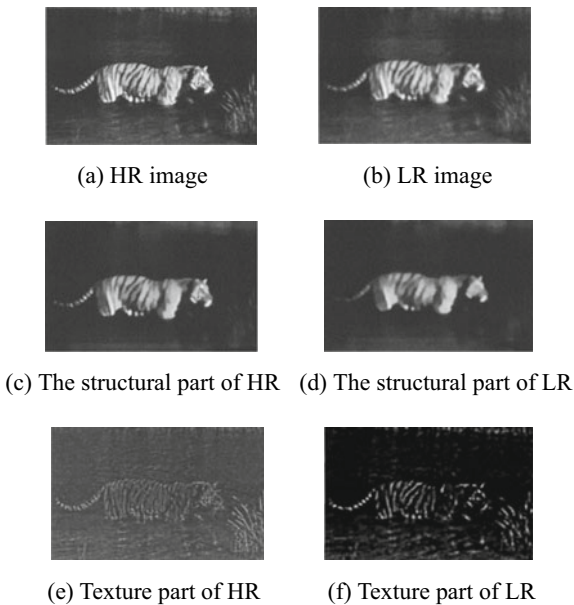


Fig. 3. Image decomposition

Third, texture image learning. First, during the training phase, a training image set is acquired. The selected HR image is preprocessed, such as color space conversion, and then the image is decomposed by texture to obtain a high-resolution image in the training set. The high-frequency image and the intermediate frequency image are divided into blocks by training the database. Finally, an image training database is constructed by using the intermediate frequency image block and the high-frequency image block, and all the image blocks are normalized. The Markov network simulation was also used to train the relationship between the image blocks.

The next step is the reconstruction phase. The main contents are as follows: the input image is obtained by downsampling the test image, and YCbCr conversion is performed; the Y component of the image can be processed by texture decomposition, the structure and texture information can be obtained by this method; the input part of the texture into interpolation processing to get the desired low-resolution texture image and its high-frequency information; the high-frequency information obtained by the block processing, it can be found at the same time in the training set to find the most similar image block, and corresponding high-resolution texture image blocks can be obtained; merges the texture areas of all the high-frequency image blocks and the high-resolution blocks, and takes the average of all the high-resolution texture image blocks in the overlapping portion as a result of the reconstruction of this block region. The reconstruction results are shown below, as shown in the following Fig. 4.



Fig. 4. Texture reconstruction results

Finally, the images are combined to combine the structure of image texture decomposition, image color information, and texture reconstruction structure to obtain the final reconstruction results, as shown in the following Fig. 5.



Fig. 5. The final results of the reconstruction of this method

3 Analysis of Experimental Results

This paper selects some images in the BSDS database for image reconstruction to verify the effectiveness of this method. The bicubic interpolation algorithm was used to process the original algorithm based on the learning algorithm and the method of this paper to process the input low-resolution image. The main parameters of the algorithm in this paper are as follows: The size of the image block is 5×5 , the number of image blocks that are trained is 100,000, and the overlapping pixel setting is 4.

First, select the tiger (108073) image in the BSDS image library, as shown in Fig. 6a is the input LR tiger image, Fig. 6b is the original HR image, Fig. 6c is the Bicubic-derived image, and Fig. 6d is the original image based on the obtained image of the learning algorithm, Fig. 6e is the image reconstructed in this paper.

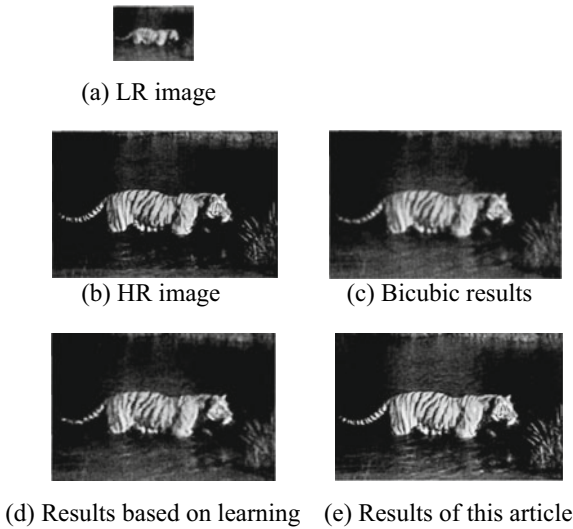


Fig. 6. Super-resolution reconstruction experiment results of the tiger (108073)

Next, select the Lena image in the BSDS graphics library, as shown in Fig. 7a is the input LR image, Fig. 7b is the original HR image, Fig. 7c is the Bicubic image, and Fig. 7d is the original result image, Fig. 7e is the image reconstructed in this paper.

This paper extracts the PSNR to objectively evaluate the reconstruction results of the image during the experiment. The greater the PSNR value, the better the image results obtained with this method.

The expression is as follows:

$$\text{PSNR} = 10 \times \log_{10} \frac{255^2}{\text{MSE}(r - f)} \quad (1)$$

In the above equation, r denotes the reconstructed HR image; f denotes the original HR image; MSE denotes the mean squared error (MSE) function of r and f .

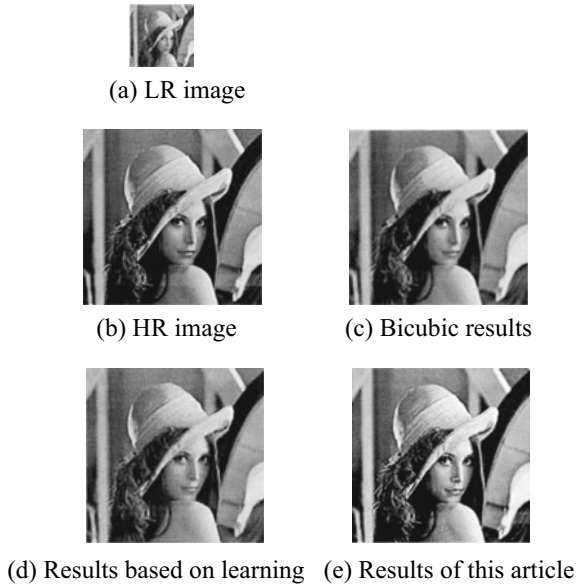


Fig. 7. Super-resolution reconstruction experiment results of the Lena

$$\text{MSE} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (r_{ij} - f_{ij})^2 \quad (2)$$

In the formula, MN represents the total number of pixels in the image; r_{ij} and f_{ij} represent the pixel values at the point (i, j) of the reconstructed HR image and the original HR image.

Comparing the reconstruction results of this paper with the results of the bicubic interpolation method and the original learning-based reconstruction method, the PSNR values are calculated, respectively. The calculated results are shown in Table 1.

Table 1. Comparison of reconstruction results of different methods under PSNR (unit: decibel)

Image name	Bicubic interpolation	Original learning-based approach	Method of this article
Tiger (108073)	21.8328	22.2340	25.3436
Lena	26.3842	25.5427	27.3952

Furthermore, considering the application of super-resolution technology in practice, the actual running speed of various algorithms is also an indicator of the merits of the evaluation algorithm. The original learning-based method and the method of this article are compared on the same platform and the running speed of the same input. The results are shown in Table 2.

Table 2. Rebuild time of this article and original learning-based method (unit: second)

Image name	Original learning-based approach	Method of this article
Tiger (108073)	66.7347	8.2574
Lena	78.3643	17.6563

By comparing the results of Table 1, we can see that the method in this paper is much better than the original learning-based method in terms of image reconstruction quality. It can be seen from Table 2 that this method can greatly improve the speed of image reconstruction and is relatively accurate for detailed information. Fewer tigers (image 108073) can speed up to eight times faster; for Lena images that contain much more detailed information and less relevant information in the training set, the speed can be increased by more than four times, which shows that this method has better timeliness. It can be seen that only the texture information after image decomposition is reconstructed, which can effectively improve the speed of image reconstruction and improve the stability and accuracy of the reconstructed image.

4 Conclusions

This paper improves the original learning-based super-resolution reconstruction method and effectively compensate for the original method's time-consuming and fuzzy edges of the reconstructed image. At the same time, it not only improves the image quality, but also increased the speed of reconstruction.

However, the effectiveness and real-time nature of reconstruction methods have always been the focus of research on the super-resolution reconstruction. Under the same conditions of training images, the larger the amount of data in the training set after block partitioning, the better the reconstruction effect will be, but this time, the image blocks in the matching and fusion process takes longer, leading to slower reconstruction speed; conversely, if the amount of data in the training set is small, the reconstruction speed will increase, and the reconstruction effect will decrease. Therefore, how to establish an effective training set, that can guarantee image quality, which also can increase the reconstruction speed, remains to be further studied.

Acknowledgements. This work is supported in part by the National Natural Science Foundation China (61601174), in part by the Postdoctoral Research Foundation of Heilongjiang Province (LBH-Q17150), in part by the Science and Technology Innovative Research Team in Higher Educational Institutions of Heilongjiang Province (No. 2012TD007), in part by the Fundamental Research Funds for the Heilongjiang Provincial Universities (KJCXZD201703), and in part by the Science Foundation of Heilongjiang Province of China (F2018026).

References

1. Yan, Z., Kun, X.U., Yong, L.I.: Remote sensing image super-resolution based on POCS and out-of-core. *J. Tsinghua Univ.* **50**(10), 1743–341 (2010)
2. Kim, K.I., Kwon, Y.: Single-image super-resolution using sparse regression and natural image prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(6), 1127 (2010)
3. Shou, Z.Y., Liao, M.L., Zhang, T.: Improved image super-resolution via sparse representation based on IBP. *Comput. Eng. Des.* (2014)
4. Freeman, W.T., Pasztor, E.C.: Learning low-level vision. *Int. J. Comput. Vis.* **40**(1), 25–47 (2000)
5. Yang, J., Wright, J., Huang, T., et al.: Image super-resolution as a sparse representation of raw image patches. In: *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2008*, vol. 2008, pp. 1–8. IEEE (2008)
6. Tang, Y., Yan, P., Li, X.: Single-image super-resolution via local learning. *Int. J. Mach. Learn. Cybernet.* **2**(1), 15–23 (2011)
7. Freeman, W.T., Jones, T.R., Pasztor, E.C.: *Example-Based Super-Resolution*. IEEE Computer Society Press (2002)



Construction of the Intelligent Small-Sized Plant

Yadi He^(✉), Yuanyuan Wang, Pengli Zhu, Shuangshuang Zhang,
Feichao Zhao, and Lidong Zhang

Northwest Institute of Mechanical and Electrical Engineering, Xianyang 712099,
Shanxi, China
sunspota@qq.com

Abstract. As intelligence is popularized in the life of people and technology develops continuously, intelligent plant become the main development direction of the modern enterprises. By taking intelligent plant of small-sized traditional factories as the case, the paper explains the conception of intelligent plant, shows the necessity of intelligent plant and establishes the systematic frame of an intelligent plant. A small-sized factory is constructed intelligently based on analysis of the systematic frame of the intelligent factory; a large database is established through an MES system, and various information flows are settled, so that intelligence of production management is realized, intelligent productions are established, an intelligent logistics system develops, and then intelligent improvement exploration of small-sized plant can be finished.

Keywords: Intellectualized · Intelligent plant · MES · Database

1 Introduction

In this century, with the development of science and technology, more and more new things have appeared in people's daily life. The Internet of Things, big data, and smart power grids have linked people together through various networks to build up broad environment of intelligent modern production and life. Artificial intelligence enters the era 2.0. Our life is glutted with the Internet of Things and large data. Traditional basic factories should continue to develop, and positively respond to the pulse of times and technology to enter the era of intelligent factories.

When mentioned intelligence, many products now include the concept of intelligence, e.g., smart homes, alpha dogs, robot arms, as small as light bulbs, and as large as the Internet, our life is fully glutted with intelligence. What is intelligence, it is still a problem, it seems that intelligence is a kind of mechanical behavior, not only mechanical movement, but also being full of certain thinking, For example, sweeping robots that can help people sweep the floor, welding robots that can weld themselves, or large computers that take on core operations. Technological developments have enabled rigid machinery to have a flexible thinking ability and help human in daily life. These are all examples of entities, and for further development, through an operating

mode that the Internet of Things, the big data and the like are connected by different types of networks, so that the Internet of Things and the big data generate thought fluctuations in operation to build a complex network system, so that an intelligent operating pattern is generated, which is like that the human brain controls every nerve thread.

The development of intelligence is successful and has epoch-making significance. Under the operation mode of intelligence, complex system construction and industrial processes that could not be completed before can be realized, which makes production more efficient and convenient, and also vigorously promotes the development of a new social model.

On the life and the networks the development of intelligence is quite rapid, but the traditional factory, which is the foundation of society, has developed slowly. In order to pursue higher efficiency and production capacity, the traditional factories must be intellectualized. In this paper, through the analysis of the concept and system of intelligent plant, the small-sized factory intelligence transformation is carried out, and the advantages and disadvantages of the current stage of intelligent plant transformation are preliminarily explored. It will provide some reference for future development.

2 Intelligent Plant

2.1 Conception of Intelligent Plant

Intelligent plant mainly constructs intelligent production system, intelligent distribution equipment, and intelligent logistics system to realize intelligent production process. This intelligent system includes physics, information, data, and logistics.

The core of intelligent plant construction is “intelligence”, and the frontier technology of various disciplines. “Intelligent” is the core of smart factories. What is “intelligent”? “Intelligent” means that the factory has sensitive and accurate perception ability, correct thinking judgment ability and effective execution method.

From a scientific point of view [1], intelligent plant is defined as follows: the production of each link of the factory, in some way integration, through the information system to simulate the production management expert intelligent activity, analyzed the problems of the production and business operation, judgment, reasoning, planning and decision-making, replace or extend the production and business operation environment in part of the mental work and intelligence activities of human expert to collect, store, perfect, sharing, inheriting and developing. Realize the transformation from human intelligence activity to information system intelligence activity.

From the point of view of engineering practice: intelligent plant dominated by information technology, combined with other related disciplines technology, realize factory production operation, production management, management decision three levels all closed-loop management of business process, and achieved the integration of the whole factory business process and business operation decision-making, implementation of automation.

2.2 The Functional Structure of Intelligent Plant

Intelligent plant can be divided into three levels, business intelligence, intelligent production management, making intelligent, these three levels can corresponding to the enterprise information integration model of ERP, MES, and PCS. These three levels influence each other and restraining each other, and promote each other. Intelligent the construction of plant is aimed at the three levels of intelligent building. The intelligent construction of technology also further enhances the information construction of the factory, making the three relations promote and develop mutually under the cohesion of MES (Fig. 1).

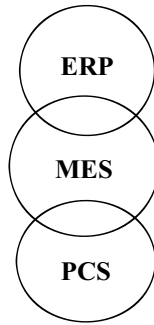


Fig. 1. The three layers of intelligence

In factories Past by raising the ERP to improve its management level and enterprise competitiveness [2]. However, this method is easy to cause the gap between ERP and PCS to increase more. The MES process plan management system in the upper and bottom control system, promote the building of MES system, effective coordination of the relationship between them, to build a good solid business model system [3].

Traditional enterprises have strengthened the construction of ERP, MES, and PCS, promote the application of a level and deepen, but also effectively carry out the layers of comprehensive construction, has completed each link of the data and information collection, arrangement work. But in order to realize intelligent plant construction but also to the level that the data collecting, sorting, information fusion, and finish the ERP, MES—PCS, PCS—MES and ERP information transmission, two lines scattered data information, collection into big data process, complete intelligent enterprise information flow in closed loop, makes the summary data to form large data form, embodies intelligence concept.

Figure 2 is an intelligent factory architecture that combines the entire production system. In this diagram, wisdom factory implements digital product design, digital manufacturing, digital management processes and business processes, as well as the comprehensive integrated optimization process, can use the engineering, manufacturing, supply chain intelligent plant model of three dimension description.

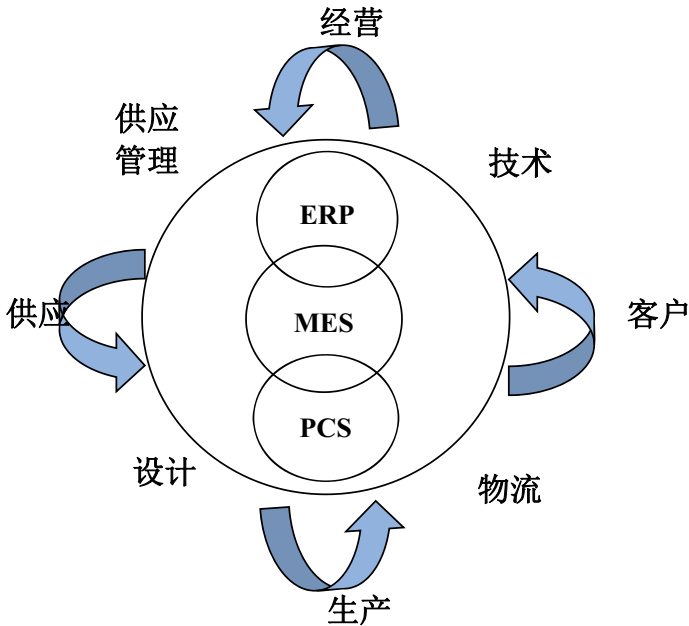


Fig. 2. The architecture of intelligent plant

2.3 Analysis of the Intellectualized Small Plant

Through the architecture of intelligent plant, this paper analyzes the intelligent construction method of small plant.

With the instance of domestic all kinds of intelligent plant construction, the intelligent information construction of the small processing plants, mainly to MES construction as the main body, and as a basis, the coordination of ERP and PCS, at the same time to establish intelligent processing of intelligent production line and logistics construction.

3 Example of Construction

In this paper, the factory that needs to be intellectualized is a traditional manufacturing midget factory. It is small in scale, has a variety of products, and is rich in orders. However, because of the small number of processing equipment and the uneven busy schedule throughout the year, orders are often overbooked and have a negative impact. Because of the lower productivity, efficiency has not always improved.

Through intelligence system construction, we should analysis the problems need to be solved from top-down. First the problem of the whole processing management network, the second problem is layout of new intelligent production line, finally establish intelligent logistics network, complete the intelligent construction of the factory.

3.1 MES

Introduces MES system for construction of intelligent plant.

MES [4] is Manufacturing Execution system, this system is a manufacturing production management ideas and management tool, that fully integrates manufacturing resources production, quality, equipment and personnel performance.

The main functions of MES include production scheduling, resource allocation, man-hour management, document control, data acquisition, quality management, equipment management, etc. it is a bridge connecting ERP and PCS. On the one hand, MES provides timely data and information for the ERP system. On the other hand, through the bottom data collection and analysis of PCS, it provides basis and guarantee for optimizing and controlling the operation of the production line.

The entire system structure of MES system is shown in the Fig. 3. There are five basic modules in this system, including database, technology, OA, project, and inspection. Database module contains all the data content in the whole production process. Technology module contains the processing technology and product design. OA system is the whole MES system coordination office. Project module enables the system to plan and quota normal operation. Inspection module to monitor the processing and the survival rate of the product.

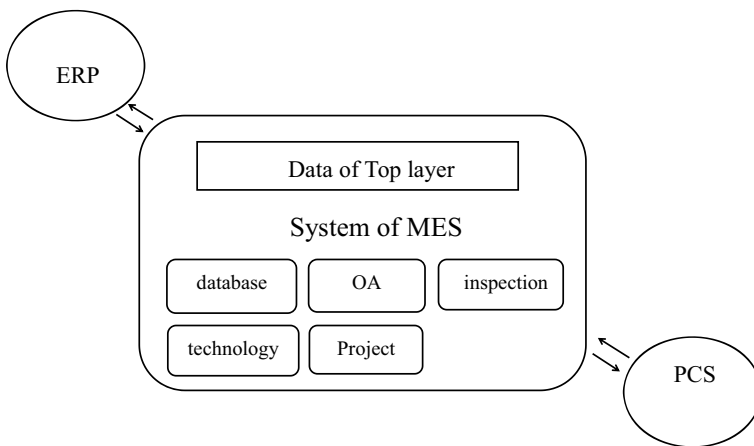


Fig. 3. System of MES

The MES system connects with the top-level data and management plan, interacts with ERP, and completes a complete MES production system by interacting with the production control PCS through five basic modules.

3.2 Database Construction

Establishing database of MES, the analysis of the need to build a factory organization structure, the decomposition process scheduling, material management, quality

management processes, the labor process of auxiliary equipment management, equipment management process, collect all kinds of data information, build a large database. Large database is the basis of MES intelligent analysis.

In the database construction, it should be as comprehensive as possible to ensure the correctness of MES intelligent analysis (Fig. 4).

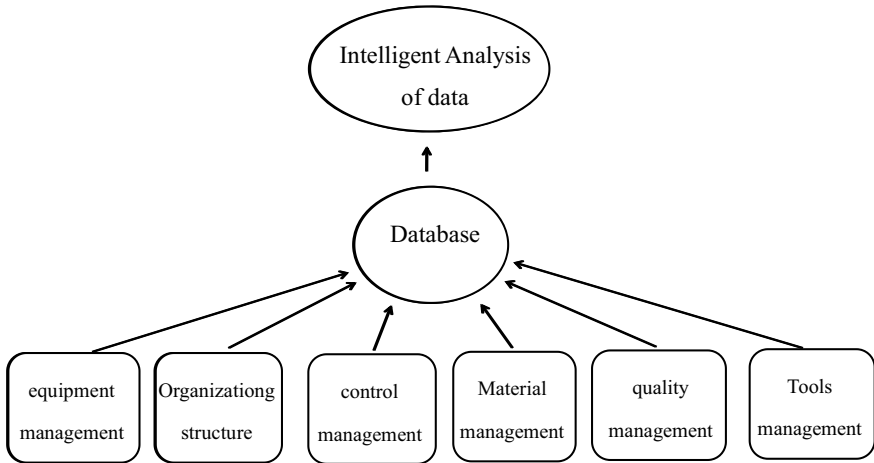


Fig. 4. Database construction

3.3 Intellectualized Production Management

The purpose of using MES in the target construction of production management is to make the factory realize intelligent management.

This on the basis of the construction management target of the main research object is based on the operation and management of monthly production plan, according to the processing condition of the factory, production capacity short-term operation plan, monthly plan can be divided into short-term assignments are usually for days, weeks work plan, and coordinate factory scheduling completed the requirements [5].

After MES based intelligent production management, the plan of the layer to adjust according to the actual production situation, fully reflect the actual production situation, timely feedback the problems arising from the operation management, and through the management to adjust work plans, implement intelligent system of production management level.

Shown in Fig. 5, the intelligent simulation combined with monthly plan according to actual condition, the use of MES system data intelligent simulation, simulation results after the combination of knowledge base, correlation analysis and give a short-term operation plan, send the management process, management to confirm the results, modified short-term operation plan, and short-term operation plan in actual application.

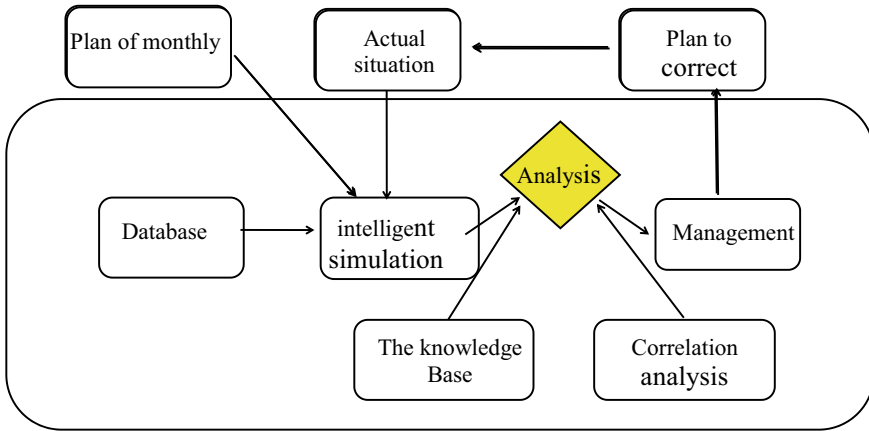


Fig. 5. The intelligent architecture of manufacturing management level

3.4 Conception of Intelligent Product Line

Intelligent production line construction based on industry 4.0. If the MES is the brain of a smart factory, intelligent production line is the heart of the smart factory, with the wisdom of MES, intelligent production line of strong operational capacity to achieve a truly intelligent plant [6].

According to the small factory in this paper, the processing product types of this factory are analyzed, and several simple intelligent production lines are preliminarily designed, which are axle processing, box processing, and assembly intelligent production lines respectively. When applied in the intelligent production line should have automatic extraction program, fault alarm, automatic detection, flow process, key points, such as automatic logistics, elastic interval set according to the key process and equipment layout, to ensure production line smart plan is completed (Fig. 6).

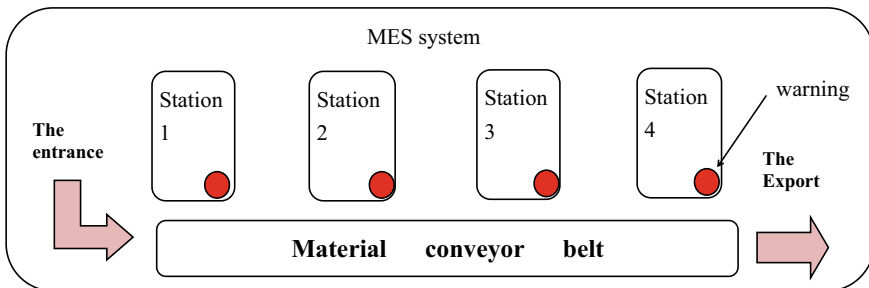


Fig. 6. The architecture of intelligent plant

3.5 Conception of Intelligent Logistics

Logistics has moved from simple goods flowing to the Internet of things [7]. The Internet of things is increasing innovation and intelligence, and then changed our way of life. The development of the Internet of things has also become a practical and effective way for enterprises to effectively reduce costs and increase efficiency, improve efficiency and reduce waste. The construction of intelligent plant cannot be separated from the construction of intelligent logistics.

This case of small-sized plants need to undertake the construction of logistics is divided into two main classes, a class for the outer zone, outside the district logistics is divided into two lines, the first for the supplier to the factory to provide spare parts, accessories, etc., the second production after the completion of the transportation to the customer for the factory, construction for the second category of logistics area, the main flow of logistics system for parts.

According to the above analysis, intelligent logistics system construction is carried out for external supply logistics and internal circulation logistics [8].

For external supply logistics lines, supplier logistics route analysis, combined with the product manufacturing cycle, to offer stable supplier logistics information platform, according to the shortage of the beat to communicate in advance, according to the requirements of manufacturing tempo send orders in manufacturing cycle, make the supply of raw materials, parts, accessories do not take up the inventory, and can guarantee supply.

Logistics for internal circulation, in view of the difficult to change the factory construction, each factory content reorganize, cooperate on downstream processing order, will undertake work position reasonable arrangement, and the processing route planning into MES construction, set up professional logistics, and logistics in the car with the intelligent data acquisition, real-time query the car number put in storage, such as working status information. Organic coordination among processes to ensure the fastest mobilization of internal logistics within the inventory capacity (Fig. 7).

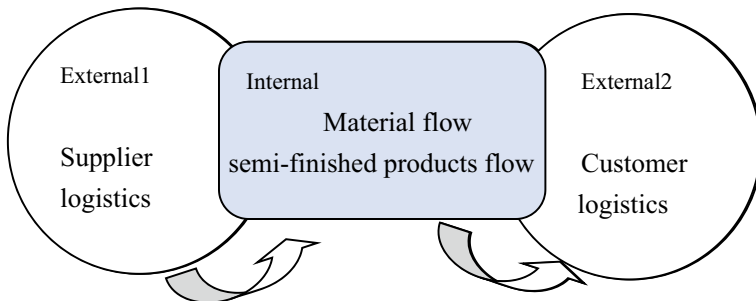


Fig. 7. The flow of logistics

4 Conclusion

- (1) In the modern era when intelligent development is rapid, the development of smart factories is urgent;
- (2) Changing the inherent development model of the factory is a very difficult thing. This smart factory construction is a positive attempt to change the original operating mode and improve efficiency;
- (3) The development of smart factories should fully mobilize the joint development of management, production lines and logistics;
- (4) In China, the development of production management intelligence gradually formed a development model with mes as the main body, but the construction of smart production lines and smart logistics should also be diversified;
- (5) The development path of smart factories should also be explored in many ways to vigorously develop the automation, informatization, and intelligence of modern factories.

References

1. Mao-sen, X.: Research on intelligent plant construction technology of process industry. *Inf. Technol. Informatiz.* **6**, 47–52 (2013)
2. Feng, L.: MES of intelligent plant key technology research. *Manuf. Autom.* 32017.04
3. Jian-chu, Y.: *Intelligent. Comput. Integr. Manuf. Syst.* **6**(5) (2000)
4. MESA. International, MES explained; a high level vision
5. Zhi-xin, W.: *Manufacturing system and application of MES* (2006)
6. Dian, L.: Research and application of intelligent gear production line management system based on FANUC FOCAS. *Manuf. Technol. Mach. Tool* 2017.05, 166–168
7. Shu-hua, D.: Intelligent logistics system construction and solution. *Inf. Res.* 2012.5, 6–15
8. Lei, Y.: Research on the construction of intelligent logistics system. *Chem. Eng. Des. Commun.* 2016(10)



Integrating Heterogeneous Datasets by Using Multimodal Deep Learning

Fariba Khoshghalbvash^(✉) and Jean X. Gao

Department of Computer Science and Engineering,
University of Texas at Arlington, Arlington, TX, USA
fariba.khoshghalbvash@mavs.uta.edu

Abstract. Rapid collection of data sources, varying in volume and structure poses a challenge for scientists to establish a practical approach to manipulating heterogeneous data sources. A multimodal learning and an integrated analysis make it possible to extract much worthwhile information from a collection of multiple simple raw data. Therefore, data integration can lead to a more reliable and robust result. High-throughput sequencing technologies, especially next-generation sequencing, leave us with multi-platform genomic data such as gene expression, SNP, CNV, DNA methylation, and miRNA expression. In this paper, we represented a multimodal deep neural network to exploit the mutual information between three different modalities to classify breast cancer patients into two groups based on their survival rate. Experimental results indicate that our method improves the classification accuracy and performs better on imbalanced data compared to the other single-modal state-of-the-art methods.

Keywords: Data integration · Omics · Deep learning

1 Introduction

Nowadays high-throughput technologies such as microarrays [1] and array Comparative Genomic Hybridization (aCGH) [2] make available to scientists an extensive collection of different types of omics data such as Gene Expression (GE), Single-Nucleotide Polymorphism (SNP), Copy Number Variation (CNV), protein–protein interaction, DNA methylation, micro RNA, messenger RNA, noncoding RNA, etc. With these advances in data collection, scientists are faced with a plethora of heterogeneous omics datasets.

It has been demonstrated that one can get a more reliable and robust result by gathering all the inter-modality information [3]. Data integration was defined as combining data from different sources and representing them in a unified view [4]. However, due to the heterogeneity, the massive scale of data, and the presence of uncertainty, errors, and missing values in the data integration has become an arduous task. Previously, methods like Bayesian Networks [5] and Kernel-based

methods [6, 7] have been applied on multi-platform genomic datasets. However, for capturing the correlation between different data modalities with high dimensions, direct application of classical machine learning algorithms does not give satisfactory results. Hence, specialized machine learning methods such as deep learning are required for the analysis of such high-dimensional multimodal data with higher degrees of accuracy.

As opposed to the old-fashioned methods that directly train the model from the original data and therefore are not able to aggregate distinct forms of inputs, it has been proved that deep learning can be used for exploiting data integration as it transforms each modality to an abstract representation [3]. In this paper, we implemented a discriminative multimodal deep neural network to integrate two different types of genomic data and clinical data to categorize breast cancer patients into two groups based on their survival rate. The result of a stratified fivefold cross-validation indicates that our model surpasses three state-of-the-art methods according to accuracy, precision, and recall scores. We also demonstrate how deep learning can cope with the unbalanced datasets.

2 Methodology

In this paper, a Deep Neural Network (DNN) architecture is used to create a multimodal network in order to integrate and exploit the information carried by three different data types at the same time. Considering the purpose of this study which is to integrate different data modalities and obtain the inter-modality information, DNN is a suitable tool to reach the abstract representation within each modality and then combine them to achieve a comprehensive result.

Previously, deep learning based architectures have been applied to join different modalities. Liang et al. [8] applied used a deep belief network for unsupervised clustering. Sun et al. [9] applied an ensemble learning and integrated the result of different networks for each modality. However, to our best knowledge, there has not been a supervised learning model that does the training part using the whole network from the input to the output. Even in the cases that the integration takes place in the middle stage, first, each modality is pretrained individually and then more layers are added on top of the trained modality-specific layers to finalize the classification task [10]. On the contrary, our Multimodal Deep Neural Network (MDNN) integrates the inter-modality information while training the parameters of the model.

Using a concise notation, the values of the nodes in each layer can be calculated by the nodes in the previous layer:

$$h^{(i)} = \sigma(W_{m_j}^{(i)} \times h_{m_j}^{(i-1)} + b_{m_j}^{(i)}) \quad (1)$$

where $W_{m_j}^{(i)}$ is the weight matrix between i th and $(i-1)$ th hidden layer, σ is the ReLU activation function, and $b_{m_j}^{(i)}$ is the bias for i th hidden layer for modality m_j .

In our model, the fusion layer is calculated by joining the last modality-specific hidden layers:

$$h^{(f)} = \sigma(W_{m_1}^{(f)} \times h_{m_1}^{(f-1)} + W_{m_2}^{(f)} \times h_{m_2}^{(f-1)} + W_{m_3}^{(f)} \times h_{m_3}^{(f-1)} + b^{(f)}) \quad (2)$$

Consecutively, the output layer that is counted as predicted labels by our model is computed as below:

$$output = \sigma'(W^{(o)} \times h^{(f)} + b^{(o)}) \quad (3)$$

where $W^{(o)}$ is the weight matrix between output layer and fusion layer, σ' is the Sigmoid activation function, and $b^{(o)}$ is the bias for the output layer.

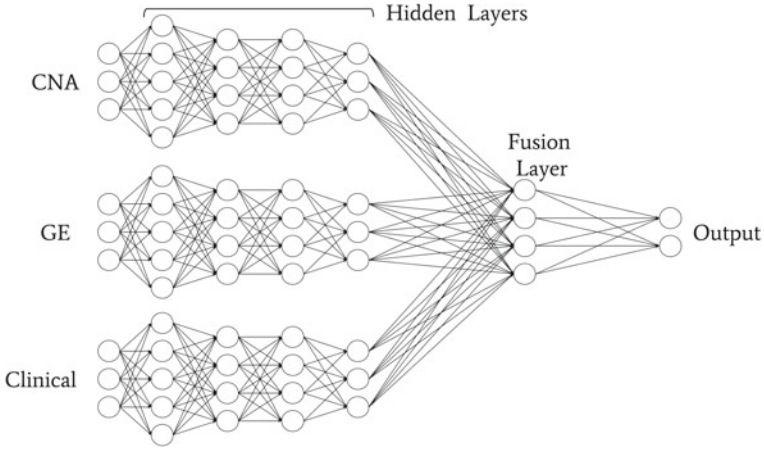


Fig. 1. An illustration of our integrative multimodal deep architecture

Table 1. Summary of the parameters of our model

# of hidden layers per modality	4
# of shared layers	1
# of hidden nodes per modality	[1000, 500, 500, 100]
# of hidden nodes in the shared layer	200
# of classes	2
Batch size	128
train\test split	0.2
train\validation split	0.2
Loss function	Binary crossentropy
Optimizer	Adadelta with initial lr = 0.1
# of epochs	100
Output layer activation function	Sigmoid
Other layers activation function	ReLU

We build a network containing four modality-specific hidden layers for each modality with 1000, 500, 500, and 100 nodes, one fusion layer with 200 nodes (A schematic illustration is shown in Fig. 1). The detailed parameters of our model are described in Table 1.

3 Data

In this study, we use three different modalities of METABRIC breast cancer dataset for 1980 that was preprocessed by Sun et al. [9] and categorized into two groups of short-term (491 patients that survived less than 5 years) and long-term (1489 patients that survived more than 5 years).

Moreover, in order to diminish the dimensionality prior to learning a model, Sun et al. [9] applied mRMR [11] on Clinical (Cli), Gene Expression (GE), and Copy Number Alternation (CNA) to reduce the number of original features. The final dataset which has been used in this paper consists of 25 features for clinical data such as age at diagnosis, size, histological type, menopausal status inferred, lymph nodes positive, stage, and mutation status, 400 features for gene expression, and 200 features for CNA. Furthermore, we normalize all of the features in these three modalities by min-max normalization and set them between 0 and 1.

4 Result

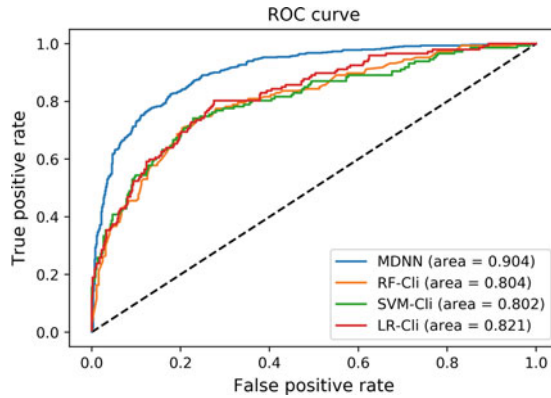
To comprehensively evaluate the performance of our model, a stratified fivefold cross-validation was applied to the aforementioned data set which left us with 396 test samples within each iteration. Following that, each training fold is divided into two groups of train (%80) and validation (%20) data.

We compare the performance of MDNN with three long-established methods: Support Vector Machines (SVM), Random Forest (RF), and Logistic Regression (LR) by calculating accuracy, precision, and recall for each model. Since the last three methods do not fulfill multimodal learning, the scores are computed for each individual modality separately.

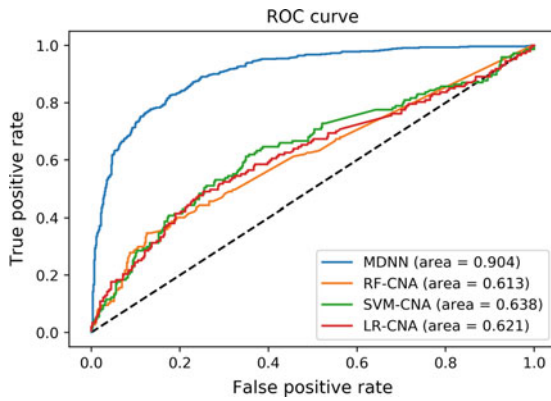
As expected, among these four learning models, MDNN outperforms the other three as it captures the shared representations between different modalities (Table 2). Moreover, Fig. 2 illustrates the ROC curves for different methods using

Table 2. Comparison of Accuracy, Precision, Recall, and F1-Score by different methods

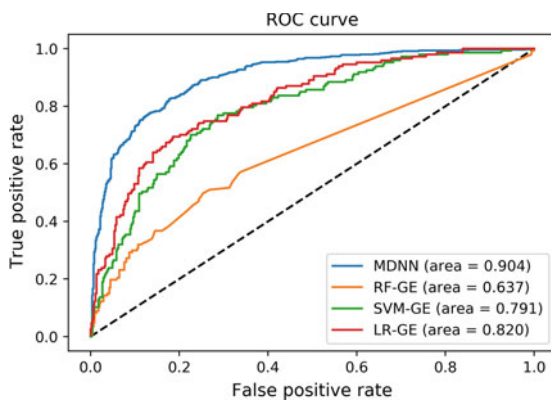
	MDNN	SVM	RF	LR	SVM	RF	LR	SVM	RF	LR
		CNA	CNA	CNA	GE	GE	GE	Cli	Cli	Cli
Accuracy	0.83	0.76	0.75	0.75	0.79	0.75	0.80	0.81	0.80	0.80
Precision	0.83	0.72	0.66	0.70	0.78	0.57	0.78	0.80	0.79	0.79
Recall	0.83	0.76	0.75	0.75	0.79	0.75	0.80	0.81	0.80	0.81



(a) MDNN vs. SVM, RF, LR for clinical data



(b) MDNN vs. SVM, RF, LR for CNA data



(c) MDNN vs. SVM, RF, LR for GE data

Fig. 2. The ROC curves of MDNN, SVM, RF, LR

different inputs. While the AUC value for MDNN is 0.904, it is much lower for other single-modal learning methods.

It may seem that using SVM for clinical data will lead to a significant result. However, if we further scrutinize the performance of these models, Table 3 indicates that MDNN best copes with unbalanced data as opposed to SVM, RF, and LR which mostly intend to vote for the group with the higher population.

Table 3. Comparison of number of samples of each class in the original test data and predicted labels by different methods

	Original	MDNN	SVM	RF	LR	SVM	RF	LR	SVM	RF	LR
			CNA	CNA	CNA	GE	GE	GE	Cli	Cli	Cli
Long-term patients	298	302	385	395	367	330	396	329	348	370	328
Short-term patients	98	94	11	1	29	66	0	67	48	26	68

5 Discussions and Conclusions

In order to integrate the latent information in distinct data modalities and exploit the shared representation for classifying breast cancer patients into two groups of short-term and long-term based on their survival rate, this paper proposes a discriminative multimodal deep architecture. This deep architecture transforms each input modality to an abstract representation which is suitable for integration. As opposed to previous work in which the learning of each modality was done individually, and the integration took place at a late stage, the proposed architecture learns the parameters of the whole model simultaneously.

It has been demonstrated that compared to other methods, the aforementioned approach overcomes three state-of-the-art methods (SVM, RF, and LR) according to the average accuracy, precision, and recall scores resulting from fivefold cross-validation. Moreover, by taking the performance of each model under scrutiny, it can be concluded that the proposed method can also cope with unbalanced data while the other three fails to do so.

For future work, we would like to apply multimodal learning to categorize unlabeled data or utilize a semi-supervised approach and build a hybrid discriminative/generative model.

References

1. Schena, M., Shalon, D., Davis, R.W., Brown, P.O.: Quantitative monitoring of gene expression patterns with a complementary dna microarray. *Science* **270**(5235), 467 (1995)
2. Pinkel, D., Albertson, D.G.: Comparative genomic hybridization. *Annu. Rev. Genomics Hum. Genet.* **6**, 331–354 (2005)

3. Srivastava, N., Salakhutdinov, R.R.: Multimodal learning with deep boltzmann machines. In: *Advances in Neural Information Processing Systems*, pp. 2222–2230 (2012)
4. Lenzerini, M.: Data integration: a theoretical perspective. In: *Proceedings of the Twenty-First ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, pp. 233–246. ACM
5. Jansen, R., Yu, H., Greenbaum, D., Kluger, Y., Krogan, N.J., Chung, S., Emili, A., Snyder, M., Greenblatt, J.F., Gerstein, M.: A bayesian networks approach for predicting protein-protein interactions from genomic data. *Science* **302**(5644), 449–453 (2003)
6. Lanckriet, G.R., De Bie, T., Cristianini, N., Jordan, M.I., Noble, W.S.: A statistical framework for genomic data fusion. *Bioinformatics* **20**(16), 2626–2635 (2004)
7. Yamanishi, Y., Vert, J.-P., Kanehisa, M.: Protein network inference from multiple genomic data: a supervised approach. *Bioinformatics* **20**(suppl_1), i363–i370 (2004)
8. Liang, M., Li, Z., Chen, T., Zeng, J.: Integrative data analysis of multi-platform cancer data with a multimodal deep learning approach. *IEEE/ACM Trans. Comput. Biol. Bioinform. (TCBB)* **12**(4), 928–937 (2015)
9. Sun, D., Wang, M., Li, A.: A multimodal deep neural network for human breast cancer prognosis prediction by integrating multi-dimensional data. *IEEE/ACM Trans. Comput. Biol. Bioinform.*
10. Pan, X., Shen, H.-B.: Rna-protein binding motifs mining with a new hybrid deep learning based cross-domain knowledge integration approach. *BMC Bioinform.* **18**(1), 136 (2017)
11. Peng, H., Long, F., Ding, C.: Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(8), 1226–1238 (2005)



A Multi-view Deep Learning Approach for Detecting Threats on 3D Human Body

Zhicong Yan¹, Shuai Feng¹, Fangqi Li¹, Zhengwu Xu², and Shenghong Li¹(✉)

¹ School of Cyber Security, Shanghai JiaoTong University, Shanghai 200240, China
{zhicongy, feng_shuai, solour_lfq, shli}@sjtu.edu.cn

² School of Information and Communication Engineering, University of Electronic
Science and Technology of China, Chengdu 611731, China
zwxu@uestc.edu.cn

Abstract. Deep Neural Network-based methods have recently shown an outstanding performance on object detection tasks in 2D scenarios. But many tasks in real world requires object detection in 3D space. In order to narrow this gap, we investigate the task of detection and localization in 3D human body in this paper, and propose a multi-view-based deep learning approach to solve this issue. The experiments show that the proposed approach can effectively detect and locate specific stuff in 3D human body with high accuracy.

Keywords: Multi-view convolution neural network ·
3D object detection · Airport security

1 Introduction

Nowadays, airport security is an indispensable requirement for safe travel, which consumes a great number of manpower, forces the passengers to wait in the long lines and waste time. Quickly and accurately detecting suspicious stuff camouflaged in human body or luggage via machine learning algorithm is one promising solution.

During the past two decades, most countries have attached great importance to the security of transport infrastructure. Scanner machines with new technology and programs are invented to enforce protection in the airports. According to the reports of Transportation Security Administration (TSA), tens of thousands of High- Definition Advanced Image Technology (HD-AIT) system had been deployed in USA airports [1], which can generate 3D millimeter wave scan data of about 2 gb per person. In the practical applications in airports, we have to predict whether any suspicious stuff is present in the human body and which part of the human body contains them. This problem can be solved as a multi-label binary classification task, while the difficulty with this problem is taking

such high-dimensional data as input. For 3D model, the dimension of input data is almost three or four orders of magnitude more than the 2D image data, posing great challenges to design the algorithm.

The convolution neural network (CNN) could efficiently classify 2D image data because it is locally connected with weight sharing structure and highly nonlinear model function [2]. But when it comes to 3D data with much higher dimension, we need to take more consideration. In 3D scenario, the classification task has been investigated by a number of works [3–5]. The state-of-the-art method MV-CNN [3] applies the multi-view strategy, which reduces the input dimension by taking multiple 2D views of 3D model as input. But when we applying the MV-CNN to the passenger screening task, we find that the network is confused by the high-dimensional 3D data because the stuff we want to detect only takes a small part in 3D image. So, we resort to efficient detection methods in 3D scenario.

For object detection and localization in 2D scenarios, He Kaiming et al. has proposed Faster R-CNN algorithm [8], which has been used extensively and investigated by various research papers [8–10]. But there is little research in the field of object detection and localization in 3D scenarios.

For the above reasons, in this paper, we propose a network architecture that can detect and locate suspicious stuff in 3D human body, which adopts the multi-view strategy in MV-CNN [3] as well. We incorporate the detection strategy to the network to make it capable of detection and localization in 3D space. Further, in order to train this network, we split the network into two stages. In the first stage, a pretrained deep CNN is used to extract score maps from images by different views of 3D human body, predicting score maps at each view, which only contain the position information of suspicious stuff. Then in the second stage, we aggregate those maps into a single feature and use a regularized Multilayer Perceptron (MLP) network for detection. This two-stage architecture ensure that only the interested objects are detected by the network. The contributions of this paper are as follows:

1. First, we describe a two-stage approach that allows the network to perform detection in 3D space.
2. Second, we validate our approach on the Passenger Screening Dataset, and the result shows that the proposed approach can efficiently detect and locate suspicious stuff in 3D human body.

The rest of this paper is organized as follows: In the Sect. 2, we provide the related research work. Section 3 describes the proposed approach in detail. In Sects. 4 and 5 we provide the experiment results and conclusion of this paper.

2 Previous Work and Literature

Our method is related to prior work on image-based CNNs and object detection methods. Next, we discuss representative work in these areas.

2.1 Convolution Neural Networks

Our work has employed the convolution neural networks, which have recently enjoyed a great success in large-scale image and video recognition. In particular, CNNs have been used as a general-purpose image feature extractors for lots of tasks, such as object classification, object detection and semantic segmentation [2, 11, 12, 15]. Without handcrafted features, CNN is easy to train with backpropagation algorithm. The Resnet [11] network and VGG [12] network structure is now the most popular CNN structures in image classification task. We choose the VGG16 network and Resnet-101 network as the base network because they offer a balance between computation and network generalization capability.

2.2 Deep Network for Object Detection

With the development in both architecture and train methods of deep CNNs, object detectors like OverFeat [13] and R-CNN [6] showed dramatic improvements in accuracy compared with traditional methods. OverFeat simply applying CNN as a sliding window function on an image pyramid. and R-CNN introduced a region proposal-based strategy, where only the interested regions can be kept and scale-normalized before classifying with a CNN, which greatly improved compute efficiency. the recent proposed Fast R-CNN [7] and Faster R-CNN [8] methods achieved more accurate result and compute efficiency, In Faster R-CNN, the selective search of interested region is altered to a region proposal network, in which a objectness score map was generated by adding an extra convolutional layer to the base network and the region proposal box was generated by this score map. In this research, we adopt this score map to help us locate the suspicious stuff in image.

2.3 View-Based 3D Shape Recognition

There has been existing work on recognizing 3D shape with CNNs, some methods resort to CNNs to extract features from multiple 2D representations [3, 4], for there are a number of advantages of using 2D representation. One main reason is the trade-off between resolution and compute efficiency [3], for example, a $16 \times 16 \times 16$ grid of 3D voxels input shares the same input size with a more refined 64×64 2D image input, so the multi-view CNN can capture more details. In addition, there are many other ideas. Charles R.Qi et al. proposed the PointNet to directly process the original point cloud data and classify the model data using an RNN-like network [5]. The slice-based CNN method proposed by F.Gomez-Donoso takes slices of the 3D models as the CNN input, which alleviates the resolution trade-off [14]. In this research, we adopt the multi-view-based 3D shape descriptor and process each view of the models individually.

3 The Proposed Approach

In this section, we present the approach of this paper. In order to predict which part of the body contains suspicious stuff, we propose a two-stage algorithm. In

the first stage, a pretrained deep CNN is used to extract score maps from images by different views of 3D human body, predicting objectness score maps at each view. Then in the second stage, we aggregate those maps into a single feature and use a regularized MLP network for detection. The complete architecture is depicted in Fig. 1.

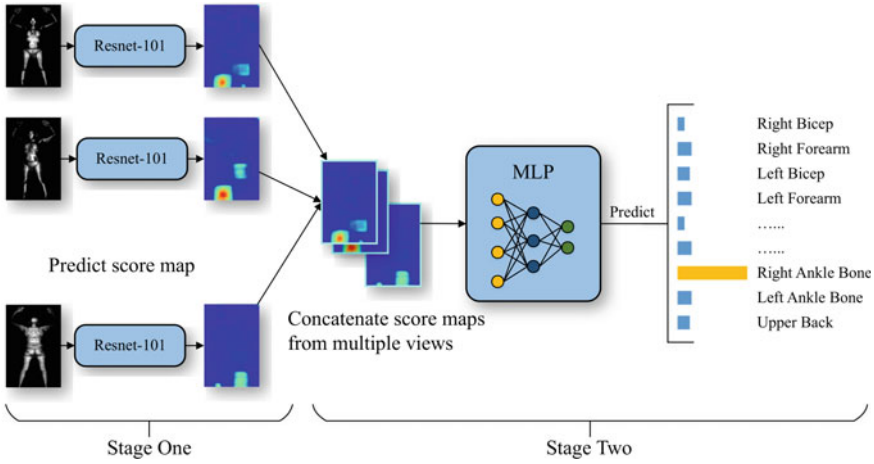


Fig. 1. The proposed two stage approach for object detection in 3D images.

3.1 Predict Score Map

We employ the VGG16 [12] network and Resnet-101 [11] network as the base network to predict a score map from each image input, which is in lower resolution than the input image. As in the original Faster R-CNN [8], the top fully connected layer in base network are removed, then RPN network is concatenated behind the last convolution layer, in which the score map prediction and bounding box regression are performed. In this task, we are not interested in the exact position of the suspicious stuff, so we remove the bounding box regression branch in RPN network.

For training RPNs, Faster R-CNN adopted anchor-based method in which every spatial location of score map was assigned an anchor box and classify each anchor box to positive or negative based on their overlap with ground truth box. But actually in the experiments, we find this method is inclined to produce false alarm. In this scene, we utilize CNN as probability regression function with the regression target of each position is defined as

$$p(p|G) = \max_{g_i \in G} \begin{cases} 1 - \left(\frac{x_p - cx_i}{w_i}\right)^2 \left(\frac{y_p - cy_i}{h_i}\right)^2, & \text{if } p \text{ in } g_i \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

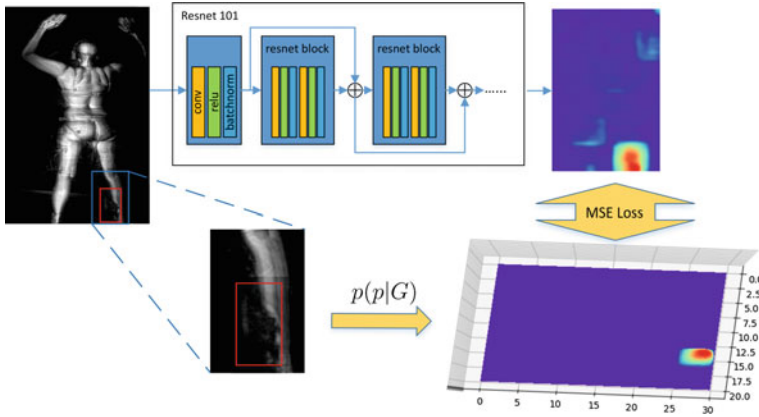


Fig. 2. In the first stage, we employ a single network to predict the score map of each view of images. we append a 1×1 convolution layer to base network, and a sigmoid activation layer to clamp the output value within $(0, 1)$.

Here, G is the set of ground truth boxes and g_i is the i -th ground truth box. x_p and y_p is the pixel position in score map w.r.t the input image. In the train phase, we apply the Mean Square Error (MSE) loss and Adam optimizer with learning rate of 10^{-4} at initial. The scheme of this stage is depicted in Fig. 2.

3.2 Detect with Multi-view Representations

In the first stage, the score maps contain the position information of suspicious stuff in multi-view representations, we claim that those information is rich enough for us to locate the suspicious stuff in human body. In the second stage, we take score maps generated from multi-view representation of a single sample as input, and introduce a MLP network with softmax activation to produce the probabilities of each body zones. In experiments, we introduce lasso regulations to reduce overfitting. So the target loss function is defined as

$$L(x, l) = \frac{1}{NM} \sum_{i=1}^N \sum_{z=1}^M [l_i^z \log f_z(x_i, \omega) + (1 - l_i^z) \log(1 - f_z(x_i, \omega))] + \lambda \|\omega\|^2 \quad (2)$$

where l_i^z denotes the label of the z -th zone in i -th sample, f is the MLP function while ω denotes the weight parameters, and λ is the weight of optional lasso penalty.

4 Experiments

In this section, we provide the training details of our models, results achieved on the split validation set, and test set result in the leaderboard.

4.1 Dataset

The Passenger Screening dataset used in this research comes from Kaggle competition at <https://www.kaggle.com/c/passenger-screening-algorithm-challenge>, This dataset contains thousands of body scans acquired by the HD-AIT system. The competition task is to predict the probability that a given body zone (out of 17 total body zones) has suspicious stuff present. For how much views are needed for efficient detection, we select 16 views as the same as in MV-CNN [3] which can balance between detection accuracy and compute efficiency. So, we rendered each human body models in 16 different views which are equally spaced around the model in advance. Figure 3a shows the rendered images of a human body and Fig. 3b shows the partition of body zones.

Since we only knew which part of body hiding suspicious stuff, the exact position of them is difficult to calculate based on the labels in the dataset. The stuff may be occluded by the human body in some views. In order to guide the algorithm to predict probabilities based on the suspicious stuff it finds, and prevent the algorithm from overfitting to the background, we hand labeled the position of suspicious stuff in the train image set, which takes 80% of the total dataset and the validate set takes the rest 20%. The position of each suspicious stuff is annotated as a box with four scalars $g_i = (cx_i, cy_i, w_i, h_i)$ which mark the center and size of the box.

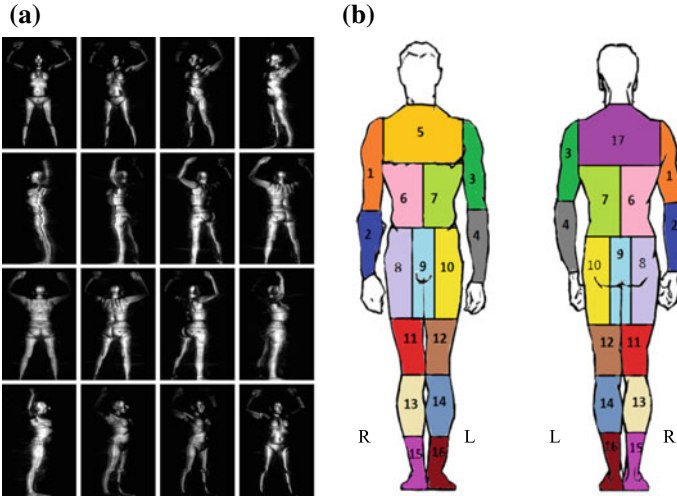


Fig. 3. Overview of passenger screening dataset. a A 3D sample from passenger screening dataset, which rendered in different views in 16 directions. b 17 different zones of human body.

4.2 Training and Implementation

We training the two networks of different stages sequentially. In stage one, we construct VGG16 and Resnet-101 network and initialize them with ImageNet [2] pretrained weights. In order to make the model more robust to various object size, we apply random scaling, cropping and flipping to the input image, and resize the images to the fix size of 400×224 in both training and testing. After training the network, we fixed the network parameters and stacked an MLP classification network on it as described before. Then we optimized the target function in Eq. 2 until it converged.

We use Tensorflow as the deep learning framework, The models are trained on a single NVIDIA GTX1070 GPU within 12 h.

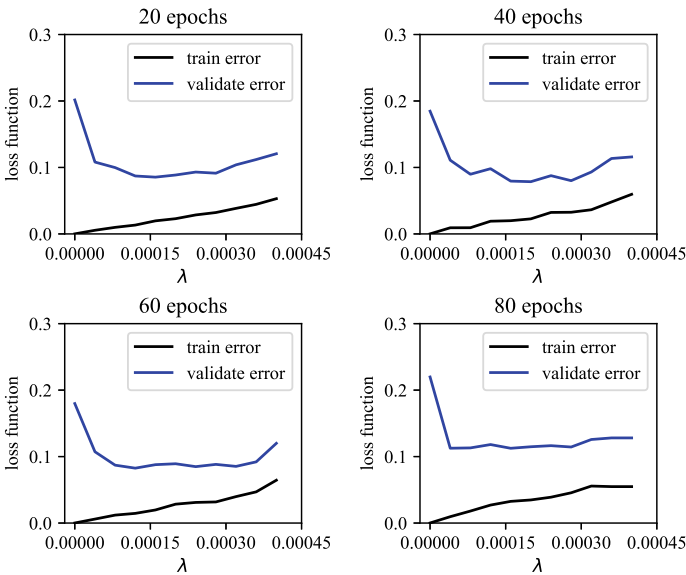


Fig. 4. The training and evaluating loss in different time of training progress and different penalty weight λ .

4.3 Results

Validation Data. We validate our method using randomly split dataset for validation, which takes 20% of the total dataset. While training stage one, we save the model parameters after training 20, 40, 60, and 80 epochs, and we train stage two with those four parameters and different regulation parameter λ . The train error and validate error with respect to λ is depicted in Fig. 4, which indicates the network of stage one starts to overfit after training 40 epochs. It also shows that the MLP detection stage needs moderate regulation to reduce

overfitting. In the end, we choose network parameters in 40th epoch and $\lambda = 0.0001$ as the final model. The result is shown in Table 1, where we compare the result of using VGG16 or Resnet-101 as base network in stage one and the result of using weight regulation or not in stage two. In all experiments, we use 0.5 as positive threshold for calculating accuracy.

From Table 1, we find the model with Resnet-101 network and regulation achieves the best performance, where the model with VGG16 network is slightly worse. From this table, we can conclude that our approach can efficiently detect and locate the suspicious stuff in 3D human body.

Table 1. Results on validation data set

Method	Validate loss	Accuracy
VGG16	0.220093	96.7263%
Resnet-101	0.179851	97.0844%
VGG16 + regulation	0.110845	97.4005%
Resnet-101 + regulation	0.086878	97.9028%

Test Data. In order to test our model in the practical application scenario, we applied our model to the test dataset which contains 2,000 new samples collected from the real world and uploaded the prediction result to kaggle server, and the result from kaggle website is listed in Table 2. The test loss is slightly larger than validate loss, which demonstrated the effectiveness of our model.

Table 2. Results on test data set

Method	Validate loss
VGG16 + regulation	0.166223
Resnet-101 + regulation	0.145261

5 Conclusion

We propose a two-stage approach based on multi-view convolution neural network to detect and locate object in 3D images, and we apply our approach to detect suspicious stuff in 3D human scan images from airport security scanner. The experiments demonstrated the proposed approach can efficiently detect and locate suspicious stuff in human body scans. Future work includes generalize this approach to different 3D images and various targets.

Acknowledgements. This research work is funded by the National Key Research and Development Project of China (2016YFB0801003) and the Sichuan province & university cooperation (Key Program) of science & technology department of Sichuan Province (2018JZ0050).

References

1. Elias, B.: Airport body scanners: the role of advanced imaging technology in airline passenger screening. Congressional Research Service, Library of Congress (2012)
2. Krizhevsky, A., Sutskever, I. and Hinton, G.: ImageNet classification with deep convolutional neural networks. In: Proceedings of Advances in Neural Information Processing Systems (NIPS), pp. 1097–1105 (2012)
3. Su, H., et al.: Multi-view convolutional neural networks for 3D shape recognition. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 945–953 (2016)
4. Qi, C.R., et al.: Volumetric and multi-view cnns for object classification on 3d data. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5648–5656 (2016)
5. Qi, C.R., et al.: Pointnet: Deep learning on point sets for 3d classification and segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 77–85 (2017)
6. Girshick, R., et al.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 580–587 (2014)
7. Girshick, R.: Fast R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 1440–1448 (2015)
8. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. In: Proceedings of Advances in Neural Information Processing Systems (NIPS), pp. 91–99 (2015)
9. Lin, T., et al.: Feature pyramid networks for object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2117–2125 (2017)
10. Dai, J., et al.: R-fcn: object detection via region-based fully convolutional networks. In: Proceedings of Advances in Neural Information Processing Systems (NIPS), pp. 379–387 (2016)
11. He, K., et al.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778 (2016)
12. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: Proceedings of the International Conference on Learning Representations (ICLR) (2015). [arXiv:1409.1556](https://arxiv.org/abs/1409.1556)
13. Sermanet, P., et al.: Overfeat: integrated recognition, localization and detection using convolutional networks. In: Proceedings of the International Conference on Learning Representations (ICLR) (2014). [arXiv:1312.6229](https://arxiv.org/abs/1312.6229)
14. Gomez-Donoso, F., et al.: Lonchanet: a sliced-based cnn architecture for real-time 3d object recognition. In: 2017 International Joint Conference on Neural Networks (IJCNN) (2017), pp. 412–418
15. Ronneberger, O., Fischer, P. and Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 234–241 (2015)



An Incremental Scheme with Weight Pruning to Train Deep Neural Network

Haonan Guo¹(✉), Zhicong Yan¹, Jichao Yang², and Shenghong Li¹

¹ Shanghai Jiao Tong University, 800 Dong Chuan Road,
Shanghai 200240, China

{haonan2012, zhicongy, shli}@sjtu.edu.cn

² Shanghai Starriver Bilingual School, Shanghai 200240, China
jasonyangjichao@163.com

Abstract. Deep neural networks have present state-of-the-art results in many different machine learning tasks. In the traditional machine learning task, we train models on the formerly prepared dataset. However, in the real-world scenarios, training data are always collected in an incremental manner, in which new samples and new classes will be added to the training data gradually. Since the traditional training method with stochastic gradient descent will suffer from catastrophic forgetting problem when training on the new data set, in this paper, we proposed a new scheme to train deep neural networks incrementally. We first train the deep model on the original dataset with a weight-pruning manner, then on the newly added training data, we train the former pruned weights while remaining the former trained core-part weights unchanged. Experiments on MNIST demonstrated that our method is efficient and can even get better performance than training from scratch on the whole dataset in the traditional manner.

Keywords: Deep neural network · Catastrophic forgetting · Weight pruning · Incremental scheme

1 Introduction

In recent years, deep neural networks have shown strong ability to learn abstract features automatically from raw data, and have dramatically improved the state of the art in many machine learning tasks such as image classification [1], object detection [2], speech recognition [3], natural language processing [4], and so on. These successes benefit from the multilayer architecture of deep neural networks, which is similar to the architecture of neural systems including optical nerves in the human visual system [5].

Most of the works on deep learning in the former literature consider the problem with a specific and complete training dataset, i.e., all of the training samples with assigned classes and associated labels are available before we start training. A variety of deep models with well-designed network architecture have presented excellent performance under such problem setting [6–8]. However, in the real-world scenarios,

since its a cumbersome work to sort out and annotated the training data manually, classes and their associated labeled data are always collected in an incremental manner. This suggests a totally different manner of training deep neural networks as illustrated in Fig. 1. Take the multiclass classification problems for instance, new classes may be added to the original classification problem during training. In these cases, the deep neural network needs to be further trained on the new data while still containing much of the knowledge it has learned from the old data.

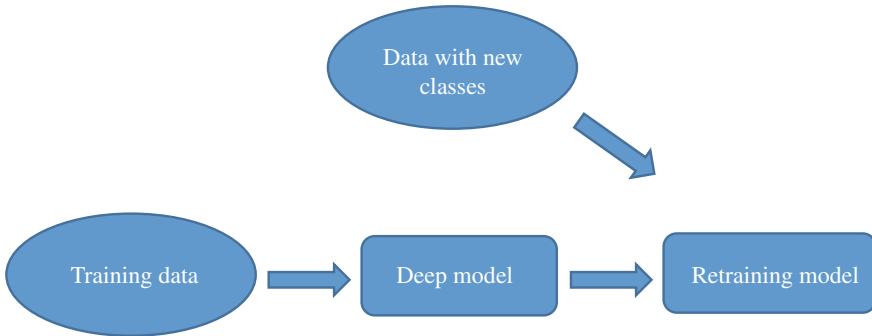


Fig. 1. Training deep neural networks in an incremental manner

Unfortunately, directly training the existing neural network on the new labeled data may results in a total disaster [9] indicates that when directly retrained the deep neural network on new data, it would raise “catastrophic forgetting” problem, which refers that the existing features learned from earlier data will be destroyed while retraining on the new data. After a few training epochs, almost all the previous knowledge will be disappeared. This can be interpreted by the cascaded architecture of deep model, a slight change on the parameters in the former layer can be amplified through the following layers.

On account of the former mentioned technical challenges, there has been a relatively little study on putting learning in such a dynamic and evolving context [10] proposed a creative algorithm that grows a deep convolutional neural network incrementally. It utilizes the kernel characteristic of the deep convolutional neural network and organizes the network in a hierarchically way. However, this method can be difficult to extend to apply to other architecture types of deep neural networks. In this paper, we considered the more general feedforward deep neural network and proposed a new method to train deep neural networks incrementally.

Our works are partially motivated by the recent works on model compression. A serious of works on model compression [11–13] have demonstrated that a large amount of weights in deep neural networks are redundancy and the models with a smaller number of parameters could perform as good as the complex model with much more parameters. Some parameter reduction methods have been proposed, one of the most succinct is to find the weakly connected weights and pruned them iteratively during training. In our method, instead of pruning the weakly connected weights,

we divide the deep neural network into two parts: the core part that contains the important weights and the edge part that contains the unimportant weights. While the data with new classes involved, we retrain the edge part of the model while keeping the core part weights unchanged. In this way, the model could learn features of the new data before the most important part of the learned features are completely destroyed. The experiment results on MNIST dataset show that our method can overcome the catastrophic forgetting problem and address the incremental learning for deep neural networks effectively.

The rest of this paper is organized as follows: in Sect. 2, we introduce the fundamental of the deep neural network and the model compression method which motivated our works. The details of our method are described in Sect. 3. And the experiments have been conducted to demonstrate the effectiveness of our method in Sect. 4. We conclude our work in Sect. 5.

2 Background

2.1 Deep Neural Network

Deep neural networks are generally cascaded by linear weighted summation layer and nonlinear activation functions alternately, followed by a classification layer at the end. The outputs of the last layer were feed into the next. Each layer can transform the feature in the last layer to a more abstract level. Through training on the dataset, the deep model can learn features which are beneficial for solving the specific tasks automatically.

In the model, each layer consists of several neural units. Let ℓ denotes the current layer, then the relationship between the outputs and the inputs can be formulized as

$$x^\ell = f(u^\ell), \text{ with } u^\ell = W \cdot x^{\ell-1} + b \quad (2.1)$$

where x^ℓ denotes the output vector of the current layer. $x^{\ell-1}$ denotes the input vector to the current layer, which is simultaneously the output of the last layer. W is the weight matrix, and b is a scalar bias term. $f(\cdot)$ denotes an element-wise nonlinearity activation function which can be chosen as rectifier linear $\max(\cdot, 0)$, or sigmoid function.

Let X and Y denote the input and output matrix of the model, then the overall calculation of the whole model can be formalized as

$$Y = f(\theta, x) \quad (2.2)$$

where θ denotes the ensemble of the parameters W and b . For a specific task, Y should be matching to the target vector y , and the gap between them can be measured by a loss function $E(Y, y)$, such as Euclidean distance or cross-entropy loss. The learning process can be formalized as the minimization of the loss function. Parameters should be adjusted to fit the training data and get less training error as well as generalization error on the test data.

Since the loss function is complicated and non-convex with millions of parameters. We cannot solve the optimization problem analytically. The optimization of the parameters is generally executed iteratively through stochastic gradient descent.

2.2 Training with Model Compression

Though the recent deep neural networks have present the state-of-the-art results in many machine learning tasks, these models with millions of parameters made it difficult for training and the real-world implementation. A smaller model may have the advantage of faster training and less storage. Recently, several works focus on the model complexity reduction problem [11–13], these works demonstrated that for most of the widely used model architecture, the large amount of weights are not all necessary. They believe that the models with a smaller number of parameters could perform as good as the complex model with much more parameters. Some parameter reduction methods have been proposed to simplify the model while preserving the performance [11] presented a parameter pruning method with weight norm regularization and weight pruning. The training steps are illustrated in Fig. 2. They remove all connections after an initial training phase, then retrain the sparse network so the remaining connections can compensate for the connections that have been removed. The phases of pruning and retraining may be repeated iteratively to further reduce network complexity. This training method can improve the energy efficiency and storage of neural networks without affecting accuracy by finding the right connections. In this paper, we utilize a similar training process in our first phase of training on the original task.



Fig. 2. Training steps with weight pruning

3 Methods

We sequentially train our model on the original task and the incremental task. There are additional training data which belong to new classes in the incremental task. For the original task, we train our deep neural networks on a similar process with the pruning method introduced in Sect. 2. After the parameter update with each mini-batch of training data, we find out the low-weight connections with weights below a threshold. Once a connection is continuously marked as low weight for a critical time, we identify the connection as weakly connected and prune it in the following training process, i.e., set the value of the weight to zero in the feedforward propagation and set the gradient to zero in the backward propagation. The prune operation should proceeded all the way with the training process until the performance on the valid set stops improving.

Through the first phase of training, we get a sparse connected deep neural network with the proper amount of parameters. The redundant weights were pruned and the remaining weights were trained sufficiently on the currently given dataset. We denote the remaining weights as a core part of the neural network and the pruned weights as edge parts. For the new involved data of new classes, in order to contain the well-learned features through training on the early dataset as more as possible, we retrain the former pruned weights and keep the core part of the model unchanged. So new features can be learned to adapt the newly involved data while the former learned knowledge is destroyed in a tiny degree. The training process is shown in Table 1.

Table 1. Training processes of our method

<i>Train model on the original task</i>
Step 1: the follow steps are processed iteratively to prune the weak connections
(a) select a stochastic mini-batch from the training data set
(b) calculate the gradients of the weights and biases by backpropagation
(c) update the weights and biases with a proper learning rate
(d) identify the weak connections and prune them until the desired sparseness being achieved
Step 2: train the pruned model with the conventional SGD method
<i>Train model on the new added task</i>
(a) assigned the former pruned weights as edge-part of the model and the remaining weights as core-part of the model
(b) train the edge-part of the deep neural network with the new added training data while remaining the core-part unchanged

To accelerate the pruning process, we adopt the L2 regularization in the pruning phase. Since L2 regularization can penalize large parameters, and push the weights to be weak connections. Apart from the hyper-parameters on the traditional neural networks, like momentum rate, L2 rate and so on we have two new parameters: threshold value and critical time for the identifying of the weak connections. All the hyper-parameters needs to be tuned on the valid set to get the best performance.

The threshold values can seriously impact the amount of pruned connections when the prune operation was first processed. If they are set to be too large, a huge amount of weights will be pruned after the first prune process, while much of them are still around the initial values and have not been fully trained. Since our prune operation is irreversible, it can get a bad connected architecture. And if the threshold values are set to be too small, the effect of pruning the redundancy connection will be weakened. So we should tune the threshold values properly on account of the mean value of the weights for each layer and the ultimate sparseness we want to achieve.

4 Experiments

We test our method on deep neural networks for classification problem on the MNIST data set. The MNIST data set consists of 28×28 pixel handwritten digit images of 0–9. The task is to classify the images into 10 digit classes. There are 60,000 examples in the training set and 10,000 in the test set. We select 10,000 examples from the training set for validating, which can help us find the optimal hyper-parameters (learning rate, threshold values for pruning and L2 parameter in our experiments). Some examples of the data set are shown in Fig. 3. Our deep models have the input layer of 784 dimensions and we employ the z-score operation before they are fed into the input layer.



Fig. 3. Some examples in the MNIST dataset

4.1 Analysis of Catastrophic Forgetting

First, we demonstrate that our proposed method can overcome the catastrophic forgetting problem when training on the new data. We divided the whole training set into two parts: the samples with label 0–4 as the original part and the samples with 5–9 as the newly added part. For the original part, we train the deep neural network with weight-pruning method described above. The pruned weights were assigned as edge part of the model and remaining weights were assigned as the core part of the model. Then for the newly added part data, we train the edge part of the model while keeping the core part of the model unchanged. We also compared with the traditional manner which to retrain the model straightly, i.e., the former model was used to initialize the parameters. The contrast experiments were implemented for deep neural networks with different hidden layers, each with 512 units. Table 2 shows the test error rate for retraining with the two different manners, respectively. We can see that if we retrain the model straightly, the test error rate can get better on the new data, however, the test error on the original data fall extremely. With our proposed method, we get a comparative test error on the new data, while the test error on the original data is about 10% suppress the former manner. It demonstrates that our method can overcome the “catastrophic forgetting” problem to some extent.

Table 2. Analysis for catastrophic forgetting on deep neural networks with different hidden layers

Hidden layer	Test error on the original data	Retraining straightly		Retraining with the proposed method	
		Original data	New data	Original data	New data
512 * 3	0.66	62.4	1.33	52.1	1.45
512 * 4	0.64	62.5	1.39	47.8	1.41
512 * 5	1.70	63.8	1.33	50.6	1.45

4.2 Incrementally Training on the Whole Dataset

In the section, we will demonstrate that our method of incrementally learning can get better performance than the traditional training manner of training on the whole dataset straightly. For the incremental learning process, we first train the model on the samples with labels 0–4 with the weight-pruning manner. Then we train the edge part of the model on the whole dataset. As a comparison, we compare our method with the traditional method of training on the whole dataset from scratch. Table 3 shows the test error on deep neural networks with 3–5 hidden layers, each with 512 neural units. It’s obvious that our method can suppress the traditional training method slightly. However, our incremental training manner is more accord with the real-world scenario.

Table 3. Test error for incrementally training versus training from the scratch

Hidden layer	Incrementally training with the proposed method	Training from the scratch
512 * 3	1.56	1.81
512 * 4	1.61	1.97
512 * 5	1.70	1.93

5 Conclusion

In this paper, we present a new scheme to train deep neural networks which is similar to the learning process of human beings in the real-world scenario. We first train the deep neural network with weight pruning on the existing dataset and the pruned weak connections will be denoted as edge-part of the model. When new samples and new classes of data are added to the training data, we train the edge-part of the model while remaining the core-part unchanged. In this manner, we can adjust the model to fit the new samples while avoiding loss knowledge we have already learned. The experiments results demonstrate that our proposed method can defeat the catastrophic forgetting problem to some extent and the incrementally training manner can even get better performance than the traditional manner of training on the whole dataset straightly.

Acknowledgements. This research work is funded by the National Key Research and Development Project of China (2016YFB0801003)

References

1. He, K., Zhang, X., Ren, S., et al.: Delving deep into rectifiers: surpassing human-level performance on imagenet classification 1026–1034 (2015)
2. Sermanet, P., Eigen, D., Zhang, X., et al.: OverFeat: integrated recognition, localization and detection using convolutional networks. Eprint Arxiv (2013)
3. Hinton, G., Deng, L., Yu, D., et al.: Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Process. Mag.* **29**(6), 82–97 (2012)
4. Mikolov, T., Deoras, A., Povey, D., et al.: Strategies for training large scale neural network language models. In: *Automatic Speech Recognition and Understanding*. IEEE, pp. 196–201 (2011)
5. Hubel, D.H., Wiesel, T.N.: Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* **195**(1), 215–243 (1968)
6. Lin, M., Chen, Q., Yan, S.: Network in network (2013). arXiv preprint [arXiv:1312.4400](https://arxiv.org/abs/1312.4400)
7. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. *Comput. Sci.* (2014)
8. Szegedy, C., Liu, W., Jia, Y., et al.: Going deeper with convolutions. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9 (2015)
9. Moe-Helgesen, O.M., Stranden, H.: Catastrophic forgetting in neural networks. *Dept. Comput. Inf. Sci. Nor. Univ. Sci. Technol. (NTNU) Trondheim Nor. Tech. Rep.* **1**, 22 (2005)
10. Xiao, T., Zhang, J., Yang, K., et al.: Error-driven incremental learning in deep convolutional neural network for large-scale image classification. In: *Proceedings of the 22nd ACM International Conference on Multimedia*. ACM, pp. 177–186 (2014)
11. Han, S., Pool, J., Tran, J., et al.: Learning both weights and connections for efficient neural network. In: *Advances in Neural Information Processing Systems*, pp. 1135–1143 (2015)
12. Srinivas, S., Babu, R.V.: Data-free parameter pruning for deep neural networks (2015). arXiv preprint [arXiv:1507.06149](https://arxiv.org/abs/1507.06149)
13. Han, S., Mao, H., Dally, W.J.: Deep compression: compressing deep neural networks with pruning, trained quantization and Huffman coding (2015). arXiv preprint [arXiv:1510.00149](https://arxiv.org/abs/1510.00149)



A Deep Learning Method of Moving Target Classification in Clutter Background

Ningyuan Su, Xiaolong Chen^(✉), Xiaoqian Mou, Lin Zhang,
and Jian Guan

Radar Detection Research Section, Naval Aviation University, Erma Road 188,
Yantai, China
cx1cx11209@163.com

Abstract. The Doppler spectrums of radar echoes of targets can reflect the change of the instantaneous velocity of targets. Therefore, it can be used for analyzing the motion state of the target and classifying them. Besides, deep learning is widely used in the classification of images. This paper proposes a deep learning based method of classifying targets in sea clutter. First, we introduce the motion model of targets and analyze their Doppler spectrum, based on which, we stimulate the time–frequency images of targets’ radar echoes. Since clutters in echoes usually obey Weibull distribution, we add Weibull clutter (Mezache and Soltani) to a novel threshold optimization technique for far-away detection in Weibull clutter using fuzzy neural networks, 2007, [1]) to the echo signals. Then we classify targets with different networks using NVIDIA DIGITS, based on the images and analyze the results of classification.

Keywords: Doppler spectrum · Weibull clutter · Target classification · Deep learning

1 Introduction

Target detection in clutter has great application value in both military and civilian fields. In recent years, the use of the micro-Doppler theory in target detection has gradually gained people’s attention. Micro-Doppler reflects the changing of Doppler, and the geometric structure of the target and the integrated modulation characteristics of the fine motion of the radar echo are reflected in the micro-motion characteristics of the target. Therefore, it provides a new approach for radar target detection and feature extraction [2]. Since the Doppler spectrum of echoes of targets, such as ships, are the results of their translation and micro-motion, we can detect and classify different targets by modeling their motion and analyzing their echoes’ Doppler spectrum. Besides, deep learning theory is also widely used in the field of radar, such as gesture detection and classification, UAV target detection and autopilot. The main techniques include clutter suppression, clutter estimation CFAR detection, and so on. Among all the networks, CNN

is one of the most widely used one, which effective in feature extraction. Since the difference between targets can be reflected on their radar echo signals and deep learning is effective in classification, we can classify targets with deep learning methods.

In this paper, we introduce the motion model of the target and stimulate time–frequency images of echoes based on the model. Then we classify targets using deep learning method, including 3 different networks, and analyze the results.

2 Motion Model of Targets

In this part, we introduce two kinds of motion models, both of which are the focus on the variety of instantaneous distance between target and radar [3].

2.1 Radar Echo Model of Nonuniform Translational Target

Take marine target as an instance, radar and target are in the same horizontal plane approximately when the observation distance is long or the height of radar is low. And we assume there is no migration of target echo in a short time. Then we construct the following model.

The model consists of a target-fixed coordinate system $C_{ref} = (X, Y, Z)$, coordinate system of target motion $C_{mov} = (x, y, z)$, and RLOS coordinate system $C_{rlos} = (q, r, h)$. In the target-fixed coordinate system $C_{ref} = (X, Y, Z)$, the origin is set at the centroid of the hull and the distance between the origin and the radar is constant. In the coordinate system of target motion, the origin in the set at the centroid of the hull and axis ox , the vertical axis, is parallel to the roll axis, pointing to the bow. The axis oy , the horizontal axis, is parallel to the pitch axis, pointing to the port side. The axis oz , the altitude, points directly above the target. The axis ox , oy and oz are set as the roll axis, pitch axis and the yaw axis of the target. The coordinate (x, y, z) means the relative position of the observation point and the centroid. In the RLOS coordinate system, the line of sight is on the XOY plane. The axis h is perpendicular to axis r , and axis q meet the right-hand rule.

In order to meet the demand for remote detection and high resolution, radar emission LFM signal is used.

When the target moves toward radar, we only observe the radial velocity component. The distance is a polynomial of time, and it can be rewritten in Taylor series form, preserving only first four monomials.

$$r_s(t_m) = r_0 - v_0 t_m - a_s t_m^2 / 2 - a_{s2} t^3 / 6 \quad (2.1)$$

In the formula, v_0 is the initial speed, and a_s is the. Parameter a_{s2} is the rate of change of acceleration. After modulation and pulse compression, the output of the echo signal is

$$s_{PC}(t, t_m) = A_r \text{sinc}[B(t - \tau)] \exp(-j2\pi f_c \tau) \quad (2.2)$$

A_r is the amplitude of signal, replace τ with $2r_s(t_m)/c$, and calculate the derivative of time for the phase.

$$f_t = -\frac{2}{\lambda} \frac{d\left(r_0 - v_0 t_m - \frac{a_s t_m^2}{2} - \frac{a_{s2} t_m^3}{6}\right)}{dt_m} = \frac{2}{\lambda} (v_0 + a_s t_m + a_{s2} t_m^2 / 2) \quad (2.3)$$

Parameter λ is the wavelength of transmitting signal. The formula above shows that after modulation and pulse compression, the echo signal is a function of velocity, acceleration and the rate of change of acceleration, and can be seen as an LFM signal.

2.2 Radar Echo Model of Three-Axis Rotation Target

We still take marine target as an example and assume that radar and target are in the same horizontal plane and the motion of target includes pitching, rolling and yawing. Then the rotation can be described as the rotation matrix \mathbf{R}_{z-y-x} :

$$\mathbf{R}_{z-y-x} = \mathbf{R}(\theta_x) \mathbf{R}(\theta_y) \mathbf{R}(\theta_z) = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad (2.4)$$

The $\mathbf{R}(\theta_x)$, $\mathbf{R}(\theta_y)$, $\mathbf{R}(\theta_z)$ [3] are the matrix of rolling, yawing and pitching, respectively. And θ_x , θ_y , θ_z are corresponding rotation angles.

In order to get the movement state of the target in RLOS coordinate system, we convert the coordinate system of target motion C_{mov} to that in target-fixed coordinate system C_{ref} using a rotation matrix \mathbf{R}_{z-y-x} .

$$C_{ref} = \mathbf{R}_{z-y-x} C_{mov} = \begin{bmatrix} a_{11}x + a_{12}y + a_{13}z \\ a_{21}x + a_{22}y + a_{23}z \\ a_{31}x + a_{32}y + a_{33}z \end{bmatrix} \quad (2.5)$$

Then we convert the movement state in C_{ref} to that in C_{rlos} , based on the geometric relations of ship and radar.

$$C_{rlos} \triangleq \begin{bmatrix} q \\ r \\ h \end{bmatrix} = \mathbf{R}(\varphi) C_{ref} = \begin{bmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11}x + a_{12}y + a_{13}z \\ a_{21}x + a_{22}y + a_{23}z \\ a_{31}x + a_{32}y + a_{33}z \end{bmatrix} \quad (2.6)$$

Parameter φ is the angle between the line of sight and the direction of the hull axis. Ignoring the altitude information, we can get the RLOS distance as follows.

$$r_s(t_m) = \sin \varphi (a_{11}x + a_{12}y + a_{13}z) + \cos \varphi (a_{21}x + a_{22}y + a_{23}z) \quad (2.7)$$

Then we calculate the derivative of time for $r_s(t_m)$, and get the micro-Doppler frequency caused by the rotation of target.

$$f_r = \frac{2v_r(t_m)}{\lambda} = \frac{2}{\lambda} \frac{dr_s(t_m)}{dt_m} \quad (2.8)$$

The angular velocity of three kinds of rotation are: $\omega_x = \theta_x/t$, $\omega_y = \theta_y/t$ and $\omega_z = \theta_z/t$.

3 Deep Learning Models

3.1 LeNet

Proposed by Yann LeCun and Yoshua Bengio, LeNet is a set of deep learning framework based on image convolution, and it is applied to handwritten digit recognition [4]. It consists of two convolution layers, each of which is followed by a pooling, and two fully connected layers. LeNet is the first truly multilevel structure learning algorithm that uses spatial relative relationships to reduce the number of parameters to improve training performance.

3.2 AlexNet

AlexNet is a historic network frame, which is proposed in 2012. It consists of 5 convolution layers and 3 fully connected layers. Each of the first two convolution layers consists of a sub-layer and a response normalization layer. Each layer is followed by a ReLU operation and each of the first, second, and the fifth convolution layer is followed by a max pooling layer. Compared with previous network frames, AlexNet has some advantages to avoid saturation of activation function, such as using local response normalization [5].

3.3 GoogLeNet

GoogLeNet is a deep convolutional neural network model constructed by increasing the depth and width of the network model. GoogLeNet, consists of 22 layers with parameters and more than 100 separated layers, uses RGB three-color channel and the size of sensor pixel is 224×224 . In order to prevent the gradient from disappearing, its two losses are set at a different depth to ensure gradient the return of gradient. And all the convolution operations are followed with a ReLU to avoid overfitting and to accelerate the speed of convergence [6]. GoogLeNet also increase the width of the network by using the inception unit.

4 Experiment Designing

4.1 Simulation Data Settings

Based on the motion model above, we simulate the echo signal and convert it to time–frequency image with a short-time Fourier transform (STFT). In this program, we use Hamming Window and the window is 1/4 the length of the sampled signal. We also add Weibull clutter [1] with sixteen different values of SCRs.

Firstly, we assume that radar works in X band and the carrier frequency is 8 GHz. The types of motion include uniformly motion, uniformly accelerated motion, rotation [7] and nonuniformly accelerated motion. Then we set the carrier frequency to 2 GHz (S band), and redo the simulation. The parameters are as follows (Table 1).

Table 1. Motion data settings

Motion	Motion data settings
Uniform motion	$v(\text{m/s})$: 5, 10, 15
Uniformly accelerated motion	$v_0(\text{m/s})$: 5, 10, 15 $a_s(\text{m/s}^2)$: 1, 2, 4, 8, 16
	$v_0(\text{m/s})$: 5, 10, 15 $a_s(\text{m/s}^2)$: -1, -2, -4, -8, -16
Nonuniformly accelerated motion	$v_0(\text{m/s})$: 50, 100, 200 $a_s(\text{m/s}^2)$: -160, -40, -10, 10, 40, 160 $a_{s2}(\text{m/s}^3)$: -640, -320, -160, 160, 320, 640
Rotation 1 [7]	$\bar{\omega}_x = 0.38 \text{ rad/s}$ $\bar{\omega}_y = 0.16 \text{ rad/s}$ $\bar{\omega}_z = 0.08 \text{ rad/s}$ $T_x = 26.4 \text{ s}$ $T_y = 11.2 \text{ s}$ $T_z = 33.0 \text{ s}$ $x = 100 \text{ m}$ $y = 20 \text{ m}$ $z = 20 \text{ m}$ (The distance vector from centroid of target to the scattering point)
Rotation 2 [7]	$\bar{\omega}_x = 6.30 \text{ rad/s}$ $\bar{\omega}_y = 1.01 \text{ rad/s}$ $\bar{\omega}_z = 0.54 \text{ rad/s}$ $T_x = 12.2 \text{ s}$ $T_y = 6.7 \text{ s}$ $T_z = 14.2 \text{ s}$ $x = 25 \text{ m}$ $y = 7 \text{ m}$ $z = 7 \text{ m}$

4.2 Model Training

We classify the time–frequency images into 5 categories based on the types of motion: uniformly motion, uniformly accelerated motion, nonuniformly accelerated motion and two types of rotation. Then we make data set in digits and train three models, LeNet, AlexNet, and GoogLeNet, with following settings (Table 2) the curve of parameters as shown in Figs. 1, 2 and 3.

Table 2. Model data setting

Training epochs	Snapshot interval	Validation interval	Solver type	Base learning rate
30	1	1	SGD	0.01

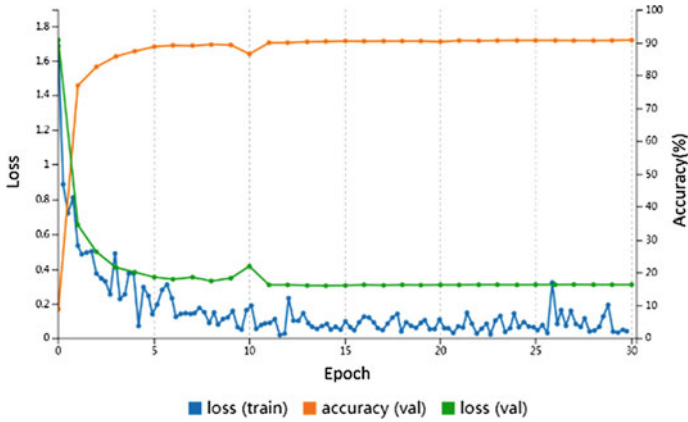


Fig. 1. LeNet training

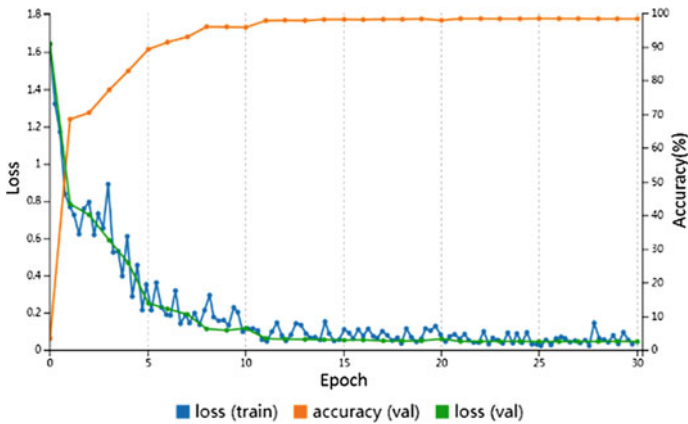


Fig. 2. AlexNet training

4.3 Model Testing

In program 1, we test three models with the same dataset, which includes the time–frequency images of uniform motion, uniform acceleration, nonuniform acceleration and two types of rotations. Program 2 consists of 16 independent tests. In each of the

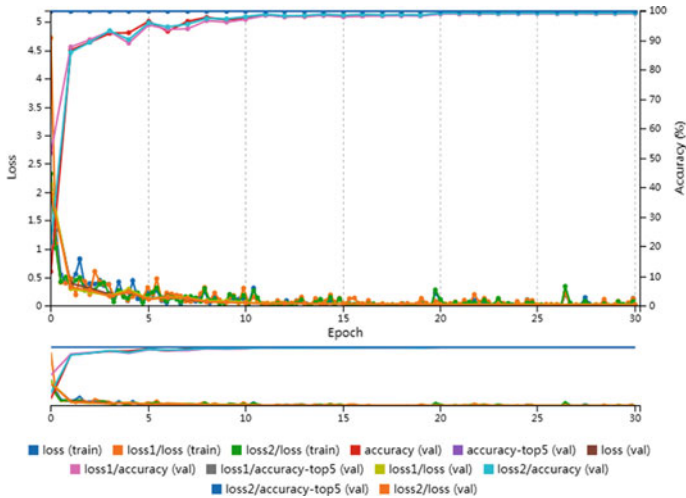


Fig. 3. GoogLeNet training

Table 3. Accuracy of different models

	Uniform motion (%)	Uniform acceleration (%)	Non-uniform acceleration (%)	Rotation1 (%)	Rotation2 (%)	Duration (s)
LeNet	61.46	93.95	95.67	91.56	90.00	487
AlexNet	81.25	98.99	83.45	86.88	75.00	509
GoogLeNet	87.50	98.99	89.59	95.63	94.69	465

tests, we use the trained GoogLeNet to classify 5 kinds of different images under a certain SCR background.

Target Classification with Different Models

We test the three trained models with test data and the results are shown in Table 3.

Target Classification Under Different SCRs

We make another 16 datasets and each of them is stimulated with different values of SCRs. Then we use the trained GoogLeNet to classify these datasets respectively. Some of the results are as follows (Table 4).

Table 4. Accuracy under different SCRs

	Uniform motion (%)	Uniform acceleration (%)	Non-uniform acceleration (%)	Rotation 1 (%)	Rotation 2 (%)
SCR = -9	58.33	95.00	78.52	72.50	100.00
SCR = -7	66.67	96.67	88.03	92.50	95.00
SCR = -5	66.67	100.00	92.25	90.00	95.00
SCR = -3	66.67	98.33	93.66	90.00	97.50
SCR = -1	66.67	100.00	94.37	97.50	100.00
SCR = 1	75.00	100.00	94.37	95.00	100.00
SCR = 3	91.67	100.00	94.72	100.00	100.00
SCR = 5	75.00	100.00	95.07	97.50	100.00

5 Conclusion

Based on the results of our tests and more detailed data, we collected during the program, our conclusions are as follows.

Among the three models we mentioned, GoogLeNet performs best in target classification. Though the accuracy of classifying uniform motion and nonuniform acceleration targets is not high (under 90%), it seldom identifies translational targets as rotations. Misses are mainly caused by confusing translational targets with different motion parameters.

The accuracy of identifying two types of rotation is high. And deep learning based target classification method can also be used to distinguishing similar targets, such as different kinds of ships. The micro-motion of different ships are similar, but the amplitude and period of their micro-motion different, which makes it possible for deep learning networks to distinguish them based on their time–frequency images.

The accuracy of classification is influenced by the SCR of the background. Pre-treatments, such as clutter suppression, can improve the performance of the networks.

Acknowledgements. This work was supported in part by The National Natural Science Foundation of China (61871391, U1633122, 61871392, 61531020), Scientific Research Development of Shandong (J17KB139), and Young Elite Scientist Sponsorship Program of CAST (YESS20160115).

References

1. Mezache, F., Soltani, A.: Novel threshold optimization technique forfar detection in weibull clutter using fuzzy-neural networks. In: 2007 IEEE International Conference on Signal Processing and Communications (ICSPC 2007), Dubai, United Arab Emirates (2007)
2. Xiaolong, C., Jian, G., You, H.: Applications and prospect of micro-motion theory in the detection of sea surface target. *J. Radars* **2**(1), 123–134 (2013)
3. Xiaolong, C., Yunlong, D., Xiuyou, L., Jian, G.: Modeling of micromotion and analysis of properties of rigid marine targets. *J. Radars* **4**(6), 630–638 (2015)

4. LeCun, Y., Bottou, L., Bengio, Y., et al.: Gradient-based learning applied to document recognition. *Proc. IEEE* **86**(11), 2278–2324 (1998)
5. Zhang, J.: Research on image retrieval based on fusion feature of AlexNet. Chongqing University of Posts and Telecommunications, Chongqing (2016)
6. Bai, Y., Wan, H., Bai, C.: Study on human behavior classification in still images based on GoogLeNet. *Comput. Knowl. Technol.* **13**(18), 186–188 (2017)
7. Gao, J.: ISAR Ship Imaging and Cross-Ranging Scaling with Multipath and Sea Clutter and Interference. Harbin Institute of Technology, Harbin (2009)

Optical System



Three-Dimensional Laser Scanning for the Bridge Deformation of Shanghai Maglev Train

Yanwen Wu¹, Lei Zhang^{2(✉)}, V. Badenko³, and R. D. Garg⁴

¹ Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai 200241, China

² MOE International Joint Lab of Trustworthy Software, East China Normal University, Shanghai 200062, China
lzhang@ce.ecnu.edu.cn

³ Water Resources and Hydrotechnical Engineering Department, Peter the Great St. Petersburg Polytechnic University, 195251 St. Petersburg Polytechnicheskaya 29, Russia

⁴ Civil Engineering Department, Indian Institute of Technology Roorkee, Roorkee 247667, India

Abstract. In China, infrastructure constructions of the city are developed continuously. The state of urban community safety and its capability are an important sign of its quality and civilization. High-precision bridge deformation detection is eagerly needed to ensure the safety of city facilities. In this paper, BP neural network is applied for high-precision 3D modeling of point cloud data obtained by 3D laser scanner along the Shanghai maglev train. Based on the 3D laser scanner technology, the deformation of the maglev train's bridge can be usually monitored. After analyzing the experiments on monitoring the bridge deformation of the Shanghai maglev train, a certain deformation effect when the maglev train is passing can be monitored. So, we will get a great deal of data from the Shanghai maglev train safety information.

Keywords: 3D laser scanning · BP neural network · Deformation monitoring

1 Introduction

With the improvement of China's comprehensive strength and infrastructure constructions, the safety of urban community has drawn special attention. As one of the countries that are greatly affected by the geological disaster, China needs to pay particular attention to the safety of infrastructures. During the operation and construction of facilities in major cities, both natural and man-made factors will cause the urban facilities to be deformed, posing a potential safety hazard. When the deformation

exceeds the range that the infrastructures can withstand, it may cause disaster and threaten people's lives and property. In order to ensure the safe use of such facilities to ensure the normal operation of the city, long-term and high-precision deformation monitoring system is required [1].

Traditional deformation monitoring mainly uses Electronic Total Station, RS, GNSS [2], and InSAR. GPS is widely applied in the field of deformation monitoring. However, the accuracy of GPS is limited due to factors such as multipath, satellite coverage, cycle slips, and so on [3]. The maximum data acquisition rate by using GPS is about 10–20 Hz, which is not suitable for deformation monitoring. For accurate monitoring, the data acquisition rate should over 100 Hz [4]. Although some supplementary instruments such as accelerometers and inclination sensor [5] were used for accuracy enhancement, the monitoring system is complex and inefficiency.

As a newly emerging technology, 3D laser scanner is widely applied to deformation monitoring due to its characteristics such as noncontact, high efficiency, and high precision. Compared with the traditional monitoring technology, it can quickly capture the deformation information and establish a real and reliable model while satisfying the deformation monitoring accuracy [6]. By using this technology, several million three-dimensional points can be obtained in just a few minutes, which are called point cloud data and conducive to the construction of the overall model.

In this paper, the 3D laser scanner is used to obtain the overall point cloud data of the bridge for the Shanghai maglev train. Then, BP neural network is applied to reconstruct the surface of the bridge for deformation detection. By analyzing the point cloud data, a certain deformation effect when the maglev train is passing can be observed. Thus, a great deal of data of Shanghai maglev trains safety information will be got.

2 Technical Principal

2.1 3D Laser Scanning Technology

The rise of 3D laser scanning technology represents epoch-making progress of mapping technology. By using this technology, the 3D data information of the target area can be got in a fast, precise, and contactless way. At the same time, error caused by optics and contact can be avoided. In such a noncontact way, massive point cloud data are obtained for large-area and high-density deformation monitoring. Based on massive data, models for heritage buildings [7] and infrastructures can be built. By preprocessing and modeling point cloud data, the deformation of the target construction can be analyzed.

As for deformation monitoring, the most important feature is high precision. Therefore, whether 3D laser scanner can be applied for deformation monitoring depends on its measurement accuracy. At present, the accuracy of this technology can reach millimeter level, which can completely replace the traditional monitoring methods in a certain area such as the health monitoring of buildings, bridges and so on,

thus saving a great deal of costs and get a more intuitive deformation analysis [6]. The experiment has been done to prove that when applied 3D laser scanner to monitor the deformation of bridge, the modeling accuracy is better than 3 mm [8]. It is acceptable accuracy for monitoring the bridge deformation.

The working principle of the 3D laser scanner is to determine the 3D coordinates of the target point by measuring the distance (distance observation S) and two angles (horizontal observation α and vertical observation β) through the device carried by the instrument. The distance is calculated based on the time that the pulse takes to travel from the equipment to the target and then return. In addition, the instrument can judge the pulse strength through the laser energy intensity reflected from the target points [9] (Fig. 1).

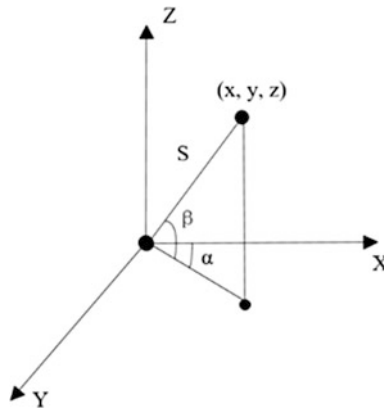


Fig. 1. The principle of the 3D laser scanner

The 3D coordinate calculation formula for the target point is as follow.

$$\begin{cases} x = S \cos \beta \cos \alpha \\ y = S \cos \beta \sin \alpha \\ z = S \sin \beta \end{cases} \quad (1)$$

2.2 BP Neural Network

BP neural network, also known as Error Back Propagation neural network, is a kind of feedforward network based on nonlinear transform unit. It was put forward by Rumelhart and McClelland in 1985. BP neural network realizes the nonlinear mapping function, thus it is widely applied to reconstruct freedom surfaces [10]. Through

training and predicting, the cumulative weights between layers can be changed to forecast the output data. By comparing the predicted output with the actual one, the pros and cons of the network construction can be analyzed. When the difference between the actual output and predict one is in the range of a preset value, the training process has been finished. Otherwise, the network will be trained again.

BP neural network is a typical hierarchical network, which consists of input, hidden, and output layers [11]. All layers are connected by a full connection method. The detailed structure is shown in the following Fig. 2.

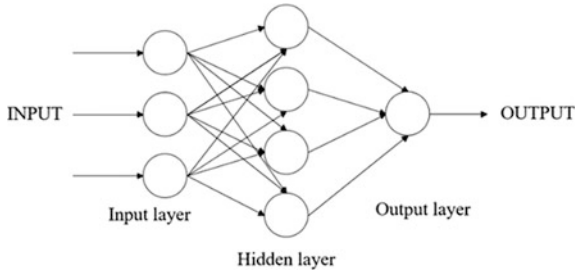


Fig. 2. The structure of neural network

In this paper, the BP neural network is used to reconstruct the surface of the bridge for the Shanghai maglev train. When the freedom surface is built, the deformation of the bridge can be monitored by comparing the models observed in different periods. High-speed maglev train causes its bridge vibrates in the vertical direction. Thus, the change of the Z coordinates are the parameters we interested in. X and Y coordinates of the bridge collected by 3D laser scanner are chosen as the input values and Z coordinates are the output one.

The process for the training consists of two steps. The first one is getting the forecast output by forwarding propagation. The second one is updating the weights between the nodes in neighbor layers by back forward propagation.

First, data should be imported into the input layer, then the hidden layer and the predictive value is got from the output layer.

The effective function, also known as the activation functions or transfer function, generally select the logistic function.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

The formulas for the process of forward propagation are as follows.

$$H_j = f\left(\sum A_{ij} \times X_i - \theta_j\right) \quad (3)$$

$$O_k = f\left(\sum B_{jk} \times H_j - \theta_k\right) \quad (4)$$

where f is the nonlinear effective function, θ is the neuron threshold, A_{ij} is the weight between node i in the input layer and node j in the hidden layer and B_{jk} is the weight of node j in the hidden layer and node k in the output layer.

Second, the weights between the nodes in neighbor layers are updated by back forward propagation. The difference between the predicted result of the output layer and the actual one is calculated. The difference is propagated back forward, thus the error for each layer can be got. According to the error, the weights of every layer are updated.

The performance of BP neural network is calculated by the loss function model. Referring to Eq. (5), it calculates the difference between the network output and its expected output after propagates through the network.

$$E = \frac{\sum (t_i - m_i)^2}{2} \quad (5)$$

where t_i is the expected output from node i , m_i is the calculated output from node i .

The self-learning process of the BP neural network includes setting and modifying the weights between two connected layers.

$$\Delta W_{ij}(n+1) = \eta \times \Phi_i \times O_j + a \times \Delta W_{ij}(n) \quad (6)$$

where η is a learning rate, Φ_i is the calculated error for the output of node i , O_j is the calculated output for the output of node j and a is the influence of the inertial term.

3 Experimental System

3.1 Data Acquisition

Shanghai maglev train departs from Longyang road station of Shanghai Metro Line 2, and extends eastward to Pudong International Airport with a total length of 29.863 km. The speed of the train can reach 430 km per hour. A train departs every 15 min.

A point along the track of the Shanghai maglev train is selected for observation. Two experiments were conducted on November 6, 2016 and March 2, 2017, respectively. The experimental equipment is RIEGL VZ-1000 and the detailed information Fig. 3 can be referred.



Fig. 3. The experiment site

The exact location of the experiment can be calculated by solving the observation and navigation data got from Trimble BD982 antenna. The result is shown in Table 1.

Table 1. Position result

Parameter	Average	Standard deviation
Longitude	121.718571°	5.53773E-06
Latitude	31.179496°	5.08526E-06
Elevation	19.723061 m	1.33764

3.2 Coarse Scanning

First, the appropriate measuring points should be selected. Adjusting the attitude of the instrument to keep it horizontal and measuring the height of the instrument. Connect the 3D laser scanner with the computer to receive measurement data. By coarse scanning the surrounding environment, the overall point cloud data of the surrounding environment can be got. The scanning parameters are shown in Table 2.

Table 2. Coarse scanning parameters

——	Vertical	Horizontal
Starting angle/degree	30.000	0.000
Finishing angle/degree	130.020	360.000
Interval/degree	0.060	0.040
The number of points/point	1668	9001
Measuring time	2'04"	
Measuring mode	450 m (300 kHz)	

We choose the measuring mode at 450 m (200 kHz), which means the Laser PRR (Peak) is 300 kHz and the maximum measuring distance is 450 m. In this measuring mode, the details of the construction can be well detected.

Through the above parameter settings, the result of the scanning is as follows (Fig. 4).

**Fig. 4.** Original point cloud

During scanning, colorful images of the scanning scene were taken. After scanning, the photos and point cloud data are registered and the following three-dimensional colorful point cloud scene can be obtained [12] (Fig. 5).

By coarse scanning, the entire scanning environment can be observed, thus scanning the part of the interest in the experiment. In order to observe the vibration effect of the maglev train on the bridge, a part of the train track bridge was selected for fine scanning.



Fig. 5. Point cloud after color matching

3.3 Fine Scanning

Scanning parameters are shown in Table 3.

Table 3. Fine scanning parameters

—	Vertical	Horizontal
Starting angle/degree	84.515	242.326
Finishing angle/degree	91.081	247.445
Interval/degree	0.060	0.012
The number of points/point	111	430
Measuring time	0'06"	
Measuring mode	450 m (300 kHz)	

In order to make the accuracy of the obtained point cloud data meet the requirements, it is necessary to filter the information which has done nothing with the deformation monitoring, such as vegetation and buildings. These redundant data can be removed by using RiSCAN PRO software.

The colorful point cloud data obtained is shown in Fig. 6.

Through surface reconstruction, the overall deformation of the train track bridge can be observed. Then, we can extract a section of the point cloud data from the surface to measure the approximate vibration of the bridge. In this experiment, the BP neural network was used for the reconstruction of the train track bridge.



Fig. 6. Point cloud data of fine scanning

4 Analysis of Deformation Monitoring

After acquiring data, the deformation of the bridge when the maglev train is passing can be analyzed. On one hand, BP neural network is applied to the 3D model reconstruction of the bridge surface. It monitors the deformation of the bridge from a macro view. On the other hand, the point cloud data is analyzed and a general vibration is calculated from a micro view.

4.1 Surface Reconstruction

Use MATLAB to achieve the surface reconstruction step of the bridge.

The data collected is divided into two parts for training and prediction. Second, set the parameters such as the hidden layer number of network, the maximum error, the maximum number of iterations and so on. Then, weights are obtained by training to make a prediction. If the predicted result does not meet the upper bound of error, retrain the network. Specific training parameters are as follows (Table 4, Figs. 7, 8, 9).

Table 4. Training parameters

Iteration	Time	Error	Gradient	Check
416	0:00:1	0.002	0.000567	0

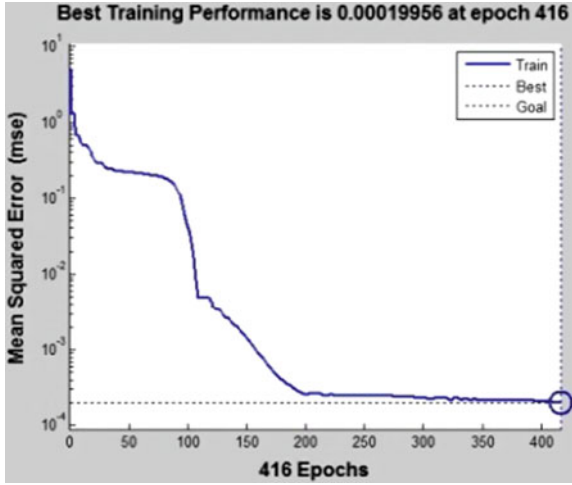


Fig. 7. Training performance

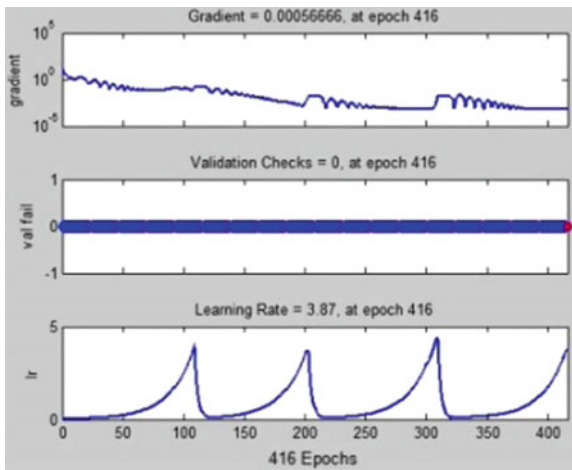


Fig. 8. Training situation

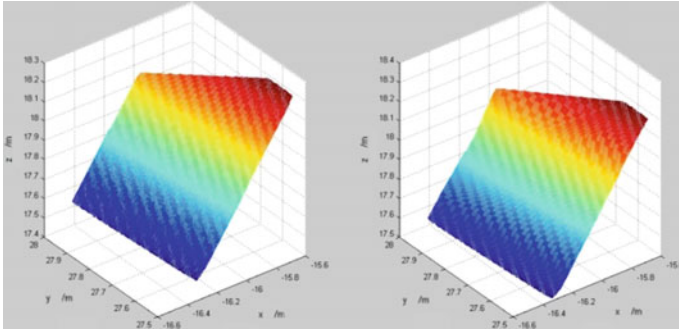


Fig. 9. Training result

The surface reconstruction obtained by using BP neural network will have certain prediction errors. The single-point prediction error is as follows:

$$E_i = z_i - z'_i \quad (7)$$

where $i = 1, 2, \dots, n$, z_i and z'_i represent the elevation before and after the forecast, respectively.

The mean square error of prediction is as Eq. (8) shows, which is as follows:

$$\sigma = \sqrt{\frac{E_1^2 + E_2^2 + \dots + E_n^2}{n}} \quad (8)$$

Through calculation, the maximum relative error of prediction points is 0.1128%. The mean square error is 0.0054 m. The accuracy meets the requirement of deformation monitoring. It proves that by using the surface reconstruction method above mentioned, measurements done in different situations can be compared to achieve the purpose of monitoring the deformation.

The point cloud data of two measurements can be displayed in the same window in RiSCAN PRO software, as shown in the figure below, to observe the change of the data.

As the following Fig. 10 shows, the two different sizes of points represent two different observations.

Due to the error of the rotation speed and measurement accuracy, it cannot be guaranteed that the measured point cloud data can completely coincide when the measured objects do not deform. However, the deviations are within a certain range. By calculating the deviation between the corresponding points, we can observe the change of the bridge when the train passes and no train passes. The deviation is calculated as follows:

$$\Delta d_i = \sqrt{(x_i - x'_i)^2 + (y_i - y'_i)^2 + (z_i - z'_i)^2} \quad (9)$$

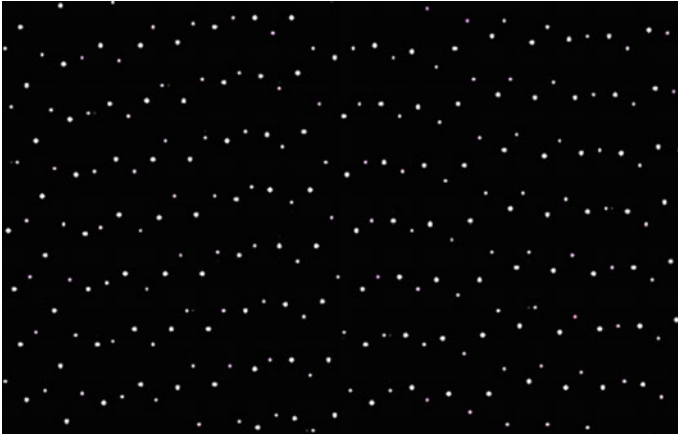


Fig. 10. Measuring contrast

where $i = 1, 2, \dots, n$, x_i , x'_i , y_i , y'_i , z_i and z'_i , respectively, represent the measured coordinates of point cloud data (Table 5).

Table 5. Contrast of the point cloud for bridge

	Average error/m	Average mean square error/m
Contrast of two experiments with no train passing	0.3976	0.2722
Contrast of two experiments with or without train passing	0.4277	0.5678

First, we measured the average error between the corresponding points of the two experiments when no train passing. Then, we calculate the difference between the two experiments. One of the experiments measures the point cloud of the bridge when no train passed and another measure the point cloud when train passing by.

It can be analyzed from the above results that although the scanning of the 3D scanner will produce a certain error, the vibration of the bridge when the train passed by can be detected. In order to get the specific vibration, relevant data should be extracted from the point cloud data for analyzing.

4.2 Vibration Analysis

Based on the above data analysis, point cloud data on the surface of the bridge were extracted at the same intervals and vibration analysis was done. The scanning process is to scan the target object from top to bottom. Then the scanner will rotate a certain angle and scan from bottom to top in the same scanning frequency for data acquisition. Ideally, the scanning interval time of the point cloud data in the same line will be the same, regardless of the accuracy of the instrument rotation angle and scanning frequency.

However, due to the width of the scanned object, the scanning interval is usually not equal to an integral multiple. Therefore, the point cloud data in an ideal situation may have an oblique offset in the row direction.

We choose a line of point cloud data in the horizontal direction to discover the vibration of the bridge. By analyzing these points, when there is no train passing, the obtained point cloud data has some special fluctuations in addition to the above inclination changes. There are various reasons for such fluctuations. For example, the bridge itself will be distorted with the outside world, followed by the instrument and the scanning error. However, by contrast, it can be found that when a maglev train passes, the fluctuation of the point cloud data is larger than before.

At the same interval, the point cloud data is sampled horizontally. Using Fourier transformed to obtain its spectrum. The comparison is as follows (Fig. 11).

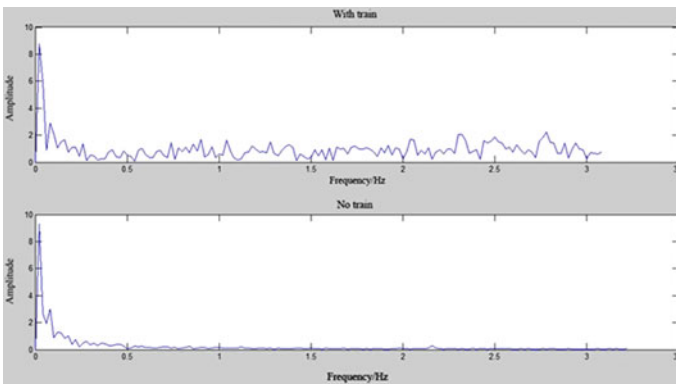


Fig. 11. Spectrum contrast

As can be seen in Fig. 10, when the train passes by, the spectrum of the bridge is more clutter than when there is no car. To a certain extent, this shows that when the train passes, it will affect the bridge (Table 6).

It can also be seen from the above table that through the two experimental data, it is proved that when the maglev train passes, it will have a certain deformation effect on the bridge. At the same time, the bridge itself will have a relatively fixed low-frequency vibration with wind or external influences.

Table 6. Vibration of bridge

	Average amplitude for experiment one	Average amplitude for experiment two
With train passing	0.9596	0.9283
Without train passing	0.2413	0.2712
Difference	0.7183	0.6571

5 Conclusion

In this paper, the deformation of the Shanghai maglev train is monitored. From a macro view, BP neural network is applied for the 3D model reconstruction of the bridge surface. The mean square error for surface reconstruction is 0.0054 m which is acceptable for deformation monitoring. From a micro view, the point cloud data is analyzed and a general vibration is calculated. A conclusion can be drawn that when the maglev train is passing, a certain deformation will affect the train track bridge. A great deal of data of Shanghai maglev train safety information can be got by this method.

However, there are some factors that influence the deformation monitoring and the calculation of the bridge vibration, such as wind, equipment error and so on. These problems remain to be solved and long-term detection is needed for further study. Some supplementary instrument can be added to enhance the accuracy of measuring the deformation for the bridge. Thus, a high-efficiency and high-accuracy deformation monitoring system will be built.

Funding. This work was supported by the Science and Technology Commission of Shanghai Municipality (No. 18511101400).

References

1. Fekete, S., Diederichs, M., Lato, M.: Geotechnical and operational applications for 3-dimensional laser scanning in drill and blast tunnels. *Tunn. Undergr. Space Technol.* **25**(5), 614–628 (2010)
2. Lovse, J.W., Teskey, W.F., Lachapelle, G., Cannon, M.E.: Dynamic deformation monitoring of tall structure using GPS technology. *J. Surv. Eng. ASCE* **121**(1), 35–40 (1995)
3. Roberts, G.W., Meng, X., Dodson, A.H.: Integrating a global positioning system and accelerometers to monitor the deflection of bridges. *J. Surv. Eng. ASCE* **130**(2), 65–72 (2004)
4. Roberts, G.W., Cosser, E., Meng, X.L., Dodson, A.: High frequency deflection monitoring of bridges by GPS. *J. Glob. Position. Syst.* **3**(1), 226–231 (2004)
5. Yigit, C.O., Li, X., Inal, C., Ge, L., Yetkin, M.: Preliminary evaluation of precise inclination sensor and GPS for monitoring full-scale dynamic response of a tall reinforced concrete building. *J. Geodesy* **4**(2), 103–113 (2010)
6. Park, H.S., Lee, H.M., Adeli, H., Lee, I.: A new approach for health monitoring of structures: terrestrial laser scanning. *Comput. Aided Civ. Infrastruct. Eng.* **22**(1), 19–30 (2007)
7. Badenko, V., Zotov, K., Zotov, D., Garg, R.D., Zhang, L., Bolsunovskaya, M., Fedotov, A.: Laser scanner survey technologies for historic building information modeling of heritage resources in Saint-Petersburg, Russia. *Constr. Unique Build. Struct.* **1**(52), 93–101 (2017)
8. Luo, D.A., Zhu, G., Lu, L., Liao, L.Q.: Whole object deformation monitoring based on 3D laser scanning technology. *Bull. Surv. Mapp.* **36**(12), 192–198 (2005)
9. Buckley, S.J., Howell, J.A., Enge, H.D., Kurz, T.H.: Terrestrial laser scanning in geology: data acquisition, processing and accuracy considerations. *J. Geol. Soc.* **165**(3), 625–638 (2008)

10. Gu, P., Yan, X.: Neural network approach to the reconstruction of free-form surfaces for reverse engineering. *Comput. Aided Des.* **27**(1), 59–64 (1995)
11. Zhang, L., Lu, J., Liang, X.: Forest fire forecasting based on a fuzzy-neural network. *J. Tsinghua Univ. Sci. Technol.* **50**(8), 1302–1306 (2010)
12. Liu, J.F., Zhang, L., Wang, Y.F., Shu, J., Shu, R.: Three-dimensional information fusion technology based on vehicle-borne laser scanning. *Sci. Surv. Mapp.* **37**(3), 174–177 (2012)



Access Control System Based on Visible Light Communication

Xinpeng Xue, Jinpeng Wang^(✉), Ying Yu, and Nianyu Zou

School of Information Science and Engineering, Dalian Polytechnic University,
No. 1st Qinggongyuan, Ganjingzi, Dalian 116034, Liaoning, China
wangjp@dlpu.edu.cn

Abstract. In recent years, with the continuous development of Visible Light Communication (VLC), its application fields are constantly expanding. In view of the tradition various gating locks have some shortcomings, especially in terms of safety needs to be improved. The visible light communication technology and access control system are joined to design a Hamming-encoded visible light communication access control system in this paper. The hardware design of this access control system is given in this paper and the system of Hamming coding is analyzed. Through MATLAB software simulation, the results show that the access control system with Hamming code can effectively reduce the bit error rate, and can provide about 2 dB coding gain. In the experiment, the system works well just like simulation. Hamming code greatly improves the performance and safety of the visible light communication access control system.

Keywords: Visible light communication · Access control system · Hamming code

1 Introduction

Access control system is defined as a system which has functions of management and control for entrances and exits [1]. In the past, access control systems were mechanical door locks, which were purely mechanical devices [2]. Therefore, no matter how the design was adopted, no matter how hard the material is, the lock could be opened by external forces, and the improper management of keys could also cause unnecessary lost. In order to overcome this problem, electromagnetic card swipe access control and code lock appeared. Although the lack of mechanical locks is well solved by electromagnetic cards and code locks, the fast wear of electromagnetic cards and the low security of code locks are still problems that cannot be ignored. Based on this, the wireless radio-frequency technology is derived, which is widely used [3]. However, the penetration of this technology is very high, it may cause the copy and intercept of the unlocking information, causing information leakage, and property loss [4]. With the vigorous development of the communication industry and LED material devices in recent years, a brand new technology, visible light communication technology, has entered our lives. Visible light communication technology uses light waves as a carrier

for information and free space as a transmission channel for information transmission [5]. LED devices have good frequency response characteristics, laying the foundation for visible light communication on devices [6]. Visible light communication uses these features of the LED, which allows LEDs to transmit information by emitting high-speed flashes that are not perceived by the human eye, and uses optical sensors to receive light information [7]. Visible light communication has the following advantages: Does not occupy today’s increasingly strained spectrum resources; Visible light communication is a point-to-point communication method, and it is not subject to electromagnetic interference [8]. Therefore, it is more safe than before. Today’s wireless RF access control has poor anti-electromagnetic interference and passwords are easily intercepted [9]. The advantages of visible light communication technology can effectively overcome these problems. Therefore, a new type of access control system based on visible light communication can be designed by combining visible light communication technology with access control.

All in all, in order to solve the various shortcomings and security problems of traditional access control systems, the paper designed a kind of access control system based on visible light communication. The system used STC89C52RC single-chip as a data transmission/reception processing unit, and used visible light as a carrier for transmission of information. Using IM/DD (Intensity Modulation/Direct Detection) modulation method, the encoding information was loaded onto the visible light to achieve the transmission of identity information. The encoding method selected Hamming error correction coding. The simulation results showed that the use of the Hamming codec after the access control system The bit error rate was significantly improved and could provide about 2 dB of coding gain. Furthermore, the system worked well in experimental tests just like simulation.

2 Overall System Design

The working principle of the visible light communication system is shown in Fig. 1.

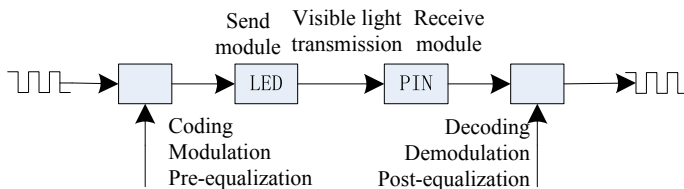


Fig. 1. Working principle diagram of visible light communication system

The visible light communication access control system mainly consists of an optical key terminal (a sending terminal of information) and a access control terminal (a receiving terminal of information). The optical key terminal is composed of a single

chip and an LED driving circuit. The single chip is responsible for storing the user’s identity information and verifying whether the password information can be passed through the access control. The LED driver circuit is responsible for loading information onto the LED for transmission. The access control terminal is responsible for receiving visible light and recording information such as time and frequency. Then the access control terminal judges whether the received information is correct by the database. Finally, the access control terminal will alarm until it receives the right key. The access control terminal is composed of a receiving circuit and a single-chip. The system block diagram is shown in Fig. 2.

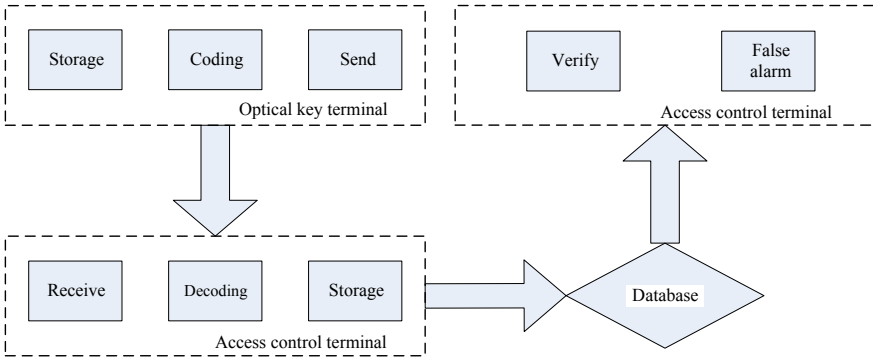


Fig. 2. The block diagram of system

3 System Software Design

3.1 Hamming Error Correction Code

Channel coding means that artificial redundancies are added to the code in accordance with certain rules [10]. This way can reduce the error rate of transmitted information, and even to improve the reliability of information systems without transmission errors [11]. It is widely used by modern communication systems. Channel coding technology can effectively improve the link performance of visible light communication systems. Common channel error correction codes include linear block codes, cyclic codes, convolutional codes and trellis codes [12]. Hamming error correction codes belong to a class of linear block codes that can correct one wrong code. Then briefly introduce the principle of encoding and decoding Hamming code, Hamming code satisfies Formula (1).

$$2^r - 1 \geq n \tag{1}$$

where n represents the total number of bits of the symbol and r represents the number of bits of the supervision symbol. The construction principle of the Hamming code is similar to the even supervision code. The supervised code adds a supervising symbol

after the existing information symbol so that the number of 1 in the total information symbol is even. It can be expressed as Formula (2).

$$a_{n-1} \oplus a_{n-2} \oplus \dots \oplus a_1 \oplus a_0 = S \tag{2}$$

Because the number of 1 is even, there must be $S = 0$ when the supervised code is transmitted accurately. In this way, it is possible to determine whether there is an error in the transmission by calculating the value of S .

There is no error at $S = 0$ and there is something wrong at $S = 1$. Formula (2) may be referred to as a supervising equation, and S may be referred to as a syndrome. In accordance with this idea, one can continue to add a supervising symbol to construct the supervising equation. At this time, two syndromes will appear. There are four cases in the combination of these two syndromes: 00,01,10,11. The 00 can be used to indicate that the transmission has experienced an error. The remaining three conditions are used to indicate the location of different error codes. Therefore, in order to indicate the condition of all one-bit errors in the bit symbol, Formula (1) must be satisfied: The introduction of r supervised symbols can have 2^r conditions. All zeros are used to indicate that the transmission has errors. The remaining $2^r - 1$ cases are used to represent $2^r - 1$ errors. If $2^r - 1 \geq n$, all n -bit single errors can be indicated. If H is set as the supervision matrix, the corresponding transmission vector of B is A , and if $E = B + A$ is an error pattern, then there is Formula (3).

$$S = BH^T = (A + E) \cdot H^T = AH^T + EH^T \tag{3}$$

Known $AH^T = 0$, so $S = EH^T$. S in Formula (3) is the abovementioned syndrome, and the syndrome can be used to judge whether the transmission has an error and perform corresponding error correction.

3.2 Software Process

Software flowchart is shown in Fig. 3. The optical key terminal single-chip completes the initialization work first and encodes the password information, the LED waits for the data transmission. The LED sends information when receiving the data sending instruction. The access control terminal uses a photodiode for receiving, and the photodiode can sense light with different brightness changes, so that the voltage across the photodiode is different. The voltage is followed and compared through the LM324 operational amplifier, and then the voltage is changed to the standard high and low-level output to the access control terminal of the single-chip for demodulation and decoding to restore the original data. Finally, the restored data is compared with the identity information stored in the access control terminal single-chip, then the access control terminal will alarm until it receives the right key.

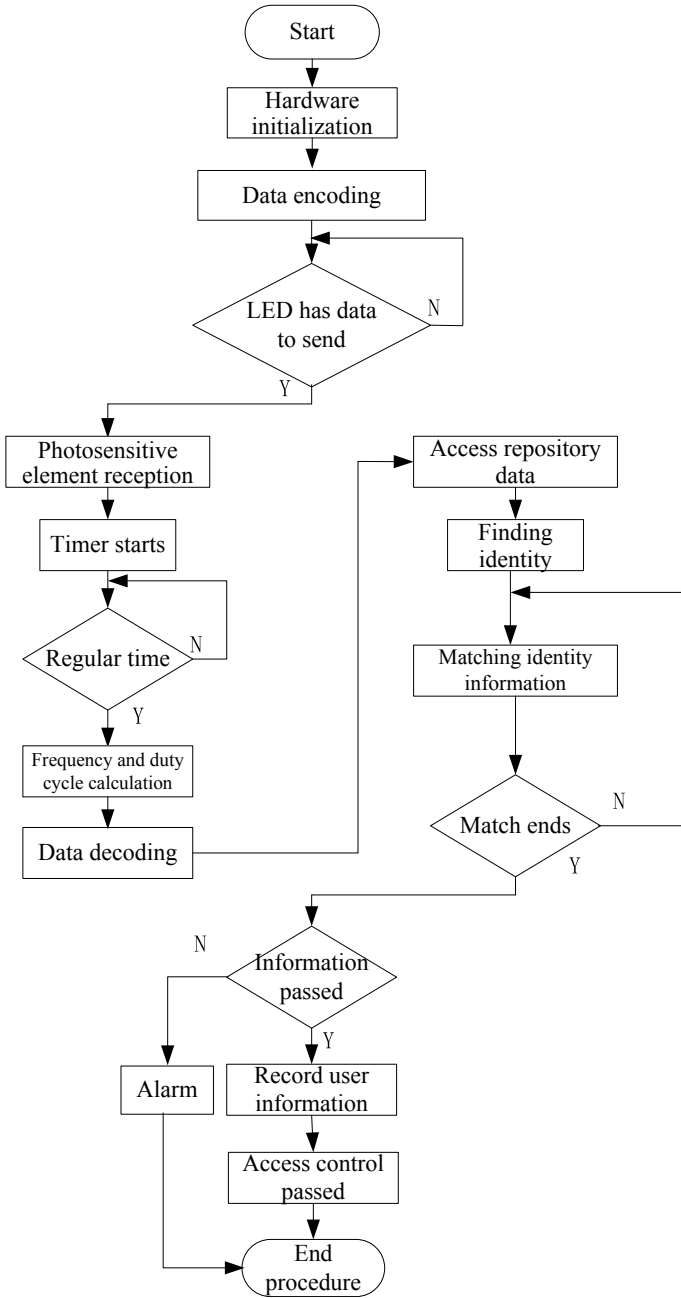


Fig. 3. Flowchart of system software

4 System Hardware Design

4.1 Single-Chip Minimum System

This design uses STC89C52 single chip as the core processor of the control part. The STC89C52 uses the classic MCS-51 core, 8K bytes of Flash, 512 bytes of RAM, 3 16-bit timers, and the serial port is a full-duplex serial port, watchdog timer, two external interrupt ports, and 32-bit IO ports. In order to meet the function of this system, two internal timers T0, T1 of STC89C52, external interruption, serial port breaks and serial port communication function are used.

The main function of two of the internal timers is to generate accurate timing and to calculate the pulse width modulation frequency and the duty cycle of the visible light. The minimum system of the single chip is shown as in Fig. 4.

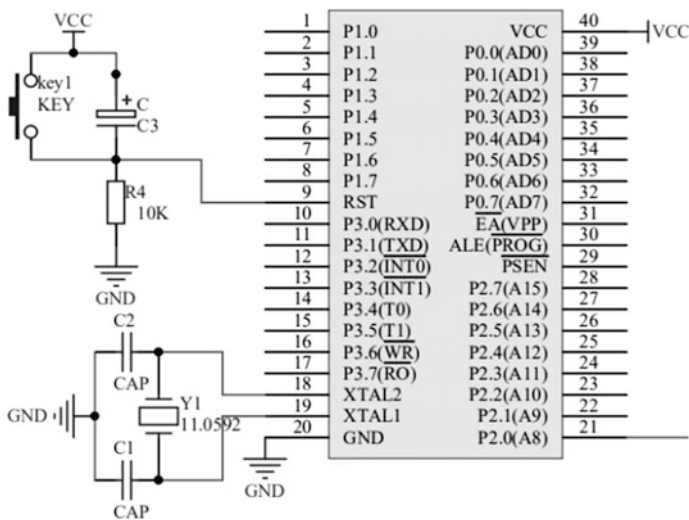


Fig. 4. STC89C52 single-chip microcontroller system

In the clock circuit, a classical external 11.0592M quartz crystal oscillator is used. Generally, the two ends of the crystal are connected with a resonance capacitor, which can enhance the anti-interference ability of the external oscillation circuit and make the oscillation capacitor start up more easily.

The single chip adopts the high-level reset method, which can trigger the single-chip reset function when at least two machine cycle high-level signals are applied to the reset pin. In order to maintain the reliable operation of the system and the convenience of debugging, the system adopts power-on automatic reset and can be reset manually if necessary.

4.2 Optical Key Terminal and Access Control Terminal

The optical key terminal is the LED driving circuit, in which the LED driving circuit is driven by a high-frequency FET, the I/O port is connected to the gate, and the gate is controlled by the I/O port. The information is sent out by controlling the duty cycle of the LED. The optical key terminal is shown in Fig. 5.

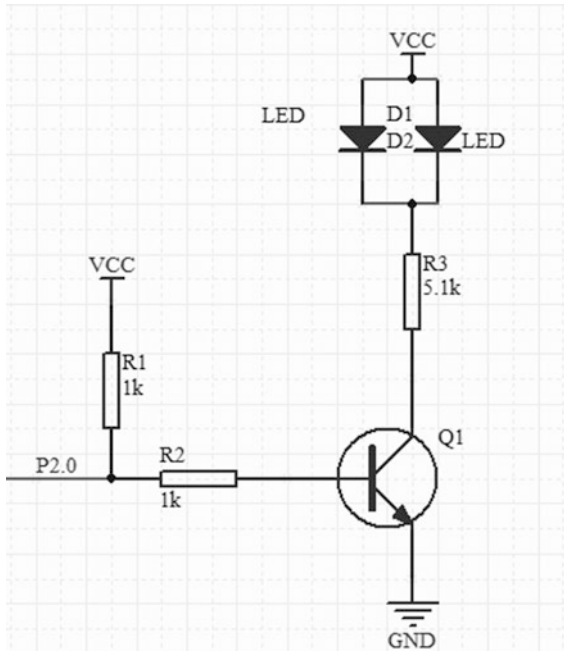


Fig. 5. Optical key terminal

The access control terminal is shown in Fig. 6. The access control terminal receives the optical signal transmitted by the optical key terminal LED through the photodiode and converts the optical signal with different duty cycles into an electrical signal. Then the signal is sent to the LM324 operational amplifier for signal amplification and voltage follow-up and comparison. Then the signal will be converted into The standard level and is sent to the single-chip P3.2 port. The P3.2 port of the STC89C52 is an interrupt trigger port, trigger timing through the input signal, and then can calculate the carrier frequency and pulse width of the optical signal and demodulate the relevant data from the coded information.

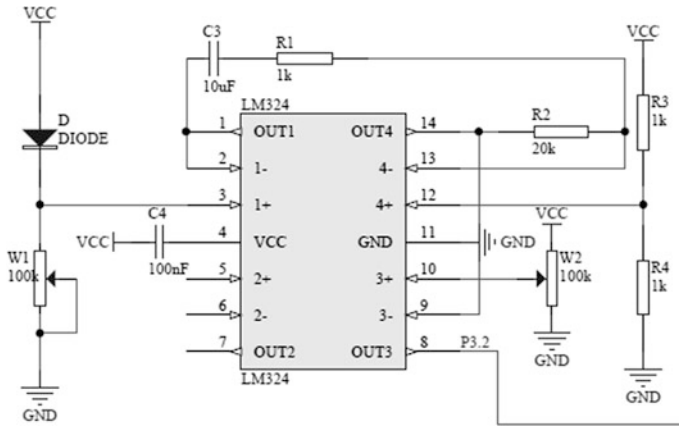


Fig. 6. Access control terminal

5 System Simulation and Experiment

According to the above design of each module, the access control system model is shown in Fig. 7.

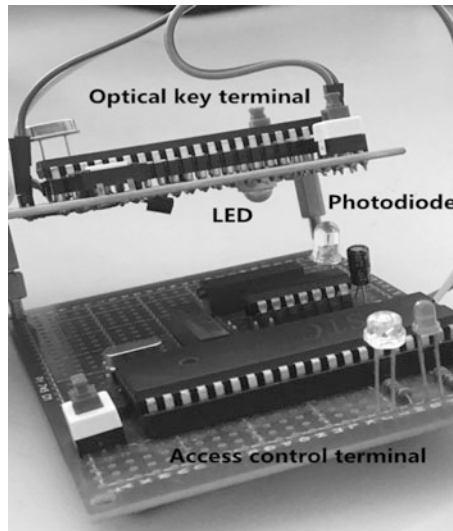


Fig. 7. Access control system

In order to verify the effectiveness of the Hamming error correction code, a system model must be constructed according to the actual situation. First, the channel of the design is analyzed. Visible light communication is mainly divided into line-of-sight

links and non-line-of-sight links. For the visible light access control system of this paper, the communication distances between the optical key terminal and the access control terminal are short and are almost unaffected by the multipath effect. Therefore, the line-of-sight link is adopted. The line-of-sight link can also be divided into a directional line-of-sight link and a nondirectional line-of-sight link. The difference between the two is that the line-of-sight link has high requirements on the angles of the sending terminal and the receiving terminal, in order to receive information accurately it must be almost aligned. After the actual hardware production and debugging, this design adopts directional line-of-sight links, and the effect of transmitting information is relatively good at this time.

We assume that the paper’s indoor visible light communication system model is in the environment of the AWGN (additive white Gaussian noise) channel model. Therefore, in the case of considering only Gaussian white noise, the Hamming [1, 2] code is programmed and the error rate is simulated and analyzed. The simulation results are shown in Fig. 8.

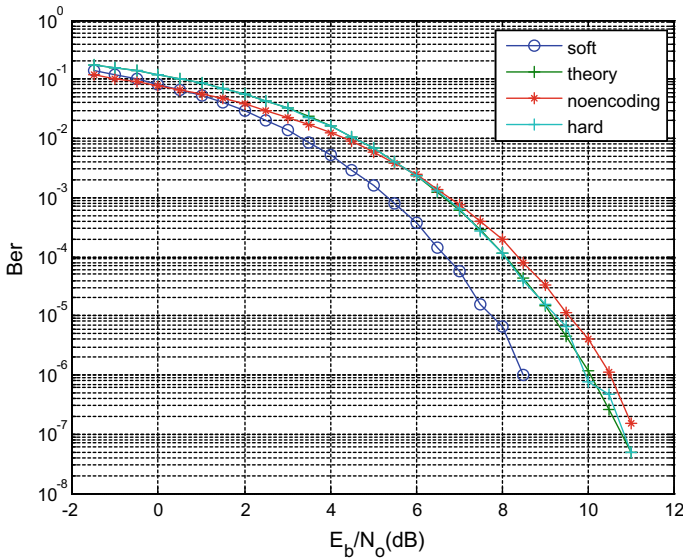


Fig. 8. Using Hamming [1, 2] coding performance curve

From the simulation result curve, it is found that the introduction of Hamming error correction code brings an SNR gain compared to the uncoded system, and as the SNR increases, the BER of the system shows a decreasing trend. The bit error rate is lower when the noise is not encoded and after Hamming [1, 2] coding. As the signal-to-noise ratio increases, the error correction coded system can achieve a lower bit error rate. Soft-decision can get nearly 2 dB of coding gain compared to un-decoded, which can further improve the performance of the access control system.

6 Conclusion

In view of the fact that the current radio frequency technology information is easily intercepted and replicated, and the rapid development of visible light communication technology, a visible light communication access control system based on the Hamming [1, 2] code is designed. First, the system hardware design is given. Then analyze the performance of the access control system using Hamming coding. The results show that the introduction of Hamming error correction code in the codec effectively reduces the error rate of the access control system and brings about 2 dB of coding gain, and the experiment shows that system coordinate with the simulation, further improving the system performance, increasing the cost of cracking, and improving the security. Under the background of the vigorous development of visible light communication technology in recent years, the method designed in this paper has certain theoretical significance and practical application value. Visible light communication technology will have a very valuable application prospect in the field of access control.

Acknowledgements. This research was financially supported by Project of the National Natural Science Foundation of China (61402069), “the Fundamental Research Funds for the Central Universities” (3132016317), 2017 Project of the Natural Science Foundation of Liaoning province (20170540059), General project of Liaoning education department in 2016 (2016J205).

References

1. Zhao, J., Li, Y., Zhang, Y., Zou, N., Wang, J.: A supplementary lighting system for plant growth with lighting-emitting diode based on DT TS & IC **61**(7), 548–551
2. Chi, N., Zhang, M., Shi, J., Zhao, Y.: Spectrally efficient multi-band visible light communication system based on Nyquist PAM-8 modulation. *Photonics Res.* **5**(06), 107–116 (2017)
3. Barry, J.R.: *Wireless infrared communications*. Springer, Berlin **85**(2), 265–298
4. Cheng, R., Yan, X.: Indoor multisource channel characteristic for visible light communication. *J. China Univ. Posts Telecommun.* **20**(04), 106–111 (2013)
5. Wang, J., Zou, N., Zhang, Y., Li, P.: Study on downlink performance of multiple access algorithm based on antenna diversity. *ICIC Express Lett.* **9**(4), 1221–1225 (2015)
6. Jovicic, A., Li, J., Richardson, T.: Visible light communication: opportunities, challenge and the path to market. *IEEE Commun. Mag.* **51**(12), 26–32 (2013)
7. Tahir, M., Siddique, A.B.: Optimal brightness rate control using VR-MPPM and its spectral analysis for VLC system. *IEEE Commun. Lett.* **16**(7), 1125–1128 (2012)
8. Shiv, D., Kahn, J.M.: Differential pulse-position modulation for power-efficient optical communication. *IEEE Trans. Commun.* **47**(8), 1201–1210 (1999)
9. Kim, S.M., Lee, H.J.: Visible light communication based on space-division multiple access optical beamforming. *Chin. Opt. Lett.* **12**(12), 14–17 (2014)
10. Liu, Y., Zheng, C.T.: Portable 100 Mbps point-to-point OOK-NRZ visible light communication devices based on white light-emitting diode illuminant. *Microw. Opt. Technol. Lett.* **54**(10), 2248–2252 (2012)
11. Yuan, M., Sha, X., Liang, X., Jiang, M., Wang, J., Zhao, C.: LDPC decoding for signal dependent visible light communication channels. *ZTE Commun.* **14**(02), 41–46 (2016)
12. Ziaei, N., Rafiee, A.: *Int. J. Adv. Res. Comput. Sci.* **5**(1), 18–20 (2014)



Portable 3D Laser Scanner for Volume Measurement of Coal Pile

Wang Zhang^{1(✉)}, Deshan Yang^{2(✉)}, Ying Li^{1,2}, and Wenhai Xu^{1,2}

¹ Dalian Maritime University, Linghai Road, Dalian, China
zhangwang@dlmu.edu.cn

² Jiaxing University, South Yuexiu Road, Jiaxing, China
deshan.yang@foxmail.com

Abstract. In order to improve the efficiency and accuracy of volume measurement of the large coal pile, a set of portable 3D laser scanner is designed. This device uses the 2D laser scanner as the measurement module supplemented by the high-accuracy sensor, to scan the 3D morphology surrounding the coal pile. Based on characteristics of the scanner and structural features of the coal pile, the data preprocessing arithmetic is optimized, which synchronously completes the scanning point noise reduction during the measurement process and saves the time required for the 3D modeling. Spatial interpolation is adopted to reconstruct the 3D model, which fills in the cavity at the coal pile top, filters excessive scanning points at the bottom, and improves the efficiency of the 3D modeling and volume calculation. As shown by the experiments, only a few minutes are required to complete the 3D morphology measurement of the coal pile, and it will generate the 3D model and volume within 20 s after the measurement.

Keywords: 3D laser scanner · Volume measurement · 3D modeling · Noise reduction · Spatial interpolation

1 Introduction

The coal storage base is significant for strengthening the stable coal supply capacity and guaranteeing national energy safety. The rapid and accurate check to the coal reserve is important for relevant enterprises to perform production accounting and planning. The main method for the coal reserve check is the volume-density method, which means to establish the 3D model of the coal pile to obtain the volume, and then calculate the mass by multiplying the volume by the density. Currently, the multi-base point laser ranging method is widely used in measuring the 3D morphology of the coal pile. The accuracy of the 3D model has a close relationship with the quantity of base points and measurement points. Therefore, the requirement of more accurate volume measurement will result in proportionally increased workload [1].

In order to improve the efficiency and accuracy of volume measurement, the 2D laser scanner is used in the 3D modeling of the coal pile [2]. Generally, the scanner is fixed at the front end of the cantilever of a bucket-wheel stacker-reclaimer [3, 4]. Many scanning points in the scanning line can be obtained via once scanning, and the accuracy of the 3D modeling will be obviously improved. However, the cooperative operation of the stacker-reclaimer is required, and the limit resulted from the installation location often generate the scanning dead zone. For this purpose, this paper designs and realizes one set of a portable 3D laser scanner for the volume measurement of the coal pile. It takes the 2D laser scanner to scan 3D morphology, and integrates independent scanning points to form a 3D point cloud under the cooperation of the location sensor and attitude sensor. The operator only needs to carry the scanner to circle around the coal pile, and then the 3D model and volume can be obtained.

2 Measurement Principle

The principle for measuring the volume of the coal pile with a portable 3D laser scanner is as shown in Fig. 1. The scanner will project scanning lines on the coal pile surface, which is composed of many scanning points with equal angle-interval. Along with the operator carries the scanner to scan surrounding the coal pile, the scanning points arranged in the linear form will cover the surface of the coal pile in order. For any scanning point P on the coal pile, assumed that its coordinate in the local coordinate system $O^L-X^LY^LZ^L$ established for the original point on the scanner is expressed as (x^L, y^L, z^L) , and the coordinate of the scanner in the 3D modeling of the global coordination system $O^G-X^GY^GZ^G$ is expressed as $(\Delta X, \Delta Y, \Delta Z)$, the coordinate of the scanning point P in the global coordination system can be expressed as (x^G, y^G, z^G) .

$$\begin{bmatrix} x^G \\ y^G \\ z^G \end{bmatrix} = \lambda R \begin{bmatrix} x^L \\ y^L \\ z^L \end{bmatrix} + \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} \tag{1.1}$$

where $\lambda = 1$ is the scale factor between two coordinate systems, and R is the rotation matrix related to the scanner rotation angle $(\varepsilon_x, \varepsilon_y, \varepsilon_z)$ [5]. Through the coordinate transformation Eq. (1.1), all scanning points on the coal pile surface can be mapped to the global coordination system, so as to integrate and form a 3D point cloud describing the morphology of the coal pile. Based on the point cloud, spatial interpolation will be carried out to form a vivid 3D model, and then divide the model into numerous prisms to calculate the volume.

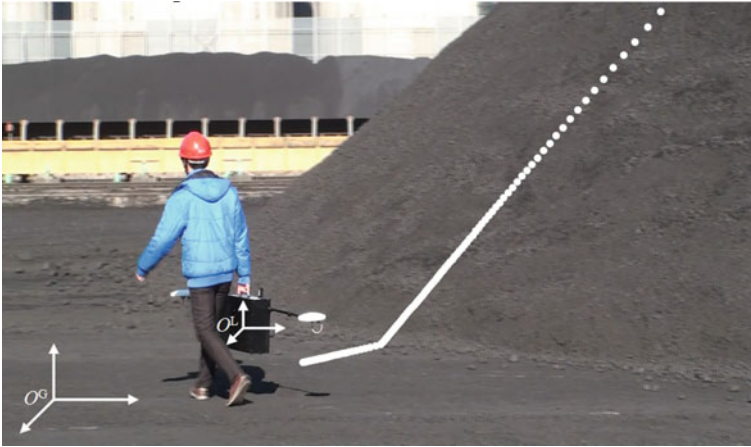


Fig. 1. Principle for measuring the volume of the coal pile

3 Data Processing

The main task for the data processing is to integrate the 3D point cloud of coal pile formed by discrete scanning points, and further to realize the 3D modeling and volume calculation. However, the preprocessing of measurement data before the integration is also necessary which mainly aims to remove noise points generated from the laser scanner due to environmental disturbance.

3.1 Scanning Point Noise Reduction

It is almost inevitable that the laser scanner for the field environment will generate noise disturbance during the measurement [6, 7]. Although the multi-echo detection technology of the LD-MRS400001 scanner can effectively reduce the noise brought by the window plate, misty rain and dust, it will generate random noise points in direct sunlight sometimes, which is inevitable for the scanner measuring surrounding the coal pile. In addition, the huge bucket-wheel stacker-reclaimer and other background objects in the storage yard will also generate continuous noise points. The scanning point noise reduction is generally carried out after generating the 3D point cloud, mainly in the observation method, fitted curve method, chord height method, and probability-based method [8]. Some of them require human-computer interaction, some involve complicated algorithms, and some have poor adaptability [9]. That makes them difficult to meet the application demand for rapid modeling and volume calculation of the coal pile. This paper investigates the original measurement data obtained by the scanner, and summarizes main expression forms of noise in scanning lines, as shown in Fig. 2. Based on this, the noise points in each scanning line are removed during the scanning process, thereby effectively improving the speed of 3D modeling.

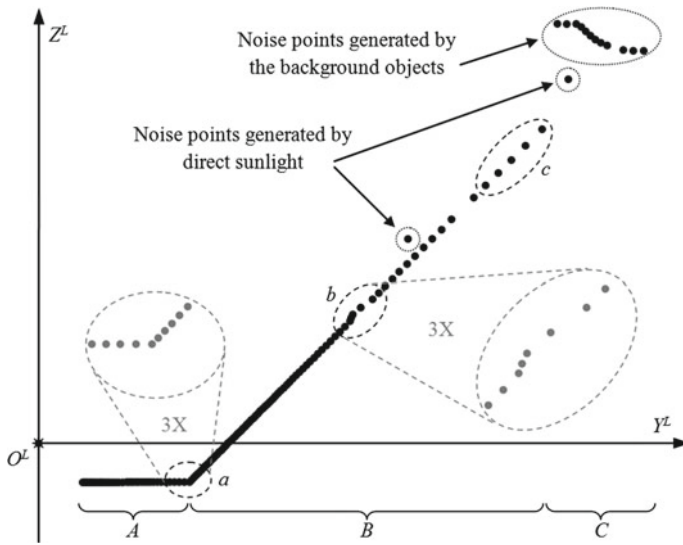


Fig. 2. Expression forms of noises in scanning lines

Generally, one scanning line can be divided into three sections. Section A is the scanning points projected on the ground, few noise points will occur in this area where is slanting underneath the scanner; Section B is the scanning points projected on the coal pile surface, noise points generated from the direct sunlight will occur in this area randomly; Section C is the continuous noise points projected on the stacker-reclaimer and other huge background objects, including noise points generated due to the direct sunlight. The noise points generated due to the direct sunlight presents the obvious mutability. Therefore, the random mutational points can be easily recognized among scanning points nearly continuously distributed, and the noise reduction can be realized by judging the spacing between adjacent scanning points:

$$P_n = S_n, \quad |S_n S_{n+1}| \leq T \tag{1.2}$$

where S_n is the scanning point measured by the scanner, T is the determination threshold, and P_n is the scanning point determined to be valid.

It is easy and rapid to reduce noises by using the fixed threshold as the criterion. However, the experiments show that it is difficult to choose a proper determination threshold T . If T is too large, the reduction effect of the noise points will be influenced. If T is too small, valid scanning points in three parts as shown in Fig. 2 will be removed as noise points. Because the scanning line is composed of scanning points with equal angle-interval, the spacing in Part a will increase due to angle change; the spacing in Part c will increase due to distance change; in addition, the spacing in Part b will increase due to occasional irregular protrusions on the coal pile surface. Therefore, it is necessary to correspondingly adjust the determination threshold value. Considering that there is a continuity between scanning point S_n and S_{n-1} or S_{n+1} , which is obviously

different with the abruptness exhibited by the random noise points, the spacing between scanning point S_{n-1} and S_{n+1} can be used as an additional criterion to optimize the process:

$$P_n = \begin{cases} S_n, & |S_n S_{n+1}| \leq T \\ S_n, & |S_{n-1} S_n| \leq |S_{n-1} S_{n+1}| \\ S_n, & |S_n S_{n+1}| \leq |S_{n-1} S_{n+1}| \end{cases} \quad (1.3)$$

As shown by the experiments, the method is effective in reducing noise points generated by the direct sunlight. But it is invalid for reducing noise points generated by the stacker-reclaimer and other background objects. These background objects will form continuous scanning points in Section C which makes it difficult to distinguish scanning points generated by the coal pile in the single scanning line. Fortunately, the cantilever of stacker-reclaimer is several meters higher than the coal pile, and a sufficient distance is kept between other background objects and the coal pile. Therefore, after the 3D point cloud of the coal pile is created, the background noise points can be removed by setting the top and bottom of the point cloud bounded by most scanning lines as the boundaries.

3.2 3D Point Cloud Creation

It is the basis for performing the subsequent 3D modeling and volume calculation to integrate discrete scanning points to a 3D point cloud [10]. The main method for the integration is to map the scanning points to the local coordinate system $O^L-X^L-Y^L-Z^L$ with the scanner as the origin to the global coordinate system $O^G-X^G-Y^G-Z^G$ used for 3D modeling. For the convenience of mapping, the original point of the local coordinate system is selected at the light outlet of the scanner, and overlap the $Y^L-O^L-Z^L$ plane with the scanning plane of the scanner. As for the n th scanning point P_n^m in the m th scanning line obtained by the portable 3D laser scanner, its coordinate $(x_n^{Lm}, y_n^{Lm}, z_n^{Lm})$ in the local coordinate system is simplified as

$$\begin{bmatrix} x_n^{Lm} \\ y_n^{Lm} \\ z_n^{Lm} \end{bmatrix} = \begin{bmatrix} 0 \\ l_n^m \cos(\phi + n\Delta\phi) \\ l_n^m \sin(\phi + n\Delta\phi) \end{bmatrix} \quad (1.4)$$

where l_n^m is the distance between the scanning point P_n^m and the outlet of the scanner, ϕ is the initial scanning angle of the scanner, and $\Delta\phi$ is the increment of the scanning angle.

When the scanner emits the m th scanning line, its coordinate in the global coordination system $O^G-X^G-Y^G-Z^G$ is $(\Delta X^m, \Delta Y^m, \Delta Z^m)$, and the rotation angle of the coordinate axis is $(\varepsilon_X^m, \varepsilon_Y^m, \varepsilon_Z^m)$. According to Eq. (1.1), the coordinate $(x_n^{Gm}, y_n^{Gm}, z_n^{Gm})$ of scanning point P_n^m in the global coordination system can be written as

$$\begin{bmatrix} x_n^{Gm} \\ y_n^{Gm} \\ z_n^{Gm} \end{bmatrix} = \lambda R \begin{bmatrix} x_n^{Lm} \\ y_n^{Lm} \\ z_n^{Lm} \end{bmatrix} + \begin{bmatrix} \Delta X^m \\ \Delta Y^m \\ \Delta Z^m \end{bmatrix} = \lambda R \begin{bmatrix} 0 \\ l_n^m \cos(\phi + n\Delta\phi) \\ l_n^m \sin(\phi + n\Delta\phi) \end{bmatrix} + \begin{bmatrix} \Delta X^m \\ \Delta Y^m \\ \Delta Z^m \end{bmatrix} \quad (1.5)$$

where the coordinate $(\Delta X^m, \Delta Y^m, \Delta Z^m)$ of the scanner is measured by the location sensor, the rotation angle $(\varepsilon_X^m, \varepsilon_Y^m, \varepsilon_Z^m)$ deciding the value of rotation matrix R is measured by the attitude sensor. Based on this, the scanning points in the m th scanning line can be mapped to the global coordination system. Along with the progress of the measurement, the obtained scanning lines are mapped to the global coordination system in order, so as to form the 3D point cloud shown in Fig. 3, which can represent the surface morphology of the coal pile.

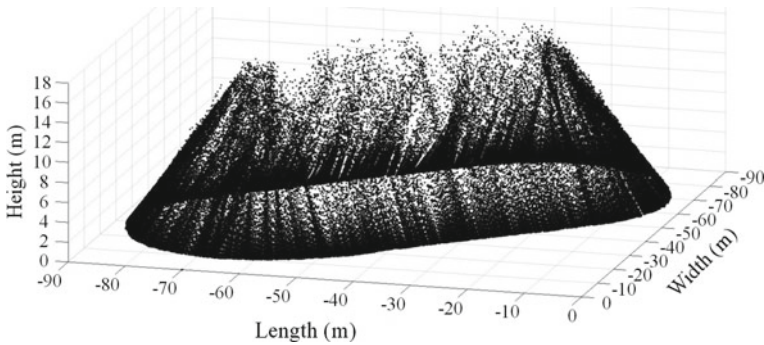


Fig. 3. 3D point cloud of the coal pile

3.3 3D Modeling and Volume Calculation

The common method for the 3D modeling based on the point cloud is triangulation which describes the surface morphology in the triangular mesh form [11, 12]. It can accurately express the morphological characteristics of the curve surface, and offer facilities for the subsequent volume calculation. The experiments show that the method based on triangulation is excellent in describing morphological details of the coal pile surface. However, the top of the 3D model constructed by it has some cavities, and the modeling speed is not ideal. By analyzing the distribution of the points in the 3D point cloud, it can be found that the scanning points at the coal pile top are sparse, while the points at the bottom are dense. As for the reason, the scraggly structure is formed when stacking the coal pile, and the increase of scanning distance results in the increment of spacing between scanning points. Therefore, when using the point cloud obtained by the portable 3D laser scanner to construct the 3D modeling, it is necessary to reconstruct the point cloud by spatial interpolation. It can not only fill up cavities at the coal pile top, but also filter excessively dense scanning points at the bottom, which makes the modeling more vivid and rapid. Considering there is no overlap of the coal pile surface in the height direction Z^G , firstly the point cloud is projected to the $X^G O^G Y^G$

plane, and then the spatial interpolation in the square mesh form will be carried out. Finally, the construction result is mapped back to the $O^G-X^GY^GZ^G$ space, and the rapid 3D modeling as shown in Fig. 4 will be realized.

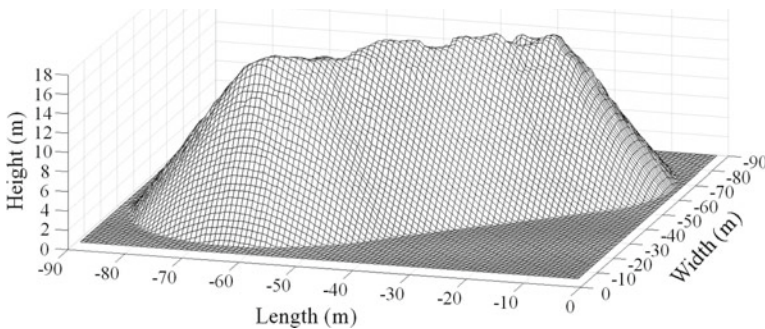


Fig. 4. 3D model of the coal pile

The projection is divided into numerous squares in the $X^GO^GY^G$ plane, the 3D model of the coal pile is divided into numerous prisms in the $O^G-X^GY^GZ^G$ space correspondingly, and then the coal pile volume will be obtained by summing over the volume of every prism. In addition, the quantity of prisms is significantly reduced because of space correspondingly, thereby effectively improving the speed of volume calculation. The field experiment was conducted at the coal storage base in Huanghua Port. The portable 3D laser scanner was used to measure several 3,000–30,000 m^3 coal piles, and the results proved that the measurement accuracy can be less than 0.8%.

4 Conclusion

The portable 3D laser scanner uses the 2D laser scanner to measure the morphology of the coal pile, which can project the scanning line composed of many scanning points on the coal pile surface once. Along with the operator carries the scanner to scan surrounding the coal pile, scanning points arranged in the linear form covers the whole surface in order. At the same time, the high-accuracy sensor is used to track the location and attitude of the scanner, to form the 3D point cloud of the coal pile by the coordinate transformation. Based on measurement characteristics of the scanner and structural features of the coal pile, this paper optimizes the data preprocessing arithmetic, synchronously completes the scanning point noise reduction during the measurement process, and obtains the ideal 3D point cloud when completing the measurement. Combining with the distribution of scanning points in the point cloud, this paper optimizes the 3D modeling arithmetic to fill in the cavity at the model top and filter excessive scanning points at the bottom. That makes the modeling more vivid and rapid, improves the speed of volume calculation, and shortens the waiting time for data processing to less than 20 s after completing the measurement. According to the measurement results of several 3,000–30,000 m^3 coal piles at the coal storage base in

Huanghua Port, the volume measurement accuracy is less than 0.8%, which can meet the demand of the coal storage base regarding the rapid and accurate check of coal reserve.

References

1. Dong, D.Y., Yang, L.M., Yuan, Z.W.: The inventory method and the inventory-making calculation of profit and loss in coal yard. *Coal Qual. Technol.* **190**, 13–15 (2014)
2. Zhang, B., Liu, C.L., Liu, Q., et al.: Application of laser disc technology of coal in power plant coal storage yard management. *Laser J.* **35**(5), 53–58 (2014)
3. Chen, S., Wang, H.B., Lu, G.D., et al.: Application of digital technology to fuel management in fossil power plant. *Electr. Power Surv. Des.* **96**, 56–60 (2016)
4. Li, C.A.: Study on windrow modeling methods for bulk yard. *Metall. Ind. Autom.* **41**(2), 1–5 (2017)
5. Annich, A., Abderrahmani, A.E., Satori, K.: Fast and easy 3D reconstruction with the help of geometric constraints and genetic algorithms. *3D Res.* **8**, 30 (2017)
6. Zhang, W.J.: 3D laser scanning technology and its application. *Stand. Surv. Mapp.* **32**(2), 42–44 (2016)
7. Jian, X., Zhou, Q., Huang, H.: Application of 3D laser scanner to 3D modeling of bulk material yard. *Port Oper.* **238**, 32–35 (2018)
8. Zhu, Q.W., Ma, Y.J.: Application and research of building modeling based on 3D laser scanner. *Geogr. Geo-Inf. Sci.* **30**(6), 31–35 (2014)
9. Kedzierski, M., Fryskowska, A.: Methods of laser scanning point clouds integration in precise 3D building modelling. *Measurement* **74**, 221–232 (2015)
10. Liu, F., Feng, Z.K., Yang, L.Y., et al.: Estimation of tree crown volume based on 3D laser point clouds data. *Trans. Chin. Soc. Agric. Mach.* **47**(3), 328–334 (2016)
11. Yu, J., Lv, P., Zheng, C.W.: A comparative research on methods of delaunay triangulation. *J. Image Graph.* **15**(8), 1158–1167 (2010)
12. Yuan, X.C., Wu, L.S., Chen, H.W.: Improvement and comparative analysis of Delaunay triangulation algorithm. *Comput. Appl. Softw.* **33**(9), 163–166 (2016)



Development of a New Type of Long Base Stress Sensor for Hull Structure

Wei Wang^(✉), Libo Qiao, and Yuliang Li

Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, College of Physical and Electronic Information, Tianjin Normal University, Tianjin, China
weiwang@tjnu.edu.cn

Abstract. In this paper, a new type of hull structure long base stress sensor is introduced. When the ship changes under external or internal forces, the fiber Bragg grating fixed between two bases is deformed. The deformation of the fiber grating and the wavelength change of the fiber grating are basically linear. By demodulating the end of the compensation algorithm, we can get the fiber wavelength and the ship structure stress. Therefore, the actual structural strength of the hull is accurately mastered, which provides a stable and reliable device for long-term detection and health monitoring of hull structure in the field of ocean ships.

Keywords: Fiber Bragg gratings · Long base stress sensor · Strain sensitivity coefficient

1 Introduction

In traffic and transportation vehicles, ships have the advantages of large load and low cost. However, ship safety problems are especially prominent in the harsh environment. Many unexpected situations are unpredictable [1]. The driver needs to know the actual structural strength of the hull accurately. Therefore, it is of great significance to detect the hull stress under normal and accidental loads.

Fiber grating sensors have good resistance to electromagnetic interference and electrical insulation. At the same time, it is safe and reliable to use in the terrible surrounding [2]. The fiber has the advantages of small transmission loss, small size, and plastic geometry. Remote monitoring can be realized, and it is convenient to form a sensor network. Because of its corrosion resistance and stable chemical performance, fiber optic sensors are suitable for marine vessels in such harsh circumstance [3].

The long base stress fiber grating sensor system is a network of stress sensors installed in the hull structure to realize the automatic monitoring of the ship structure stress change. It provides real and reliable data for shipboard decision makers on important structural parts of the ship [4–6]. As far as possible to ensure the safety of personnel and materials on board and enhance the ability of the ship to resist disasters. To improve safety performance and provide a scientific basis for secured ship safety.

The remainder of the paper is organized as follows. In Sect. 2 the basic fiber grating sensor theory will be introduced. Then optical fiber grating sensitivity principle will be

deduced base on stress and strain method. The third section of the article introduces the basic structure and operating principle of the sensor. Experimental results for stress-strain will be shown in Sect. 4. Conclusion and discussion is in Sect. 5.

2 Theory

2.1 Basic Fiber Grating Sensor Theory

The principle of fiber grating sensing is to convert the physical quantity of the measured structure into the amount of change of the Bragg wavelength reflected by the grating. The measured physical quantity can be measured by detecting the variation of the reflected Bragg wavelength. The principle is shown in Fig. 1 [7].

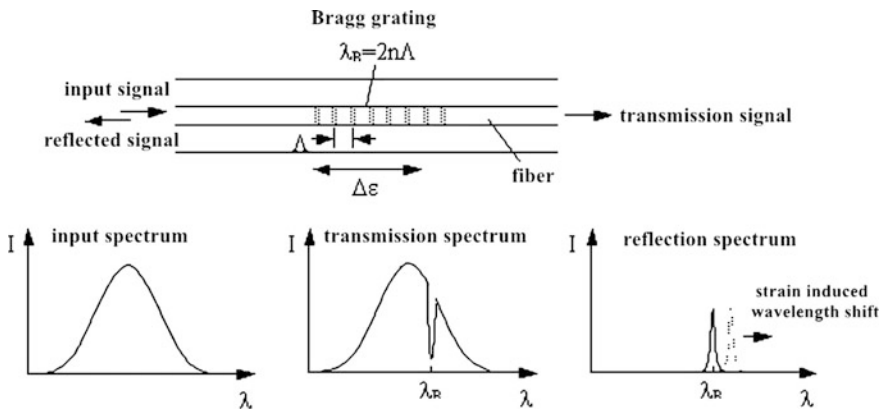


Fig. 1. Sensing principle of FBG

According to the coupled mode theory, the reflection wavelength λ_B of the fiber grating is mainly determined by the grating period Λ and the effective refractive index of the reverse coupling mode n_{eff} . So the reflection wavelength λ_B of fiber grating can be expressed by Formula 1.

$$\lambda_B = 2n_{eff}\Lambda \tag{1}$$

Any external influence (changes in temperature, stress, and pressure) that causes these two parameters to change will cause a shift in the grating Bragg wavelength (Formula 2).

$$\Delta\lambda_B = 2\Delta n_{eff}\Lambda + 2n_{eff}\Delta\Lambda \tag{2}$$

Therefore, the change of external conditions can be detected by measuring the variation of the center reflection wavelength and refractive index of FBG [8].

2.2 Optical Fiber Grating Sensitivity Theory

After the sensor deformation becomes stable, the reference sensor gives the amount of strain that occurs on the test piece ε_0 . At the same time, the wavelength variation λ_0 of the strain sensor to be calibrated is recorded. Then the strain sensitivity coefficient τ of the sensor to be calibrated can be calculated by Formula 3. Repeat two to three measurements and average. Finally, the average strain sensitivity coefficient $\bar{\tau}$ is obtained.

$$\tau = \frac{\Delta\lambda_0}{\varepsilon_0} \quad (3)$$

The sensor's strain variable is calculated using Eq. 4. L represents the reference length of the dependent variable and Δl is the deformation of the elastomer (displacement change).

$$\varepsilon_0 = \frac{\Delta l}{L} \quad (4)$$

The strain sensitivity coefficient of the fiber grating sensor is calibrated. Since the sensitivity coefficient is constant within a range, the sensor's strain amount ε can be obtained by the variation λ of the sensor wavelength.

3 Sensor Structure

Figure 2 is the structural diagram of the new type of hull structure long base stress sensor, which includes the fiber grating, connecting pipe, sensor base fixing plate, protection tube, the top seat of the sensor, sensor base movable plate, sealed lid, top wire and the sensor base welding plate. (Label 1 is the sealed lid. Label 2 is the sensor base movable plate. Label 3 is the top seat of the sensor. Label 4 is a protective tube. Label 5 is the connecting pipe. Label 6 is the sensor base fixing plate. Label 7 is the fiber grating.)



Fig. 2. Structural of the new hull structure long base stress sensor

In response to adverse conditions such as wetness and surge in the ocean, the sensor has been improved in terms of waterproof performance. The 316L stainless steel protective tube and the sealed lid are installed to improve the water tightness and anti-

destruction function of the system. Figure 3 is a diagram of the sensor housing, which functions to prevent surge damage.

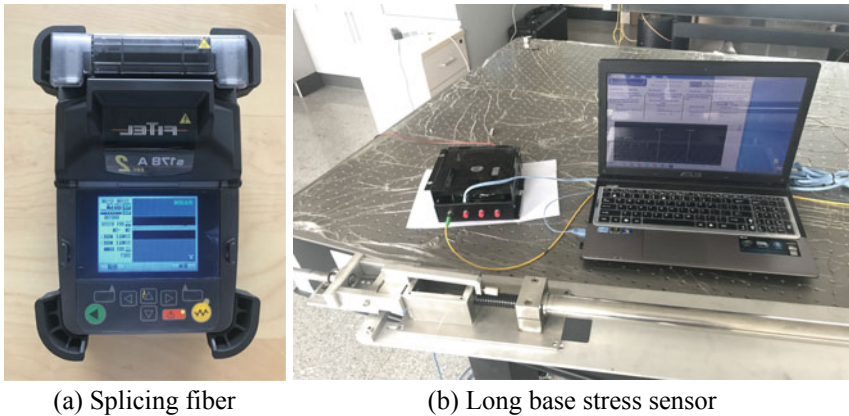


Fig. 3. Sensor housing

The key technology of the stress sensor is to use a single fiber grating. And paste the two ends of the prefabricated optical fiber grating on the base plate of the sensor base. Correct the pretension value of the fiber grating through the top wire. Ultimately the cable is connected and transmitted by the watertight flange. Finally, the stress of the ship structure is displayed on the demodulation instrument side in the form of wavelength variation of the fiber grating.

4 Sensor Test Experiment

The experiment uses a high-precision spiral micrometer to simulate the actual structural strength of the hull. By continuously moving the spiral micrometer, it represents that the hull is subject to varying degrees of strength changes. The correspondence between strain and length is used to calculate the sensitivity of the sensor, and the linearity curve of the sensor is plotted. Figure 4 is a related operation diagram of the test experiment.



(a) Splicing fiber

(b) Long base stress sensor

Fig. 4. Sensor experiment process

First, the fiber grating was beforehand pulled to 2 nm. Connect the sensor to the demodulator and observe the change in the center wavelength of the fiber through the computer. Then the high-precision spiral micrometer is continuously rotated forward.

The deformation of the sensor was observed through the demodulator, and the reading of the fiber wavelength is recorded at the same time.

Reverse moving the spiral micrometer and repeat the previous steps. Repeatedly recorded three times. The endpoint method is used to fit the calibration curve and the sensitivity of the sensor is calculated (Fig. 5). This experiment uses a high-precision micrometer, which changes by 0.5 mm for each revolution. Recording the data of the helical micrometer moving in 10 circles.

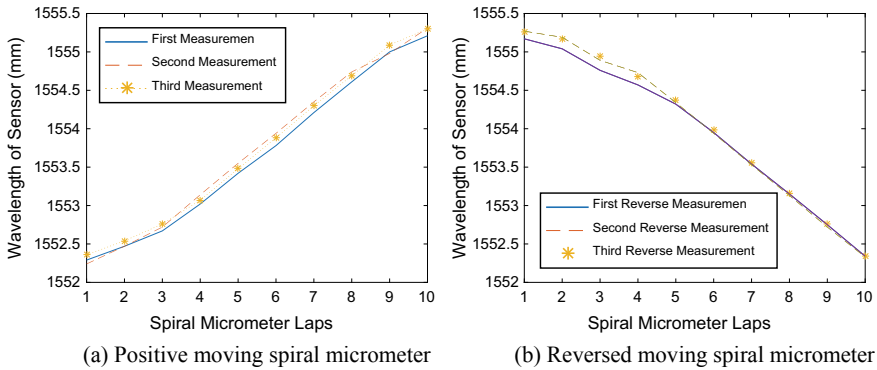


Fig. 5. The linearity curve of sensor

The maximum measurable deformation of the long base stress sensor is $2000 \mu\epsilon$. According to the image shown in Fig. 5, it can be seen that with the continuous positive movement of the spiral micrometer, the center wavelength of the optical fiber displayed by the demodulator continuously increases. Basically shows a linear change. The linearity of the sensor is 98.87%. It can be concluded that the sensor has good linearity. Using Eq. 3 to calculate, the sensor's sensitivity is $1.03 \text{ pm}/\mu\epsilon$.

Therefore, in the actual process, when the structural strength of the hull changes, the long base stress sensor can quickly capture this change and the center wavelength of the optical fiber will also change. Conversely, when the external strength of the hull returns to normal, the center wavelength of the fiber will return to its initial normal value.

5 Conclusions

In the process of designing the long base stress sensor system of ship structure, the two ends of the prefabricated fiber grating are affixed to the movable plate of the sensor base. The pretension value of the fiber grating is corrected by the top wire. Watertight flanges are connected to the transmission cable. The stress of the final hull structure is displayed at the end of the demodulator in the form of wavelength change of the fiber grating.

The design also has the function of adjusting the initial zero position. The adjustment of the sensor zero position can be realized by using the installation tool, and it has the function of collimation. The system is convenient for installing, calibration, and cable laying. The stainless steel protective pipe and sealed lids are set up in the light of the bad working conditions such as sea humidity and surge. Therefore, the water tightness and anti-damage function of the system are improved. Improve the applicability and durability of fiber grating sensors for marine transportation applications. Therefore, it provides a stable and reliable device for long-term detection and health monitoring of hull structures in ocean-going ships.

Acknowledge. This paper is supported by the Natural Science Foundation of China (61271411), Natural Youth Science Foundation of China (61501326, 61401310). It also supported by Tianjin Research Program of Application Foundation and Advanced Technology (15JCZDJC31500), and Tianjin Science Foundation (16JCYBJC16500).

References

1. Wu, Z., Wang, F., Zhao, J.: A statistical study of marine accidents and a synthetical evaluation of ship safety. *J. Dalian Mar. Coll.* (1991)
2. Moody, V.: *Fiber theory and formation*. Elsevier Inc. (2004)
3. Jiang, D.S., He, W.: Review of applications for fiber Bragg grating sensors. *J. Optoelectron. Laser* **4**(13), 420–430 (2002)
4. Chen, X.Y.: Distributed measurement system and its model for ship structure. *Ship Eng.* **5** (26), 62–66 (2004)
5. Chen, X.Y., Wang, X.Q.: Application of fiber Bragg grating sensors on smart ship structure. *Ship Eng.* **2**(28), 29–32 (2006)
6. Wang, W.: *Study of key technology on ship hull structure health monitoring with fiber Bragg grating*, Tianjin University, Tianjin (2007)
7. Liu, Y., Zhang, W., Jin, Z., Liu, J., Li, M., Li, X., Geng, B., Dong, B.: Research on the gear operating state detection based on the fiber Bragg grating sensing technology. In: *IOP Conference Series: Materials Science and Engineering* (2017)
8. Lee, B.: Review of the present status of fiber sensors. *Opt. Fiber Technol.* 3–4 (2003)



Research of Single-Phase Photovoltaic Grid-Connected System Based on MATLAB

Hao Yang^(✉)

Anhui Technical College of Mechanical and Electrical Engineering,
Wuhu 241002, Anhui, China
etbeat@163.com

Abstract. Based on the analysis of structure and common control scheme of photovoltaic grid-connected system, a solar photovoltaic (PV) cell simulation model is built according to the internal structure and output characteristics of PV cell. A control method of photovoltaic MPPT which can track photovoltaic points and the maximum power of photovoltaic cells is expounded. This method can be used to improve the efficiency of photovoltaic cells to the greatest extent. Finally, the principle and control strategy of the photovoltaic grid-connected system are analyzed in detail. According to the instantaneous value control method of grid-connected inverter, some requirements of grid connection with low THD and high power factor are achieved.

Keywords: Photovoltaic grid-connected system · PV cell · MPPT control · Instantaneous value control

CLC: TM72

Document code: A

1 Introduction

Energy is indispensability for the development of the world economy. However, because of the depletion of traditional energy sources such as coal, oil, and natural gas and increasing of the world population, the energy problem has appeared more or less in many countries all over the world, and the energy crisis has become a difficult problem to be solved all over the world. The emergence of new renewable energy has greatly alleviated the energy crisis, especially the utilization of solar energy. As a new type of renewable energy, solar energy has attracted more and more attention in recent years due to its advantages of green environment and inexhaustible energy [1].

One of the most common utilization of solar energy is grid-connected photovoltaic power generation. First, a solar array is used to collect solar energy, and converts the heat energy of solar into DC power and then outputs to the inverter. The DC energy is converted into sinusoidal alternating current energy which will be incorporated into the AC network, thus an active inverter system is formed [2].

2 Composition of Photovoltaic Grid System

MPPT two-stage mode is the common structure of photovoltaic grid system, which is generally composed of PV array, DC–DC chopper circuit, DC–AC inverter circuit, active filter circuit, and so on, as shown in Fig. 1. The DC–DC converter is used to track and control the maximum power point of photovoltaic. The DC–AC inverter is used to turn the DC inverse of the solar cell into alternating current and sends to AC network, and controls its phase be same with the grid, so that the power factor is close to 1 [3]. The active power filter is used to filter the high harmonics contained in alternating current power before the grid is connected, so that it is closer to an ideal AC sine signal.

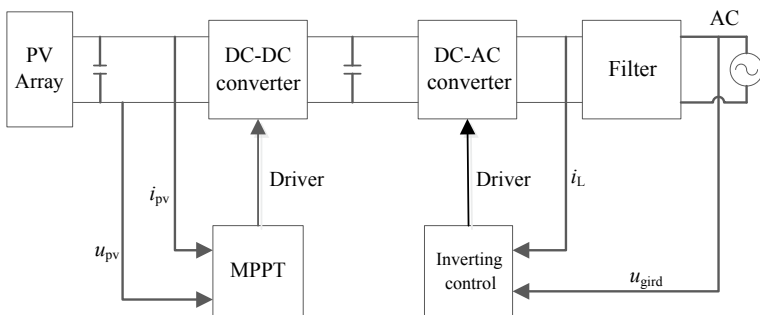


Fig. 1. Two-stage grid-connected structure mode

3 PV Cell Modeling and MPPT Control Method

3.1 PV Cell

There are many factors in nature which can affect the efficiency of PV cell, such as light intensity, temperature and so on. The output of the PV cell will appear nonlinear due to PN junction parameters and external environmental factors. The relationship between output voltage and output current is shown in Eq. (1) [4, 5].

$$I = I_{ph} - I_o \{ \exp[q(U + IR_s)/AKT] - 1 \} - (U + IR_s)/R_{sh} \quad (1)$$

Thereinto, U is the output voltage of PV Cell; I is the output current of PV cell; I_{ph} is the photocurrent current of PV cell; I_o is the reverse saturation current of diode. R_s is the series equivalent resistance; R_{sh} is the parallel equivalent resistance; K is Karl Seidman constant; T is the absolute temperature of PV cell; A is a dimensionless curve fitting constant with value range of $1 \leq A \leq 2$. According to the engineering mathematical model, the PV cell simulation model is built in Simulink, as shown in Fig. 2.

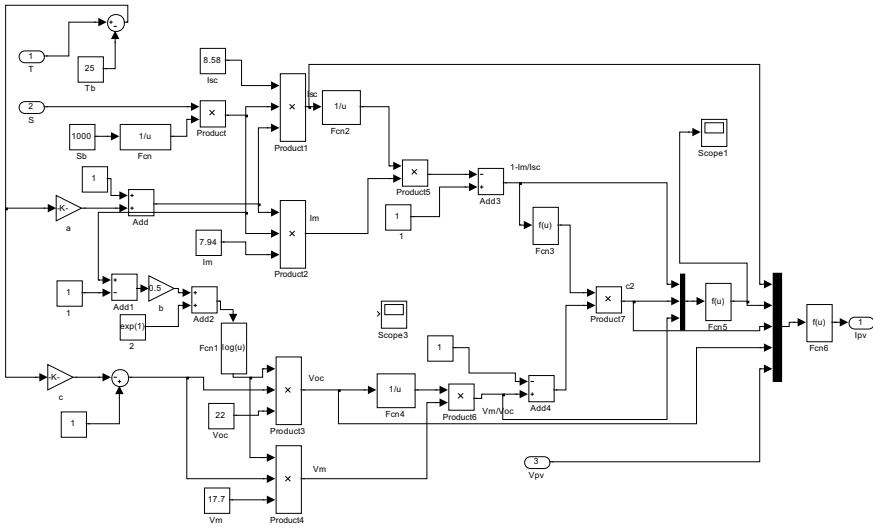


Fig. 2. PV cell simulation model

When the light intensity is 1000 W/m^2 and the ambient temperature is $25 \text{ }^\circ\text{C}$, the I-V and P-V characteristic curves of PV cells are shown in Figs. 3 and 4. With the increasing of PV cell output current, the output voltage gradually decreases, and the PV cell output power will increase gradually with the increasing of PV cell output voltage. After the maximum power point, the output power decreases sharply as the output voltage continues to increase.

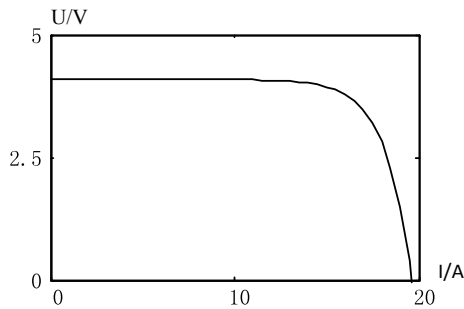


Fig. 3. I-V output characteristic curve

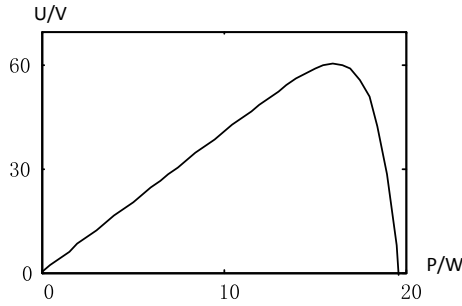


Fig. 4. P-V output characteristic curve

3.2 MPPT Control Method

MPPT control can effectively solve instability of system working points caused by the nonlinear output voltage of PV cell, so that the maximum power can be output under any conditions, therefore, MPPT control is widely used in the field of PV grid system. There are some methods of MPPT control, such as CVT, P&O or IC³ and so on. In this paper, the disturbance observation method is used to control the output voltage. Because the Boost circuit has the advantages of the simple driving circuit and always working in the continuous current state, this circuit is usually used to achieve maximum power point tracking. The MPPT simulation model built in the Simulink toolbox is shown as Fig. 5, in which $L = 10^{-3}$ H, $R_1 = 30 \Omega$, $C = 3 * 10^{-3}$ F.

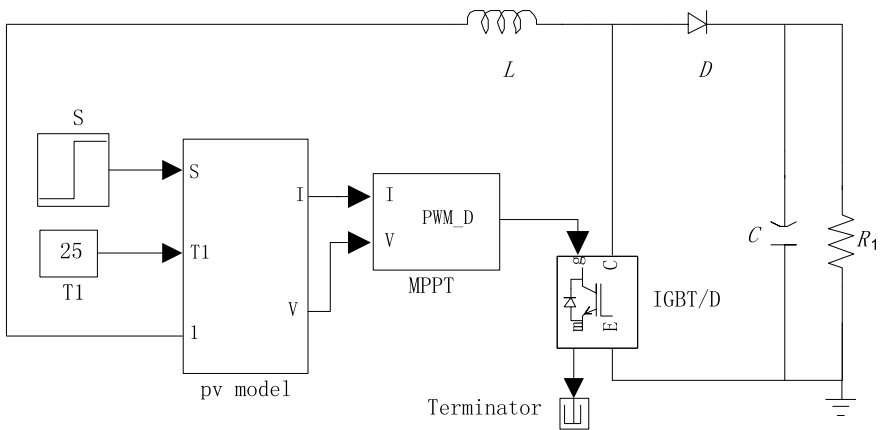


Fig. 5. MPPT simulation model

When the light intensity changes from 600 to 1000 W/m², the output voltage waveform of the simulation model is shown as Fig. 6.

It can be seen from the waveform in Fig. 6 that the P&O control has a good dynamic response and can track the maximum power point of PV cell well. The control

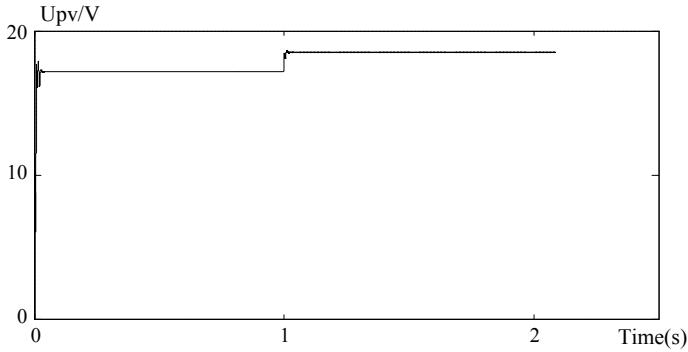


Fig. 6. MPPT voltage output waveform

scheme is simple and easy to be realized, and the control precision is more accurate than CVT. However, there are some disadvantages of this control scheme that the output voltage will be slightly oscillating near the maximum power point, and the control precision is greatly influenced by the disturbance step.

4 Control Strategy of Two-Stage Photovoltaic Inverter

The two-stage type PV inverter circuit is a voltage type full control bridge inverter circuit as shown in Fig. 7, where U_{dc} is the steady voltage of PV array after Boosting and control, U_o is the output voltage of inverter, U_{grid} is the grid voltage, R_2 is the equivalent resistance of the line (generally small, assuming that is 0.01Ω), L is the filter inductance which is $6 * 10^{-3} \text{ H}$ and is used to filter out the higher harmonics in output voltage, at the same time, to balance the voltage difference between the output voltage and the grid voltage [6].

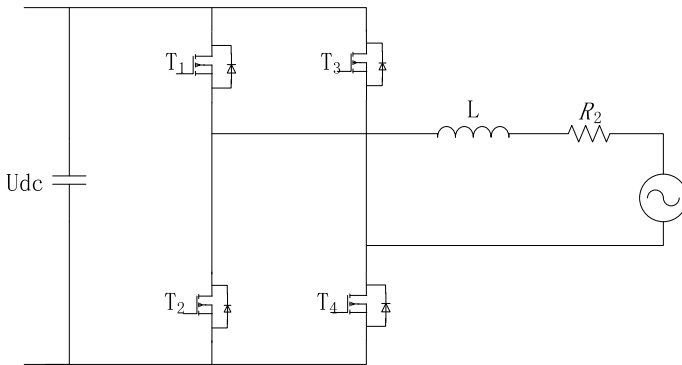


Fig. 7. Voltage type full control bridge inverter circuit

There are four switch tubes of the full bridge inverter circuit that are divided into two groups T1, T4 and T2, T3 which are connected alternately, using SPWM driving mode. After inverting, the AC voltage signal is filtered out of the high harmonic through the filter inductor; after adjusting the voltage phase, the AC voltage signal becomes AC sinusoidal signal with less harmonic content and the same frequency as the voltage of the power grid, which is incorporated into AC network.

There are many methods to control the inverter circuit. In this paper, the instantaneous value control method which has a better dynamic response is used. This control method has been widely used in the field of active inverter because of its advantages such as fast dynamic response speed, simple control and less harmonic content in the output voltage waveform. The control logic schematic diagram of SPWM instantaneous current is shown in Fig. 8 [7].

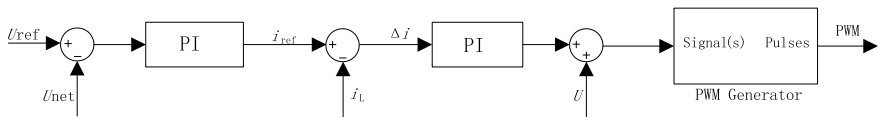


Fig. 8. Control logic of SPWM instantaneous current

The difference between the reference voltage and the output voltage of the sampling inverter is defined as the instantaneous error. The instantaneous error signal is adjusted by the PI signal regulator as the current reference signal i_{ref} which is compared with the instantaneous value i_L of the sampling inductor current to produce the error signal Δi . After the PI adjustment and the voltage feedforward compensation, the signal is compared with the triangle carrier, that is, the SPWM drive signal is produced which is used to control the power switch tube in the inverter circuit. According to the above control methods, a grid-connected control model is built in MATLAB, as shown in Fig. 9, where $P = 0.01$; $I = 100$; $K = 1/400$.

Figure 10 shows the simulation results of AC grid-connected power generation system under an illumination intensity of 1000 W/m^2 and ambient temperature of 25°C .

Where i_L is the grid-connected current, U_{net} is the power grid voltage. After 0.06 s, the inverter output has the same frequency, same phase and with THD (the total harmonic base) of $3.6\% \leq 5\%$, which meet grid connection requirements.

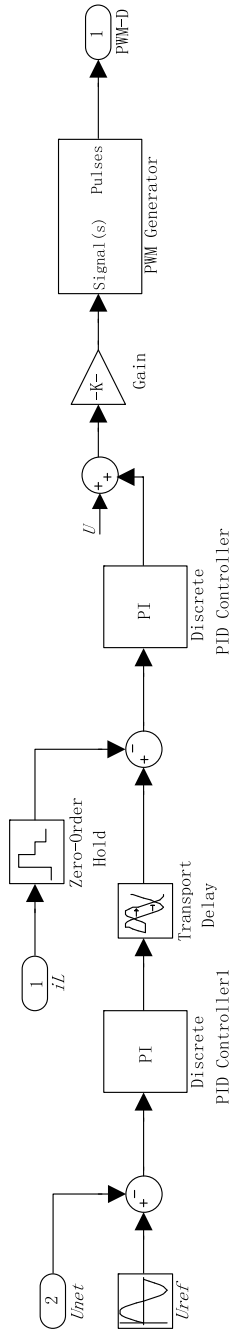


Fig. 9. Instantaneous current control model

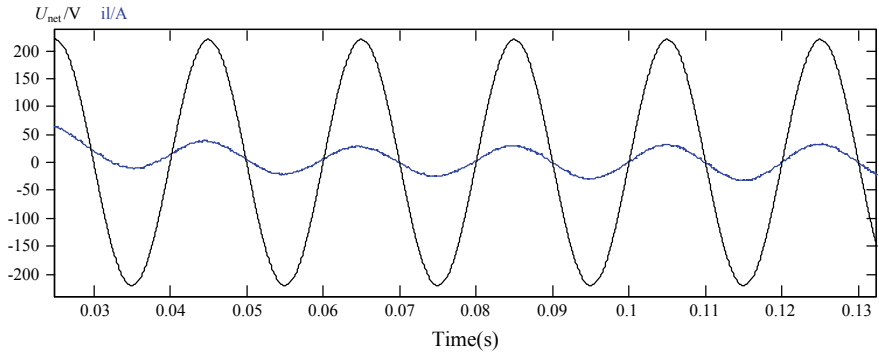


Fig. 10. Output waveform of grid-connected current and voltage

5 Conclusion

As shown in Fig. 10, the voltage-type full bridge inverter which is combined with P&O control method and current instantaneous value control method can make the output power of PV cell track the maximum power point well, and better achieve the demand of output voltage and grid. However, this control scheme still has some shortcomings, such as contradiction between disturbance step size and speed and control accuracy, which need further optimization.

Fund Project. The Key Research Project of Natural Science for Anhui Provincial Universities (KJ2017A747).

References

1. Yong, L.: Solar photovoltaic grid-connected power generation system. Shanghai Electr. Power **1**, 49–53 (2005)
2. Peiqing, Z.: Research on single-phase photovoltaic grid-connected system, pp. 34–35. Southwest Jiaotong University, Chengdu (2010)
3. Donghui, L., Hexiong, W., Xiaodan, Z., et al.: Study on several key issues of photovoltaic grid-connected power generation system. Power Syst. Prot. Control **38**(21), 208–214 (2010)
4. Changjiang, W.: Universal mathematical model of photovoltaic cell based on matlab. Electr. Power Sci. Eng. **25**(4), 11–14 (2009)
5. Chengyu, H., Quanzhu, Z., Yonghong, D., et al.: Research on the matlab simulation of single-phase photovoltaic two-stage grid-connected system. Power Technol. **35**(5), 553–555 (2011)

6. Yan, S., Yin, M., Li, Q., et al.: Research on related technologies of solar photovoltaic grid-connected systems. *Appl. Electron. Compon.* **11**(1), 77–80 (2009)
7. Zhang, H., Shan, L., Ren, J.: et al. Study on photovoltaic grid-connected inverter control system. In: *International Conference on Power Electronics*, pp. 210–212 (2010)

Yang Hao Male, lecturer, Master, Research orientation: power electronics and power drives.



Analysis of Output Characteristics of Photovoltaic Arrays Under Shaded Conditions

Hao Yang^(✉)

Anhui Technical College of Mechanical and Electrical Engineering, Wuhu,
Anhui 241002, China
etbeat@163.com

Abstract. The output power–voltage curve of photovoltaic array under uniform illumination is a single-peak value, the voltage–current U-I curve is a single-step type. Once the photovoltaic array is obscured by clouds, mud, trees, etc., the photovoltaic array appears like partial shadows, and the P-U curve output of the photovoltaic array will appear like multiple peak values, and the U-I curve appears multiple steps, which will reduce the output power of the photovoltaic array. By using MATLAB/Simulink to build the model of the photovoltaic array under the shadow condition and carry out several simulations, we can get some characteristics of the P-U curve and the U-I curve, and summarize the rules. It provides a theoretical basis for the maximum power point tracking (MPPT) under the shadow condition.

Keywords: Photovoltaic array · Partial shading · P-U characteristic curve

1 Introduction

Because solar energy is renewable, clean, and inexhaustible, it is used all over the world and has gradually become the main part of the world energy composition.

Photovoltaic cells convert light energy into electric energy through light for use. The output characteristics of photovoltaic cells are easily affected by the environment. Under uniform illumination, the P-U curve of photovoltaic cells has a single-peak characteristic, and its maximum power is easy to trace. However, when photovoltaic cells are obscured by trees, clouds, buildings, mud, etc., shadows are formed on photovoltaic cells [1]. In this case, the output characteristics of photovoltaic cells will change, so the study of the output characteristics of photovoltaic panels under shadow conditions is of great significance to the study of the maximum power point tracking (MPPT) technology of photovoltaic cells and the improvement of the power-generation efficiency of photovoltaic cells.

In this paper, the output characteristics of photovoltaic arrays under various shadow conditions are simulated by using MATLAB/Simulink software. The law and characteristics of the output characteristics are analyzed through the simulation results, and the reasons for this rule are analyzed.

2 Photovoltaic Cell Model

Figure 1 is the equivalent model of photovoltaic cells with a single diode, according to document [2], the output characteristic equation of its mathematical model is Formula (1).

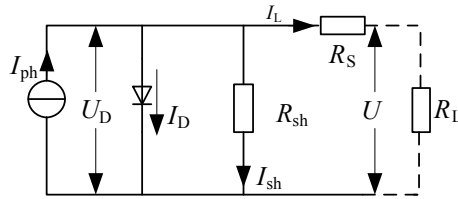


Fig. 1. The equivalent circuit diagram of photovoltaic cells

$$I_L = I_{ph} - I_0 \left[\exp\left(\frac{q(U + I_L R_s)}{AkT}\right) - 1 \right] - \frac{(U + I_L R_s)}{R_{sh}} \tag{1}$$

$$U_D = U + I_L \cdot R_s \tag{2}$$

In the Formula, U_D is the end voltage of the equivalent diode;

I_L is the output current of photovoltaic cells;

I_{ph} is the photoinduced current of photovoltaic cells;

I_0 is the reverse saturable electric current of diode;

q is electron charge (1.6×10^{-19} C);

R_s is a low resistance value for series resistance, less than 1 Ω ;

A is a kind of curve constant, and the main reaction is internal PN section, the range of value is 1 ~ 5;

K is the Boltzmann constant -1.38×10^{-23} J/K;

T is the absolute temperature of the environment where the photovoltaic cells are located ($t + 273$ K);

R_{sh} The Bypass Resistance of Photovoltaic cells in parallel;

In the Formula (1), I_{ph} , I_0 , and R_{sh} are not only related to the temperature and intensity of the battery, but also difficult to measure, Therefore, this model is not used in engineering [3].

Since the photovoltaic cell manufacturer only gives the battery parameters under the standard test conditions (illumination intensity $S_{ref} = 1000$ W/m², temperature $T_{ref} = 25$ °C), maximum power P_m , maximum power point operating voltage U_m , the maximum power point operating current I_m , the open circuit voltage U_{oc} , and short electric circuit current I_{sc} . Therefore, it is necessary to modify the engineering model under any conditions.

Since the value of R_{sh} is very large and the value of R_s is very small, it can be neglected in practical calculation, then there is $I_{ph} \approx I_{sc}$, $U_D \approx U$, as a result, the Formula (1) can be simplified as follows:

$$I_L = I_{sc} \left[1 - C_1 \left(e^{\frac{U}{c_2 U_{oc}}} - 1 \right) \right] \tag{3}$$

In the formula, $C_1 = I_0/I_{sc}$, $C_2 = \left(\ln \left(\frac{1}{C_1} + 1 \right) \right)^{-1}$

At the maximum power, put $I_L = I_m$ and $U = U_m$ into Formula (3) to get

$$I_m = I_{sc} \left[1 - C_1 \left(e^{\frac{U_m}{c_2 U_{oc}}} - 1 \right) \right] \tag{4}$$

Owing to $\frac{U_m}{c_2 U_{oc}} \gg 1$, neglecting the -1 Item of Formula (4), C_1 can be got

$$C_1 = \left(1 - \frac{I_m}{I_{sc}} \right) e^{-\frac{U_m}{c_2 U_{oc}}} \tag{5}$$

When the circuit is open, $I_L = 0$, $U = U_{oc}$, put Eq. (5) into Eq. (4), it is

$$I_L = I_{sc} \left[1 - \left(1 - \frac{I_m}{I_{sc}} \right) e^{-\frac{U_m}{c_2 U_{oc}}} \left(e^{\frac{1}{c_2}} - 1 \right) \right] \tag{6}$$

Owing to $e^{\frac{1}{c_2}} \gg 1$, neglecting the -1 Item of Formula (6) to get:

$$C_2 = \frac{\left(\frac{U_m}{U_{oc}} - 1 \right)}{\ln \left(1 - \frac{I_m}{I_{sc}} \right)} \tag{7}$$

So we can calculate C_1 and C_2 according to the parameter U_m , I_m , U_{oc} , I_{sc} provided by the manufacturer under the standard test condition, and then get the characteristic curve of photovoltaic cell tested under the standard condition by the Formula (3).

However, the above equations are only derived under the standard test conditions, because the environment is changing, so we should deduce the characteristic equations under any temperature and light.

Suppose that ΔS is the difference of light intensity, and ΔT is the difference of temperature, then

$$\Delta S = S - S_{ref} \tag{8}$$

$$\Delta T = T - T_{ref} \tag{9}$$

From the above points, it can be deduced that its general modified expression is

$$I'_{sc} = I_{sc}(1 + a\Delta T) \tag{10}$$

$$I'_m = I_{sc} \frac{S}{S_{STC}} (1 + a\Delta T) \tag{11}$$

$$U'_m = U_m(1 - c\Delta T) \ln(1 + b\Delta S) \tag{12}$$

$$U'_{oc} = U_{oc}(1 - c\Delta T) \ln(1 + b\Delta S) \tag{13}$$

A, B, and C are the environmental compensation coefficients, according to the reference [4], we can get $a = 0.0025 \text{ (C)}^{-1}$, $b = 0.5 \text{ (W/m}^2\text{)}^{-1}$, $c = 0.00288 \text{ (C)}^{-1}$, the three parameters of a, b, c are put into I'_{sc} , I'_m , U'_m , U'_{oc} , and then into Formulas (5) and (7) to find out “ C_1 ” and “ C_2 ”, and then put into Formula (3), finally the output characteristic curves under arbitrary conditions can be obtained.

3 Modeling and Simulation Analysis of Photovoltaic Array

Because the operating voltage and operating current range of the single photovoltaic cell is small, it is generally not used alone as a power source, so the single photovoltaic cell needs to be connected in series and parallel in a certain way and then encapsulated, and PV Modules can be used as power source.

A photovoltaic module can generate a voltage of about 16 V, when the application requires a higher voltage and current, a photovoltaic module cannot meet the need. So it is necessary to connect PV modules in series, parallel and then install them on the bracket, thus forming the PV Array, Fig. 2 is the relationship of the single photovoltaic cell, the photovoltaic module, and the photovoltaic array.

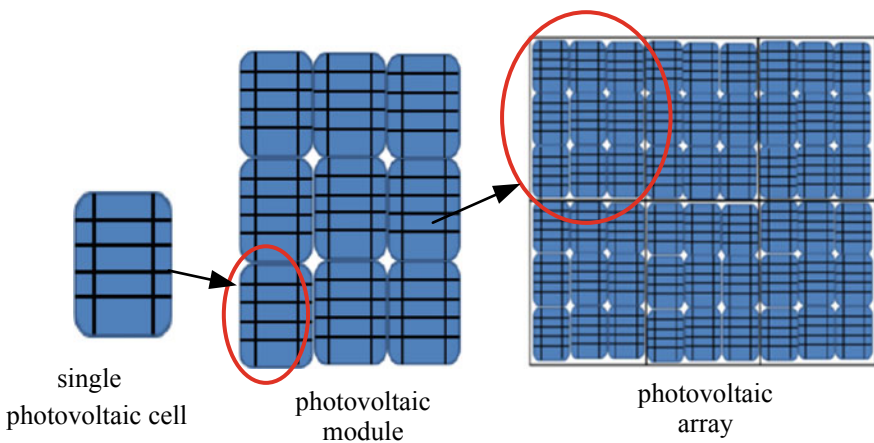


Fig. 2. The relationship of the single photovoltaic cell, the photovoltaic module, and the photovoltaic array

Because the cell characteristics of the single photovoltaic cell, the photovoltaic module, and the photovoltaic array are similar, we can enlarge and reduce only the electricity output of the photovoltaic module for the electricity output of the photovoltaic array, then the output voltage U_{pv} and the current I_{pv} of the PV array are as follows:

$$\begin{cases} U_{pv} = m \cdot U \\ I_{pv} = n \cdot I_L \end{cases} \quad (14)$$

Among them, m and n are series and parallel number of PV modules, respectively, and U and I_L are the voltage and current output by PV modules, respectively.

According to the engineering model of photovoltaic module, two simulation models of photovoltaic module are built in MATLAB/Simulink, in order to study the influence of partial shadow on photovoltaic module, the simulation model of PV module produced by Yingli company is selected for simulation, and each module parameter is $U_m = 30.23$ V, $I_m = 7.94$ A, $U_{oc} = 37.47$ V, $I_{sc} = 8.53$ A, $P_m = 240.8$ W.

3.1 Factors Causing Multi-peak Phenomenon of Photovoltaic Array

In order to avoid the hot spot effect [5], it needs to parallel bypass diode reversely at both ends of PV modules to prevent PV module from burning when shadows appear. The following are simulated in the following cases:

1. For the condition of photovoltaic module in series with uniform illumination, the simulation model is shown in Fig. 3. Assuming the illumination intensity of photovoltaic modules PV1 and PV2 is 1000 W/m^2 , and the temperature T is $25 \text{ }^\circ\text{C}$, the P-U curve of the output of photovoltaic array is shown in Fig. 4.

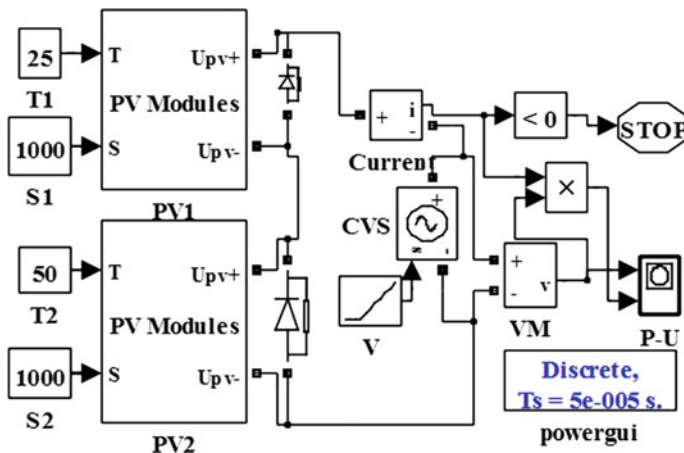


Fig. 3. Simulation model of photovoltaic module under standard test

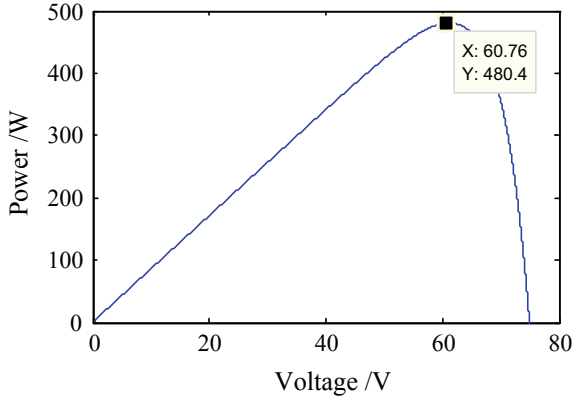


Fig. 4. P-U curves under standard test

2. When the photovoltaic module is connected in series, it is assumed that the PV1 in Fig. 3 is obscured, The intensity of illumination is 600 W/m^2 , the temperature is unchanged, the intensity, and temperature of illumination of the PV2 are constant, that is, the simulation model only needs to change the S1 in Fig. 3 to 600 W/m^2 , at this time the output P-U curve is shown in Fig. 5.

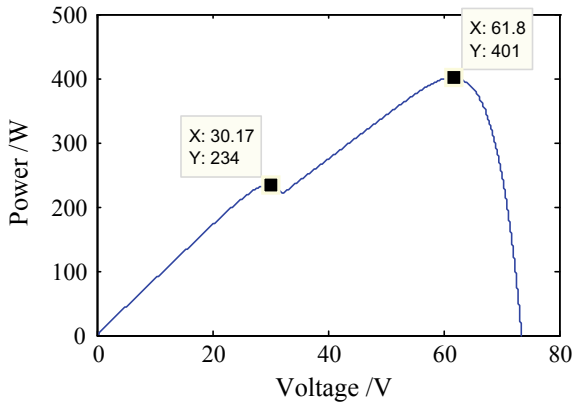


Fig. 5. P-U curve of photovoltaic module in series shadow

3. For the temperature change of the photovoltaic module in series, assuming that the temperature of the PV1 module in Fig. 3 is $45 \text{ }^\circ\text{C}$, the illumination intensity and temperature of PV2 are invariant, that is, the simulation model diagram only needs to change T1 in Fig. 3 to $45 \text{ }^\circ\text{C}$, the P-U curve at this time is shown in Fig. 6.

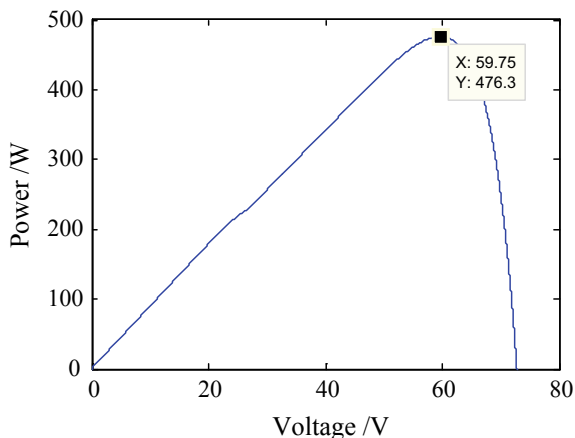


Fig. 6. P-U curve of temperature change in photovoltaic module in series

- For the condition of the appearance of shadow when photovoltaic modules are in parallel, the simulation model is shown in Fig. 7. Assuming that the module PV2 is obscured, the illumination intensity becomes 600 W/m^2 , the temperature is $25 \text{ }^\circ\text{C}$, and the illumination intensity is 1000 W/m^2 , the P-U curve of photovoltaic array output is shown in Fig. 8.

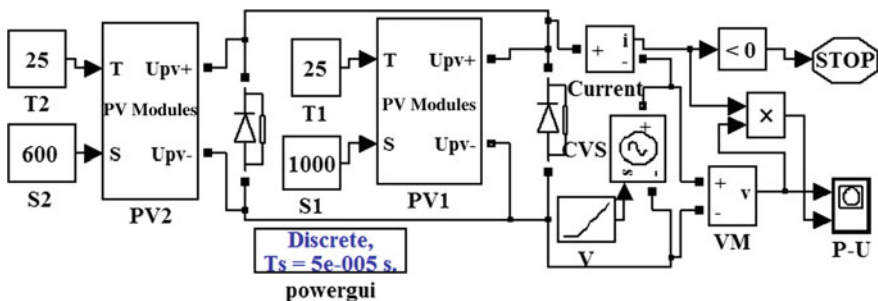


Fig. 7. Model of photovoltaic module in parallel shadow

- For the condition of temperature variation of photovoltaic modules in parallel, suppose that the temperature of PV2 in Fig. 7 becomes $50 \text{ }^\circ\text{C}$, the illumination

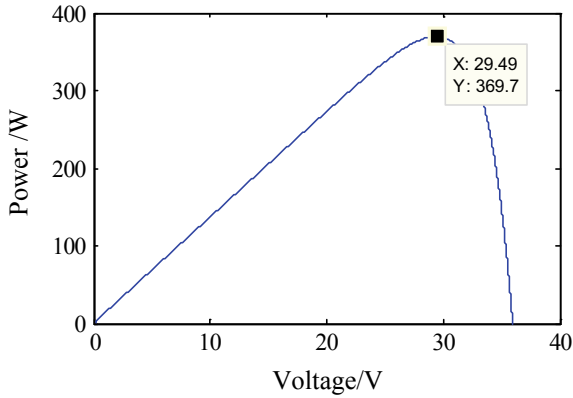


Fig. 8. P-U curve of photovoltaic module in parallel shadow

intensity is 1000 W/m^2 , the illumination intensity and temperature of the PV1 light are constant, that is, if the T2 in Fig. 7 is changed to $50 \text{ }^\circ\text{C}$, and S2 changed to 1000 W/m^2 , the P-U curve of the output of the photovoltaic array is shown in Fig. 9.

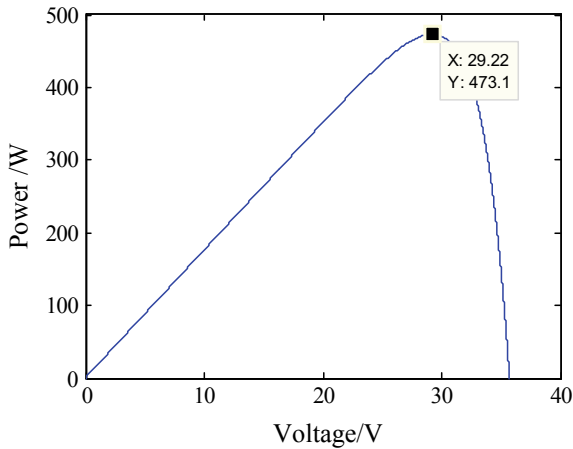


Fig. 9. P-U curve of parallel temperature variation in PV module

It can be obtained from Fig. 4 that the output P-U curve of the PV array has a single-peak value under uniform illumination, and the maximum power is 480 W, which is consistent with the parameters provided by the manufacturer, thus verifying the correctness of the photovoltaic cell engineering model.

Compared with Figs. 5 and 8, it can be seen that when the shadow appears in the series module, the P-U curve of the photovoltaic array output appears multiple peaks, but in parallel, when the shadow appears, there are no multiple peaks.

Compared with Figs. 6 and 9, it can be seen that no matter whether the photovoltaic module are in series or parallel, and the difference in temperature does not cause the P-U curve of photovoltaic array output to peak, so when the model is analyzed, the effects of temperature differences on photovoltaic arrays can be ignored.

3.2 Output Characteristic Law of Photovoltaic Arrays

From the previous section, it is known that the multiple peaks appear only on the series modules, so this section uses four photovoltaic modules in series for making a simulation to observe the characteristics of the photovoltaic array output, and the photovoltaic cell parameters are unchanged. The simulation model is shown in Fig. 10.

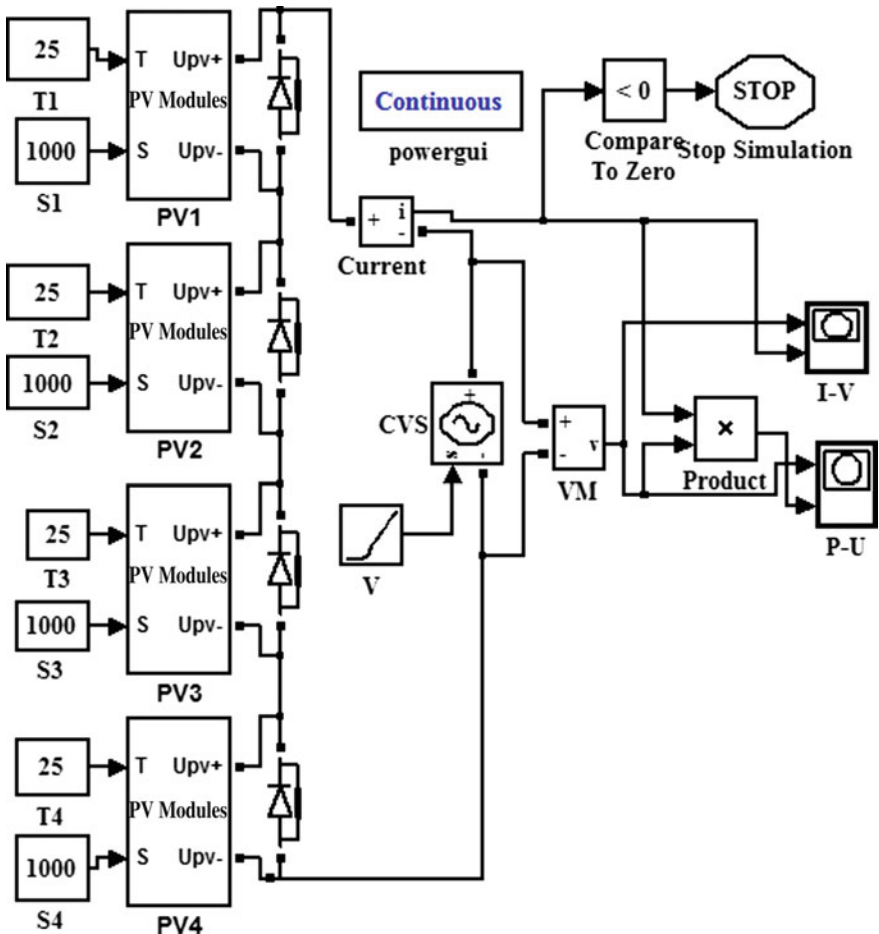


Fig. 10. Simulation model of photovoltaic array

In order to observe the output characteristics of the photovoltaic array under the condition of partial shadow, the intensity of illumination is simulated according to the following sets of data.

1. S1 is 1000 W/m^2 , S2 is 1000 W/m^2 , S3 is 1000 W/m^2 , S4 is 600 W/m^2 , The output P-U curve and I-U curve are shown in Fig. 11 and Fig. 12, respectively.

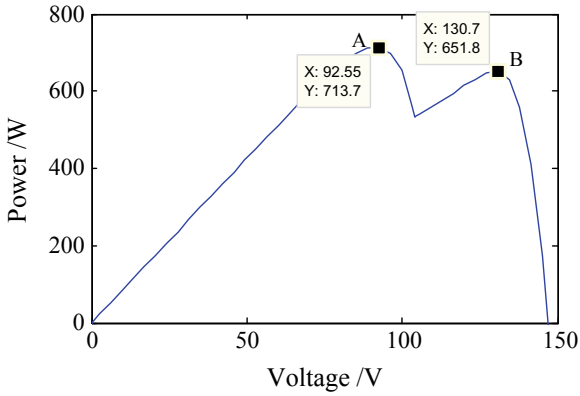


Fig. 11. The P-U curve after being obscured

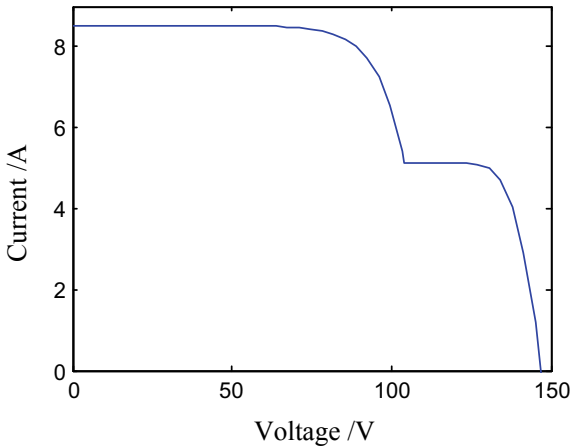


Fig. 12. The I-U curve after being obscured

From Fig. 11, we can see that the coordinates of the two peak points are A (92.55713.7), B (130.7651.8), the voltage difference of two peak points, $\Delta U_{AB} \approx 130.7-92.55 = 38.15 \approx 0.9 \times U_{oc}$ and from Fig. 12, the I-U curve of photovoltaic array output has two steps.

- S1 is 1000 W/m², S2 is 1000 W/m², S3 is 800 W/m², S4 is 500 W/m², the output P-U curve and I-U curve are shown in Fig. 13 and Fig. 14, respectively.

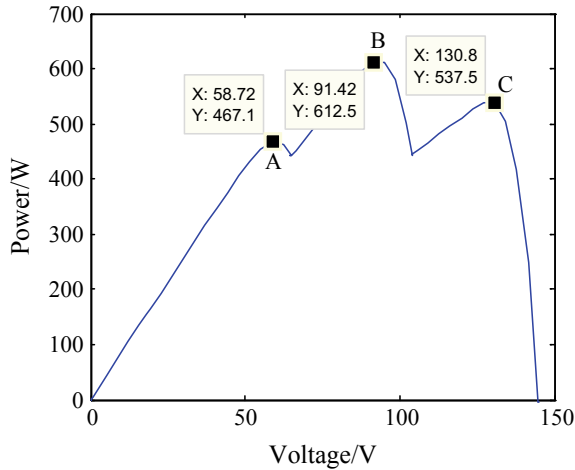


Fig. 13. The P-U curve after being obscured

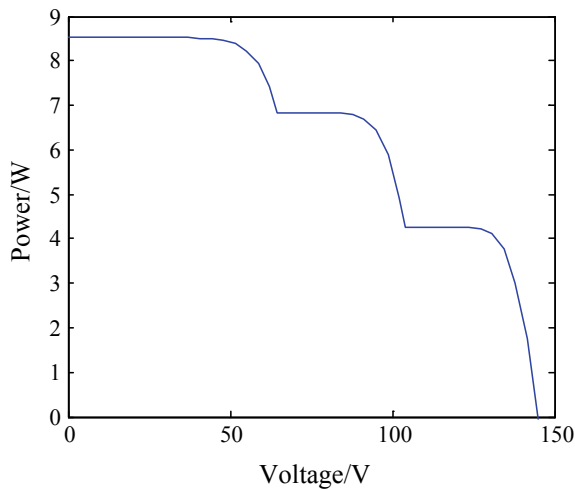


Fig. 14. The I-U curve after being obscured

It can be seen from Fig. 13 that the coordinates of the three peak points are A (58.72467,1), B (91.42612,5), C (130.8537,5), The voltages between peak points are $\Delta U_{AB} \approx 91.42-58.72 = 33.08 \text{ V} \approx 0.9 \times U_{oc}$, $\Delta U_{BC} \approx 130.8-91.2 = 39.6 \text{ V} \approx 0.9 U_{oc}$, from Fig. 14, it can be seen that the I-U curve has three steps.

3. S1 is 1000 W/m^2 , S2 is 800 W/m^2 , S3 is 600 W/m^2 , S4 is 400 W/m^2 , the output P-U curve and I-U curve are shown in Fig. 15 and Fig. 16, respectively.

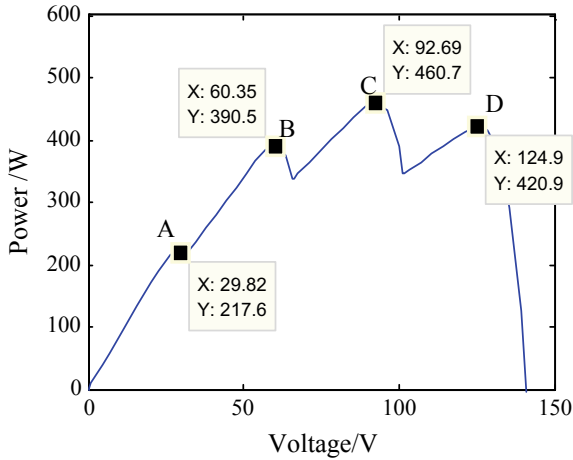


Fig. 15. The P-U curve after being obscured

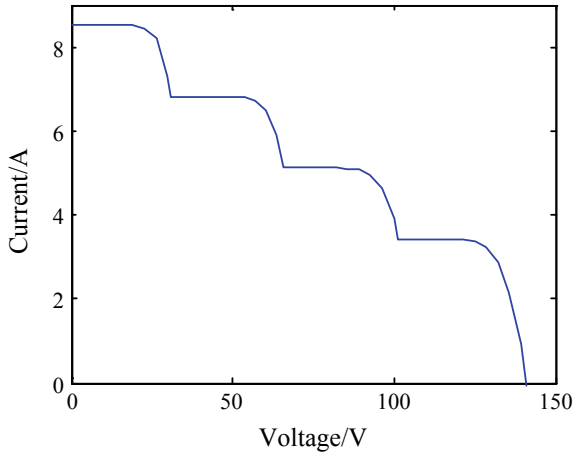


Fig. 16. The I-U curve after being obscured

As shown in Fig. 15, the coordinates of the four peak points are A(29.82, 217.6), B(60.35, 390.5), C(92.69, 460.7), D(124.9, 420.9), respectively, the voltages between peak points are $\Delta U_{AB} \approx 60.35 - 29.82 = 30.53 \text{ V} \approx 0.9 \times U_{oc}$, $\Delta U_{BC} \approx 92.69 - 60.35 = 32.34 \text{ V} \approx 0.9 \times U_{oc}$, $\Delta U_{CD} = 124.9 - 92.69 = 32.21 \text{ V} \approx 0.9 \times U_{oc}$, respectively, as shown in Fig. 16, the I-U curve of PV array output has four steps.

4. S1 is 1000 W/m^2 , S2 is 700 W/m^2 , S3 is 700 W/m^2 , S4 is 700 W/m^2 , the output P-U curves and I-U curves are shown in Fig. 17 and Fig. 18, respectively.

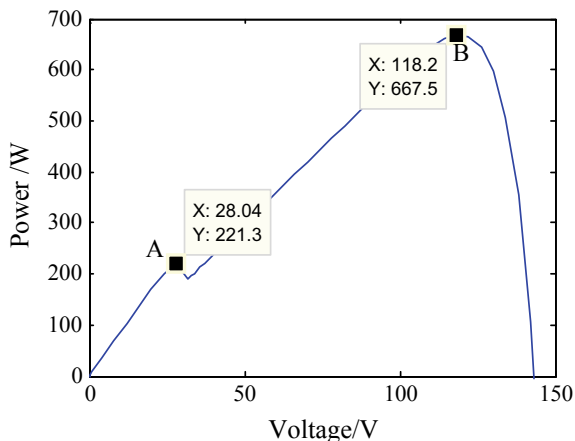


Fig. 17. The P-U curve after being obscured

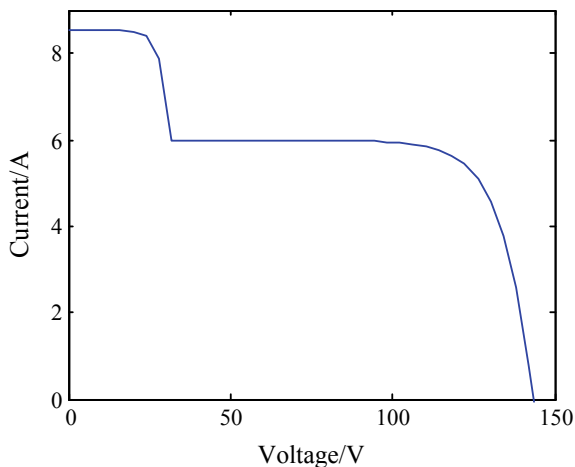


Fig. 18. The I-U curve after being obscured

From Fig. 17, it can be seen that the coordinates of the two peak points are A(28.04, 221.3), B(118.2, 667.5), respectively, and the voltages between the peak points are $\Delta U_{AB} \approx 118.2 - 28.04 = 90.16 \text{ V} \approx 3 \times 0.9 \times U_{oc}$, respectively. There are two steps in Fig. 18.

From the above four sets of simulation results, we can get some characteristics:

1. When the photovoltaic array is partially obscured, the output's P-U curve will appear multiple peaks and I-U curves will appear multiple steps.
2. As we can see from Figs. 11 to Fig. 18, when the photovoltaic array appears several kinds of light, the P-U curve of the photovoltaic array output has several peak points, and the I-U curve has several steps.
3. From Figs. 11, 13, 15 and 17, we can draw a rule that the voltage difference between adjacent peak points is approximately equal to the integer times of $0.9U_{oc}$, in which U_{oc} is the open circuit voltage of the photovoltaic module.
4. It can be seen from the I-U curve of the four groups, the curve on the left side of the I-U curve corresponding to each peak point is about a straight line, the slope is zero; the curve on the right side of the peak point is a curve, the slope is less than zero.

The above characteristics and rules can provide a theoretical basis for the maximum power point tracking in the shadow condition, and improve the dynamic performance of the maximum power point tracking.

4 Conclusion

Through the modeling of photovoltaic cells, the engineering application model of photovoltaic cells is given, and the correctness of the model is verified by MATLAB/Simulink simulation. Second, the output characteristic curve of the photovoltaic array under the partial shadow is simulated by MATLAB/Simulink, and some rules and characteristics of the P-U and I-U curves of the photovoltaic array output are obtained by observing several groups of simulation data. These laws and characteristics can provide a theoretical basis for multi-peak maximum power point tracking (MPPT).

Fund Project. The Key Research Project of Natural Science for Anhui Provincial Universities (KJ2017A747).

References

1. Xu, T.: Research and Analysis of Maximum Power Point Tracking for Photovoltaic Power Generation under Shadow. Lanzhou Jiaotong University (2016)
2. Wang, Q., Xu, H., Zhang, Q., Zhang, H.: Research on improved MPPT based on variable step conductance increment method. *Manag. Technol. Electr. Appl. Energy Effic.* (21), 10–15 + 47 (2017)
3. Liqiao, W., Sun, X.: Photovoltaic Generation Technology in Distributed Generation System, pp. 8–12. Mechanical Industry Press, Beijing (2014)

4. Ningning, C.: Research on Multi-Peak Maximum Power Tracking of Grid-Connected Photovoltaic Inverter. Nanjing University of Science and Technology (2014)
5. Li, S.: Research on Mismatch Analysis and Optimal Control of Photovoltaic System under Shadow Condition. Hefei University of Technology (2016)

Author Biography

Yang Hao Male, lecturer, Master, Research orientation: power electronics and power drives.



A Novel Fiber Polarization Splitter Based on Orthogonal Dual Core

Yu Hou^(✉)

Department of Physics, TangShan Normal University, Tangshan 063000, China
198468hy@163.com

Abstract. A novel polarization splitter is proposed in a photonic crystal fiber with orthogonal dual core. The device is designed by using index matching coupling theory. The proposed splitter exhibits some features, which has short splitting length (the minimum splitting length is 0.3 cm), low loss (the silicon fiber has been widely used in optical fiber communication) and so on. Moreover, it is easy to fabricate over its counterparts because the right core is equivalent to the left core rotated by 90° . It is extremely useful for the future applications of the broadband polarization device.

Keywords: Polarization splitter · Index matching coupling · Photonic crystal fiber

1 Introduction

Polarization splitters are able to split a light signal into two orthogonal polarization states have been attracting much research interest for their potential applications in the fields of optical fiber communications and sensors. Conventional splitters have been achieved by utilizing metal overlap, mode evolution in optical waveguides, e.g., fused biconical photonic crystal fiber coupler [1]. The uprising of photonic crystal fiber (PCF) technology allows for a brand-new light guidance mechanism in a way that could not be ever imagined. Owing to their distinguished optical properties, PCFs offer a new opportunity to control light propagation through a periodic arrangement of refractive indices along the fiber axis. Various types of PCF polarization splitters have been proposed in either two-core or three-core schemes [2, 3]. These devices can be achieved through two approaches: (1) Directional coupler based on two identical cores [1, 2]. This device works due to the fact that the x and y polarization states have different coupling lengths, so a light signal can be completely split into two cores after a certain distance. This device shows a broader operating wavelength but a disadvantage is that the splitting length is longer than other ways. The size of these devices usually is in the order of tens of centimeters. (2) Polarization-selective coupler [4, 5], in which one polarization can couple between the two cores, however, the other one cannot. The splitting length can be effectively shortened by this means. Nonetheless, the bandwidth of these devices is rather narrow because the splitting operation only works in the proximity of the crossing point between the two dispersion curves.

According to the reference [5], the strong coupling only occurs in a wavelength range of a few nanometers.

Based on index matching coupling theory, a new method to achieve a polarization splitting is proposed. It can make one polarization mode couple between two cores in a broadband. It should be noted that the theory of the Index Matching Coupling has been widely used to design two-dimensional photonic crystal waveguides. Fan et al. first developed the Index Matching Coupling theory for the resonant cavity in 2001 [6]. According to this theory, they designed a power splitter with up to 100% transmittance. Thereafter, based on the resonant cavity theory, many research teams designed a good variety of waveguides, such as Y-type waveguide [7], cross-shaped waveguide [8]. A novel polarization splitter with a broader band and shorter splitting length is designed using this method. By numerical simulation, we found that the strongest coupling occurs by the wavelength of $1.55 \mu\text{m}$; the operating bandwidth is $0.9 \mu\text{m}$, which is about several times broader than the result of general reports; the minimum splitting length is 0.3 cm , which is shorter than the conventional directional coupler. This polarization splitter shows more excellent integrated performance than ever before.

2 Design and Characteristics Analysis

To design a polarization splitter using Index Matching Coupling method, the orthogonal structure of the fiber cores should be ensured. In this paper, we adopted dual-core PCF in which the two cores are perpendicular. The cross section of the fiber based on orthogonal dual core is shown in Fig. 1. To facilitate the analysis, the left core is called core A and the right one is called core B.

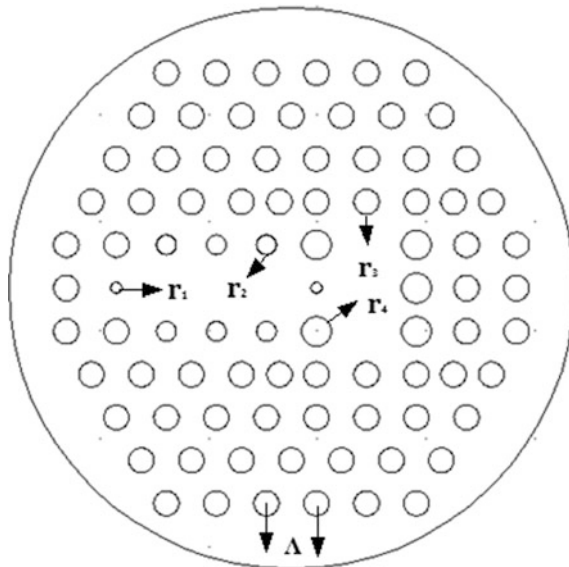


Fig. 1. Cross section of the fiber based on orthogonal dual-core PCF

If one polarization wants to couple between the two cores, it must meet the index matching coupling condition. So we make some adjustments to the size of the air holes near core A and core B. The size of those air holes is exhibited in Table 1. Core A is a rectangle and its long axis in the X direction. We reduce some air holes around the core A, which has two goals: (1) it would make it easier for the selected mode coupling in the two fiber core (2) it can increase the refractive index of the y-polarization mode. The core B is equivalent to the core of core A rotated by 90° , except some air holes in the core B is bigger than core A. The reason for this is that the bigger air holes can reduce the refractive index of y-polarization mode in core B. Thus the converse and overlap of the dispersion curve can be achieved. The size of macro air holes in the fiber core is of the order of magnitude of the optical wave, so in this condition, it can be regarded as optical wave enter to the two vertical placed anisotropic medium. The background material is silica with a constant refractive index of 1.45.

Table 1. The parameters of the fiber

Parameters	Value (μm)	Parameters	Value (μm)
r_1	0.5	r_2	0.9
r_3	1.15	r_4	1.35
A	4.6		

Figure 2. shows the different forms of index curve for the different fiber structures (which are shown in the inset). This figure can also be regarded as a process of making the index curve converse. Figure 2a plots the effective index of the directional coupler with two identical symmetrical cores. The mode of two polarizations is degenerate. In Fig. 2b, the fiber cores are changed to two identical high birefringence cores. Figure 2c shows the index curve of the fiber designed by us. It shows the dispersion diagram of equivalent core A and B. The index curves of x- and y-polarization mode is split because the structure is a high birefringence fibers. The core B is equivalent to the core A rotated by 90 degrees, so the dispersion curve of core B converses to core A. on every wavelength of the operating band, the value of the effective index for comparison is $X_A > Y_A, Y_B > X_B$; To make $Y_A \approx Y_B$, the size of the air holes of core B should be adjusted. Thus in the broad operating band, the y-polarization mode will be strong coupling between the two cores while the coupling of x-polarization mode cannot take place owing to the fact that the index of x-polarization mode keeps a large difference.

In order to validate the above analysis, Fig. 3 shows the coupling states of x and y polarizations at a different wavelength. We can see clearly that, at every wavelength, y-polarization mode is strong coupling between the two cores while the x polarization mode cannot be couple because the indices of x-polarization mode for core A and core B have a large difference.

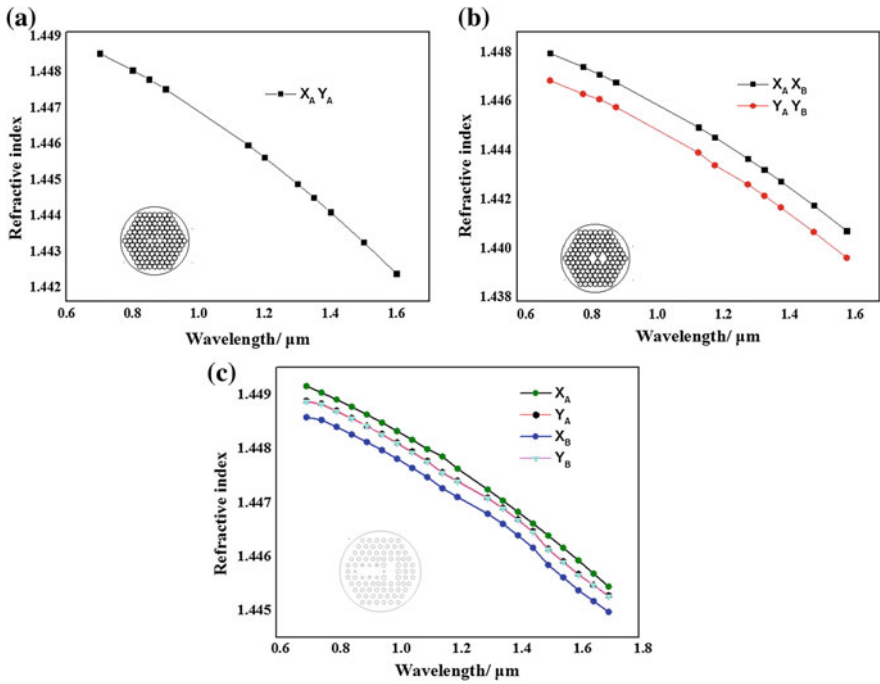


Fig. 2. The different forms of index curve for the different fiber structures. **a** Dispersion curve for the symmetrical structure. **b** Dispersion curve for the asymmetrical structure. **c** Dispersion curve for the fiber design by us

The device we proposed is a polarization-selective coupler based on index matching theory. The core idea is that, to make the coupling of one polarization occur in a broad operating band, meanwhile, the coupling of the other ones cannot take place. Therefore a broadband polarization-selective coupler is constructed. Its splitting length is equal to one coupling length of y-polarization mode. The coupling length is defined as a single-polarization power is completely transferred from one core to the other. It is expressed to be $L_C = \lambda/2|n_e - n_o|$, here the same, n_e and n_o are the effective indices for the polarization even and odd modes, respectively. In this paper, we select y-polarization mode. Through the calculation, the splitting length is shown in Fig. 4. We can see that the coupling length of y-polarization mode gradually decreases with the increase of the wavelength. When the wavelength equals to $1.6 \mu\text{m}$, the coupling length has a minimal value 0.3 cm .

3 Discussion and Conclusion

The working bandwidth of traditional fiber polarization splitters is rather narrow because the splitting operation only works in the proximity of the crossing point between the two dispersion curves. According to the reference [5], the strong coupling

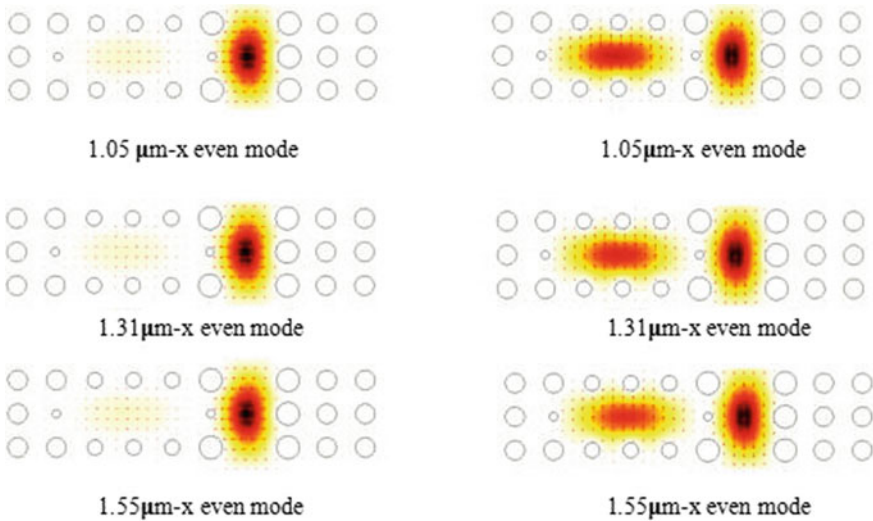


Fig. 3. The coupling states of x and y polarizations at different wavelengths

only occurs in a wavelength range of a few nanometers. In this paper, by subtle design, we make the dispersion curve of y-polarization mode in core B overlap with that of core A. So y-polarization mode can couple between the two cores over a very wide wavelength range. At the same time, x-polarization mode cannot be couple because the indices of x-polarization mode for core A and core B have a large difference. The fiber polarization splitter is extremely useful for the future applications of the coupling device.

In particular, the fiber with orthogonal dual core has an advantage that it is simple to design and easy to fabricate than the counterparts in the communication band. We can easily realize the coincidence of y-polarization index curve by adjusting the size of the hole or the lattice constant. Moreover, the size of the device is small as the light signal can completely split through one coupling length of y-polarization mode; therefore the shorter coupling length is ensured. Simultaneously, this structure itself features ultralow confinement loss and high birefringence [9]. This point thanks to the combined effect of the cores with asymmetric rectangular structure and the cladding with circular air holes of a asymmetric array. Lastly, the background material is silica, which has little absorption to the optical wave, so the loss of this device can be further reduced. Taking all the factors into consideration, the splitter we proposed exhibits some features, which has a short splitting length (the minimum splitting length is 0.3 cm), low loss (the silicon fiber has been widely used in optical fiber communication) and so on. Moreover, it is easy to fabricate over its counterparts because the right core is equivalent to the left core rotated by 90° . It is extremely useful for the future applications of the broadband polarization device.

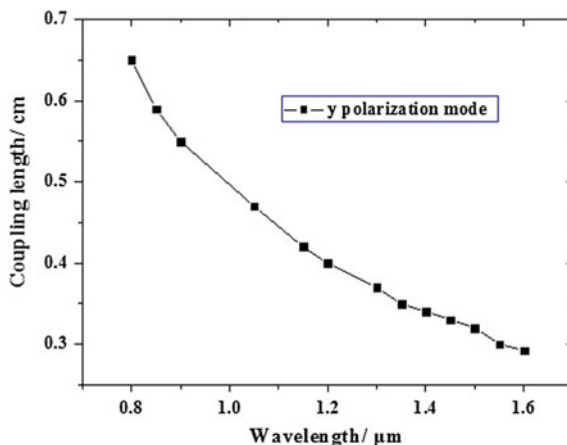


Fig. 4. The splitting length of y polarization

Acknowledgements. The work was jointly supported by the Science and Technology Projects of Hebei Province (Grant No. 16274522).

References

1. Lee, B.H., Eom, J.B., Kim, J., et al.: Photonic crystal fiber coupler. *Opt. Lett.* **27**, 812–815 (2002)
2. Zhang, L., Yang, L.: A novel polarization splitter based on the photonic crystal fiber with nonidentical cores. *IEEE Photon. Technol. Lett.* **16**, 1670–1673 (2004)
3. Xu, Q., Wang, M., Lin, S., et al.: A novel polarization splitter based on dual core photonic crystal fiber. *Optik* **127**, 10442–10449 (2016)
4. Dai, D., Wu, H.: Realization of a compact polarization splitter-rotator on silicon. *Opt. Express* **41**, 2346–7351 (2016)
5. Li, J., Mao, Y., Lu, C., Tam, H.Y.: Polarization splitting of photonic crystal fiber with hybrid guidance mechanisms. *IEEE Photon. Technol. Lett.* **23**, 1358–1361 (2011)
6. Fan, S.H., Johnson, S.G., Joannopoulos, G.D.: Waveguide branches in photonic crystals. *J. Opt. Soc. Am. B.* **16**, 162–165 (2001)
7. Zhou, J., Chang, Q., Mu, D., Yang, J.H.: Theoretical investigation of waveguide power splitters with parallel output ports in two dimensional square-lattice photonic crystals. *J. Opt. Soc. Am. B.* **26**, 2469–2472 (2009)
8. Wu, B., Zhang, H., Wang, Q., Chang, S.J.: Multifunctional photonic crystals cross waveguide for terahertz waves. *J. Opt. Soc. Am. B.* **27**, 505–510 (2016)
9. Cao, Y., Li, R.M., Tong, Z.R.: Investigation of a new kind of high birefringence photonic crystal fiber. *Act. Phys. Sin.* **62**(2013), 098702-098702-5



Theoretical Research on the Strain Characteristics of the Selective-Filling Birefringent Photonic Crystal Fiber

Jin Zuo, Tingting Han^(✉), and Jingping Yang

Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China
hanting608@163.com

Abstract. In this paper, the detailed strain characteristics of a selective-filling birefringent photonic crystal fiber (SF-PCF) were analyzed theoretically. The SF-PCF was achieved by selectively filling two symmetrical air holes around the fiber core of a PCF with a high index liquid. The group birefringence and phase birefringence presented unique characteristics different from the traditional index-guiding birefringence fiber. Hence, the transmission and sensing characteristics of the Sagnac interferometer based on the SF-PCF presented diversity and high sensitivity. Finally, the highest strain sensitivity of $-9.56 \text{ pm}/\mu\epsilon$ was theoretically achieved.

Keywords: Photonic crystal fiber · Birefringent fiber · Selective-filling · Strain sensor

1 Introduction

The photonic crystal fiber is a new type of optical fibers with novel structures, novel guiding mechanism, and flexible design and has been widely used in many versatile sensing technologies [1]. Thereinto, high birefringent (HiBi) PCFs [2, 3] possessing high birefringence and polarization maintaining characteristics, always were applied into the Sagnac Interferometers to measure different parameters, for example, the strain parameter. Furthermore, different HiBi fibers based SIs can exhibit unique strain sensitivities. For example, the index-guiding HiBi PCFs, normally being compounded of only one material, presented temperature-insensitive properties, which can effectively eliminate the problem of cross-sensitivity. And the temperature-insensitive strain sensitivities of $0.23 \text{ pm}/\mu\epsilon$ [4], $1.1 \text{ pm}/\mu\epsilon$ [5], and $2.5 \text{ pm}/\mu\epsilon$ [6] were achieved by different index-guiding HiBi PCFs, respectively. At the same time, the air holes structures of the PCFs provided greater convenience for injection of the active functional materials. This opened up a new area for the realization of HiBi PCFs with unique birefringence and sensing characteristics. Guo et al. [7] illustrated a solid-core polarization maintaining PCF by selectively filling a single hole with a high index

liquid, and achieved 3 pm/μϵ strain sensitivity. In [8, 9], a new type of selective-filling HiBi birefringent fibers was analyzed, and unique birefringence properties were revealed. Furthermore, a different filling pattern determined diverse birefringence properties and sensing characteristics.

In this paper, we theoretically analyzed the birefringence and sensing characteristics of a kind of selective-filling photonic crystal fiber (SFPCF) which is designed by selectively filling two symmetrical air holes around the fiber core of an index-guiding PCF with a high index liquid. The modal effective refractive index, phase birefringence and group birefringence of the fundamental modes were all calculated by using the finite element method. The birefringence presented unique properties different to the traditional index-guiding HiBi PCFs. Furthermore, the SFPCF was applied into a Sagnac interferometer, and the strain sensing characteristics of the SI transmission spectra were analyzed at different strains and wavelengths. Finally, a strain sensitivity of −9.56 pm/μϵ was achieved.

2 Theoretical Simulation of the Birefringence

Figure 1 showed the microgram of the cross section of the SFPCF. The effective index of silica background was 1.444 at 1.550 μm. The diameters of the air holes, the adjacent holes distance, and the diameter of the core were 3.7 μm, 5.8 μm, and 7.9 μm, respectively. This pure silicon PCF consisted of a five-ring air hole arranged in a regular hexagon which can be seen in Fig. 1a as well. The red holes are the filled holes with the index of 1.56 at 1.550 μm.

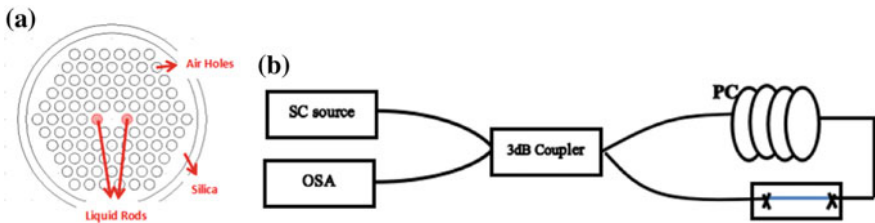


Fig. 1. **a** The microgram of the cross-section of the SFPCF. **b** The Sagnac interferometer device

The SFPCF was placed into the SI device, as shown in Fig. 1b. The SI includes a polarization controller (PC), a 3-dB optical coupler (OC) with a splitting ratio of ~50:50 at a wavelength range from 1.3 to 1.6 μm and a section of SFPCF. And the transmission interference spectrum of the SI can be expressed by the following equation:

$$T_T(\lambda) = \frac{1 - \cos(\delta)}{2} \tag{1}$$

where δ stood for the phase difference and was expressed as

$$\delta = 2\pi LB(\lambda)/\lambda \quad (2)$$

L was the filling length.

The wavelengths of the output interference dips λ_m with minimum value satisfied the phase matched condition:

$$\frac{2\pi B(\lambda_m)L}{\lambda_m} = 2\pi m, m = 0, \pm 1, \pm 2, \dots \quad (3)$$

where m was the integer.

When the axial force F was applied onto the fiber, the produced axial strain ε_z was:

$$\varepsilon_z = F/AE \quad (4)$$

where, A represented the area of the pure silica part in the fiber cross-section, and was calculated as $11213.19 \mu\text{m}^2$. E denoted the elastic modulus, and was about 7.2×10^4 MPa for silica glass. If the external pull was applied to the optical fiber, the refractive index of the silica background can be changed due to the elastic effect, which can be expressed as

$$\frac{\Delta n_{bg}}{n_{bg}} = -\frac{n_{bg}^2}{2} [p_{12} + \nu(p_{11} + p_{12})] \varepsilon_z = -p_e \varepsilon_z \quad (5)$$

where p_{11} and p_{12} were the elastic tensor Kerr coefficients, p_e was a valid elastic-optic coefficient, the value of p_e was 0.22 and it was the elastic coefficient of the silica background.

The calculated strain sensitivity interference dips change with external force sensitivity:

$$S(F) = \frac{d\lambda}{dF} = \frac{\lambda}{B_g} \frac{\partial B}{\partial F} \quad (6)$$

Figure 2a illustrated the curves of the modal effective refractive index of the core fundamental modes which were in the x and y polarized directions with wavelength, respectively. The core energy will leak into the filled liquid rods if the wavelength was close to the boundary of the anti-resonant region which can be seen at about $1.3 \mu\text{m}$. Figure 2b demonstrated the modal field distributions of the modes at some wavelengths. Point a represented the fundamental core mode. Points b and c at around $1.3 \mu\text{m}$ indicated the mode couplings between core fundamental mode LP_{01} and two-high-index-rod waveguide LP_{31} odd modes. Point d represented the mode coupling between core LP_{01} mode and two-high-index-rod waveguide LP_{12} odd mode.

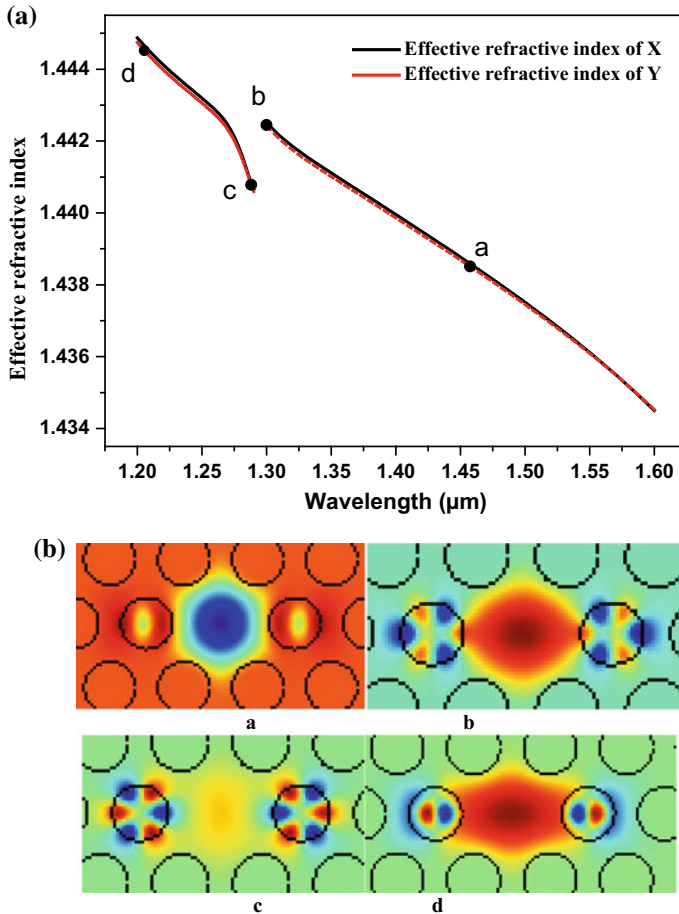


Fig. 2. **a** The curves of the modal effective refractive index of the core fundamental mode with wavelength. **b** The mode field diagram at *a*, *b*, *c*, and *d* wavelengths

3 Strain Response of the Birefringence and SI Transmission Properties

The phase birefringence B and group birefringence B_g were two basic parameters to describe birefringence characteristics. And both the mathematical expressions were shown as follows:

$$B(\lambda) = n_x(\lambda) - n_y(\lambda) \quad (7)$$

$$B_g(\lambda) = B(\lambda) - \lambda \cdot \frac{dB(\lambda)}{d\lambda} \quad (8)$$

According to the two formulas, the strain response of B and B_g were calculated and plotted in Figs. 3a and 4a. In order to observe the image trend clearly, we enlarge the two images which can be seen in Figs. 3b and 4b. As the strain increased, the B shifted to longer wavelength and increased. Figure 4a showed us the B_g had zero points at certain wavelengths.

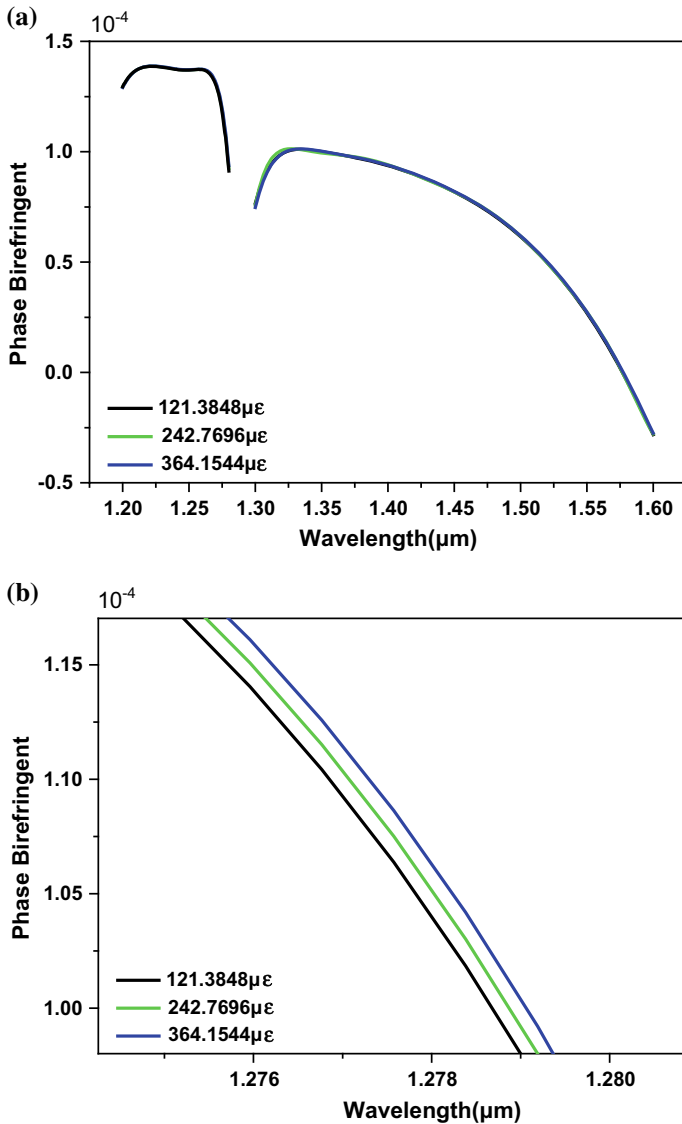


Fig. 3. a Curves of the phase birefringence with wavelength from 121.3848 to 364.1544 $\mu\epsilon$. b The enlarged curves of phase birefringence

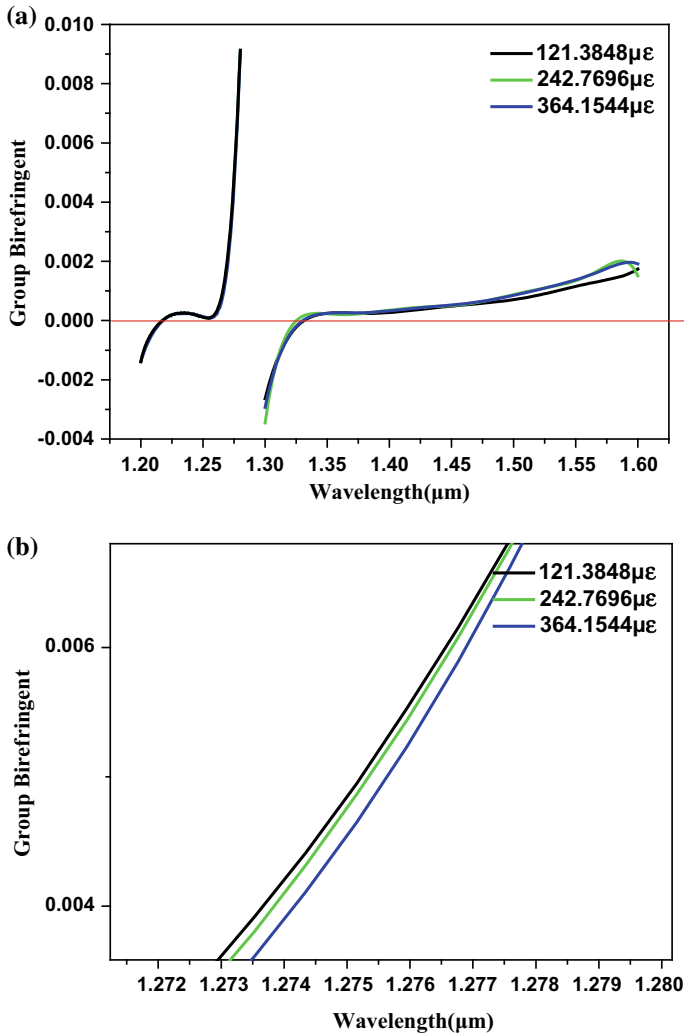


Fig. 4. **a** The curves of group birefringence with wavelength from 121.3848 to 364.1544 μm . **b** The enlarged curves of group birefringence

The output spectra of the SFPCF-based SI from 60.6924 to 242.7696 μm were illustrated in Fig. 5. With the strain increasing, the transmission interference dips will shift to the longer wavelength. Figure 6a showed the wavelength variation of two dips with strain. The data of dips at about 1.38 and 1.435 μm were analyzed and fitted with 7-order polynomial. The Adj. R-Square of them were 0.99018 and 0.99805, respectively. And we took the first-order derivative of these fitting curves which can be seen in Fig. 6b. As can be seen in the graph, the curves showed the fluctuation of the strain sensitivity which ranged from 2.73 $\text{pm}/\mu\text{m}$ to -9.56 $\text{pm}/\mu\text{m}$ and from 1.29 $\text{pm}/\mu\text{m}$ to -3.37 $\text{pm}/\mu\text{m}$, respectively. The maximum strain sensitivity of -9.56 $\text{pm}/\mu\text{m}$ was achieved.

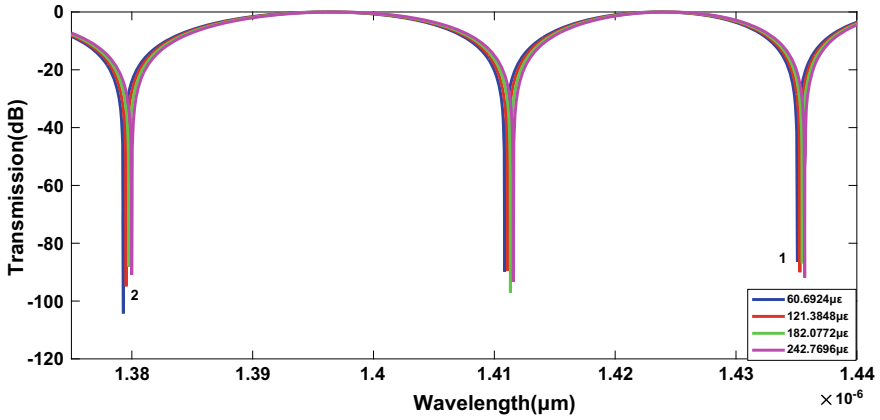


Fig. 5. The output spectra of the SFPCF-based SI from 60.6924 to 242.7696 μm

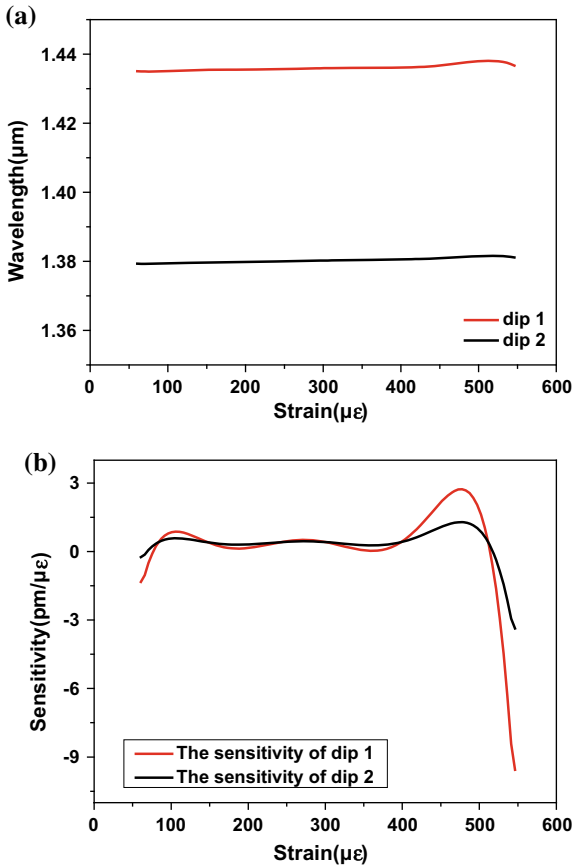


Fig. 6. **a** The specific wavelength variation of the dips with strain. **b** The strain sensitivities of dips 1 and 2

4 Conclusion

In conclusion, we theoretically investigated the strain response of a selective-filling birefringent photonic crystal fiber based SI. And the group birefringence and phase birefringence of the SFPCF were explored. Based on the unique birefringence characteristics, the strain sensitivity of $-9.56 \text{ pm}/\mu\text{e}$ was achieved.

Acknowledgements. This work was supported by the National Natural Science Foundation of China under Grant No. 11404240, and by the Doctoral Scientific Foundation of Tianjin Normal University under No. 52XB1307.

References

1. Culshaw, B.: The optical fibre Sagnac interferometer: an overview of its principles and applications. *Meas. Sci. Technol.* **17**, R1–R16 (2006)
2. Mortimore, D.B.: Fiber loop reflectors. *J. Lightw. Technol.* **6**(7), 1217–1224 (1988)
3. Liu, Y., et al.: High-birefringence fiber loop mirrors and their applications as sensors. *Appl. Opt.* **44**(12), 2382–2390 (2005)
4. Dong, X.Y., Tam, H.Y.: Temperature-insensitive strain sensor with polarization-maintaining photonic crystal fiber based Sagnac interferometer. *Appl. Phys. Lett.* **90**(15) (2007). Art. no. 151113
5. Frazao, O., Baptista, J.M., Santos, J.L.: Temperature-independent strain sensor based on a Hi-Bi photonic crystal fiber loop mirror. *IEEE Sens. J.* **7**(10), 1453–1455 (2007)
6. Zhou, D.P., Wei, L., Liu, W.K., Lit, J.W.Y.: Simultaneous measurement of strain and temperature based on a fiber Bragg grating combined with a high-birefringence fiber loop mirror. *Opt. Commun.* **281**(18), 4640–4643 (2008)
7. Guo, J., Liu, Y., Wang, Z., Han, T., Huang, W., Luo, M.: Tunable fiber polarizing filter based on a single-hole-infiltrated polarization maintaining photonic crystal fiber. *Opt. Exp.* **22**(7), 7607–7616 (2014)
8. Han, T., et al.: Unique characteristics of a selective-filling photonic crystal fiber Sagnac interferometer and its application as high sensitivity sensor. *Opt. Exp.* **21**(1), 122–128 (2013)
9. Han, T., et al.: Control and design of fiber birefringence characteristics based on selective-filled hybrid photonic crystal fibers. *Opt. Exp.* **22**(12), 15002–15016 (2014)



Design and Analysis of a Novel Metamaterial Structure Realized in Low-Frequency Band

Xiu Zhang, Honghao Sun, and Xin Zhang^(✉)

Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
tjnumark@126.com

Abstract. In this paper, the double negative metamaterial is designed and analyzed using finite element method. Most of the structure of metamaterial is applied in radar stealth technology and superlens. Compared with the traditional material, the metamaterial can focus on electromagnetic waves, which can improve the wireless power transfer (WPT) system. In previous research, the metamaterial always operated in GHz and THz which is not suitable for the WPT system. To address this problem, a novel structure of metamaterial which works in MHz is proposed. The simulation results indicate that the novel structure of metamaterial can achieve negative permeability and negative permittivity at about 50 MHz.

Keywords: Metamaterial · Negative permeability · Negative permittivity · Finite element method

1 Introduction

Metamaterial is a kind of supermaterial, which refers to artificial synthetic materials which do not exist in nature. The permeability of ordinary natural materials is positive, and the propagation rules of electromagnetic waves in materials meet the “rule of the right hand”, so they are called right-handed materials. Different from the right-handed materials, the propagation characteristic of the metamaterial is opposite which also called left-handed material [1].

In 1967, the former Soviet physicist, Veselago, studied the electromagnetic properties of the permittivity and magnetic permeability which was negatively worth the material [2]. He found that the magnetic field, electric field, and wave vector met the left-hand rule in a specific frequency band, known as left-handed material. However, in nature, no such permeability and permittivity are found to be left-handed materials, so there has been no progress in the study of left-handed materials for a long time. Until 1996, Professor Pendry demonstrated that the permittivity of a periodically arranged metal rod was negative when the frequency of the electromagnetic wave was less than the plasma frequency. In 1999, he proved that the magnetic permeability and permittivity of the Split Ring Resonator (SRR) in a periodic arrangement were negative at the

resonant frequency [3]. Combining the two materials, in 2001, the properties of left-handed materials are realized in the microwave section. In 2006, Pendry and others proposed the theory that the use of super-materials to change the direction of the electromagnetic field is arbitrary, and the scholars of the University of Duke made a Circular Cloaking Device based on this theory. At present, many academic groups are exploring the mystery of left-handed materials.

Because of its unique electromagnetic properties, left-handed materials have applications in physics, materials science, optics, and mobile communications. The current research in foreign countries mainly focuses on the innovative concept of artificial structure design, and the study of left-handed materials in infrared and visible bands. At home, all universities and institutes began to study this field before and after 2004. At present, the main direction of the work is aimed at basic research on the application of left-handed materials [4].

Although the study of left-handed materials has become a hot topic in the current scientific and technological research, the current research is either still in the primary stage. It is mainly focused on the simulation and experiment of the left-hand material in a simple medium structural unit. The application of real life is not mature, so a great deal of research on the design of the structure is needed. In order to speed up the actual use of left-handed materials, the trend of research is mainly focused on the design of theoretical models and the design of three-dimensional left-handed materials [5].

In this paper, a structure of I-split ring resonator based on SRRs structure is proposed. The structure is placed on the single side of the medium plate [6]. In the case of the electromagnetic wave incident, the metal structure is integrated with the electric resonator and the magnetic resonator on the side of the medium plate, so that the left-handed characteristics of the material are produced in a certain frequency band.

2 Theoretical Analysis

In the simulation study of left-handed materials, the equivalent medium theory is the primary method to describe the results of left-handed materials [7]. It is further the theoretical guidance of the method of extracting the equivalent parameters. In this method, the left-handed materials with a limited size are regarded as homogeneous media with equivalent permittivity and permeability. The scattering parameters obtained from the simulation can extract the equivalent permeability and dielectric constant of the left-hand material, so as to intuitively determine the frequency band position and bandwidth of the left-hand material. Because this method is simple and easy to analyze, and can quantitatively describe the specific parameters of left-handed materials, many organizations use this method to design left-handed materials.

This method was first proposed by the Smith group [8]. If the plane wave is vertically incident to the left-hand material with a thickness of d , the relation between the scattering parameter (S parameter) and the refractive index n of the left-handed material, the impedance z is as follows:

$$S_{11} = S_{22} = \frac{i}{2} \left(\frac{1}{z} - z \right) \sin(nk_0d) \quad (1)$$

$$S_{21} = S_{12} = \frac{1}{\cos(nk_0d) - \frac{i}{2} \left(z + \frac{1}{z} \right) \sin(nk_0d)} \quad (2)$$

where k_0 is free space wave number.

From the formulas (1), (2) the refractive index n and the impedance z of the left-handed material can be obtained as follows:

$$n = \pm \frac{1}{k_0d} \left[\cos^{-1} \left(\frac{1 - S_{11}^2 + S_{21}^2}{2S_{21}} \right) + 2\pi m \right] \quad (3)$$

$$z = \pm \sqrt{\frac{(1 + S_{11})^2 - S_{21}^2}{(1 - S_{11})^2 - S_{21}^2}} \quad (4)$$

where m is a positive integer.

The equivalent permittivity ε and permeability μ of the left-handed material can be calculated from the equivalent refractive index and impedance:

$$\varepsilon = \frac{n}{z} \quad (5)$$

$$\mu = nz \quad (6)$$

3 Design and Analysis

3.1 Model Design

There are two main structural designs of left-handed materials: one is resonant structure, the other is the composite left and right transmission line structures. The resonant structure is the dielectric unit in a certain frequency electromagnetic wave incident, excited the metal structure of the electric resonance and magnetic resonance, to achieve the production of double negative materials. The composite left and right hand transmission line is adjusted by the composition of the transmission line, changing the distribution parameters of the medium, so as to achieve the negative dual nature in the specific frequency band.

The structure of this paper is designed based on the resonant structure. The Split Ring Resonator and metal line are placed on both sides of the medium plate, so that the left-hand characteristics of the composite materials are realized at a specific frequency [9–12]. The structural model material is shown as shown in Fig. 1.

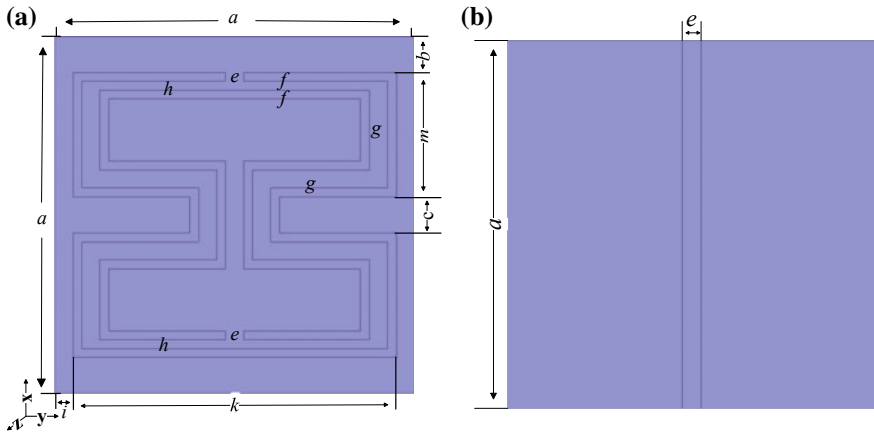


Fig. 1. **a** The front of the model structure; **b** the back of the model structure

The SRR and metal wire are placed on both sides of the medium plate (FR4, $\epsilon = 4.4$, loss tangent of 0.02). The thickness of the medium plate is 2 mm. The metal line and the SRR are made of copper with a thickness of 0.1 mm. Therefore, the thickness of the new structure is 2.2 mm in the experiment. The length of the border length is listed in Table 1.

Table 1. Parameters of the structure

Edge number	a	b	c	e	f	g	h	i	k	m
Length (mm)	200	20	20	10	5	10	5	10	180	70

3.2 Analysis

In order to obtain the electromagnetic properties of the model, the finite element method is applied to analyze the model to get the S parameters and phase parameters of the structure. By observing the position of the passband and the band gap of the S parameters, the frequency band of the left-hand characteristics can be roughly speculated. In order to further obtain specific results, such as permittivity and permeability, the equivalent parameter extraction method mentioned above is applied to calculate.

The model is placed inside a vacuum rectangular air cavity. The metal wire, medium plate, and metal ring are placed in the negative direction of the Z -axis in turn. Medium plate is placed along the XY surface, and the metal bar is placed along the direction of the X -axis. Then, we need to set the boundary conditions, set the two planes parallel to the XY surface as the perfect magnetic wall, and set the two planes parallel to the YZ surface as the perfect electric wall, and the two planes parallel to the XZ surface are set to wave port [13].

The polygonal open metal structure has similar physical properties with the open resonator, which produces the reverse magnetic field and achieves negative permeability μ . The metal wire on the back forms ELC resonance and produces negative permittivity ϵ [14].

Through simulation analysis, the real and imaginary parts of the S parameter can be obtained as shown in Fig. 2.

Using the parameter extraction method mentioned above, the electromagnetic parameters of the model structure are extracted as shown in Fig. 3 which shows the amplitude and phase changes of S parameters respectively. The amplitude and phase of the observed S parameters are abruptly in the vicinity of 50 MHz, and the existence of the left-hand band can be known. Only by changing the S parameters of the two results above, the result is not very reliable. Therefore, put the obtained S parameters into the formulas (3) and (4), and the equivalent impedance z and the equivalent refractive index n are obtained, then we can obtain the permittivity ϵ and magnetic permeability μ [15].

The refractive index, equivalent impedance, permittivity, and magnetic permeability are shown in Figs. 4 and 5. It can be seen that the frequency band of negative refractive index is 46–52 MHz.

Figure 4a shows the curve of refractive index changing with frequency, and determines the frequency band of negative refractive index is 46–52 MHz again. Figure 4b shows the change curve of the wave impedance. It also can be observed that the equivalent permeability is -3.6 , and the permittivity is -1.85 when the frequency is 50 MHz.

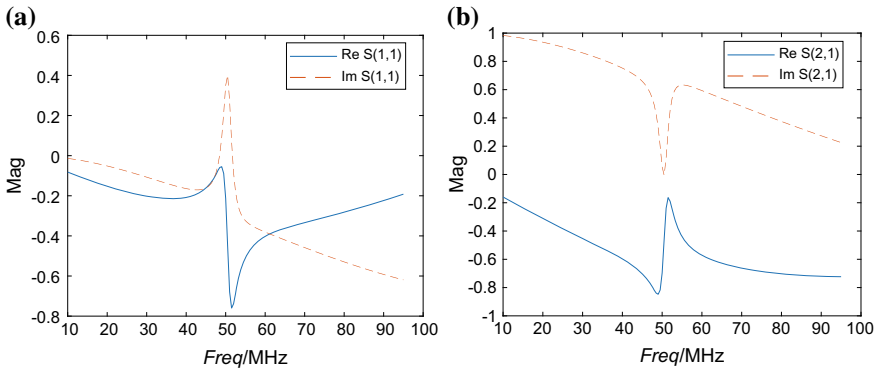


Fig. 2. a Real and imaginary parts of S_{11} ; b real and imaginary parts of S_{21}

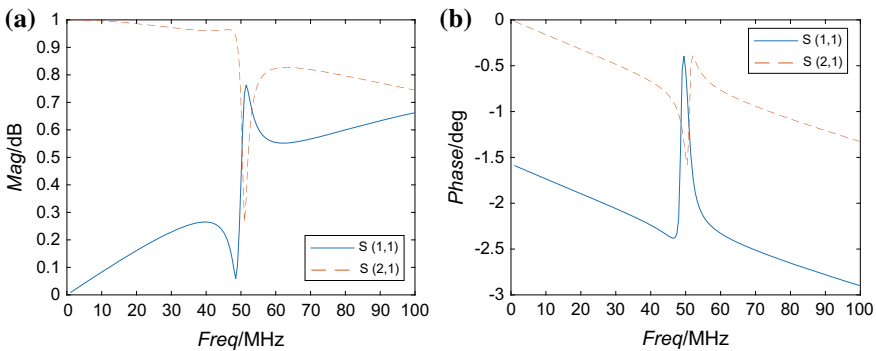


Fig. 3. a The magnitude of the S parameter; b the phase of the S parameter

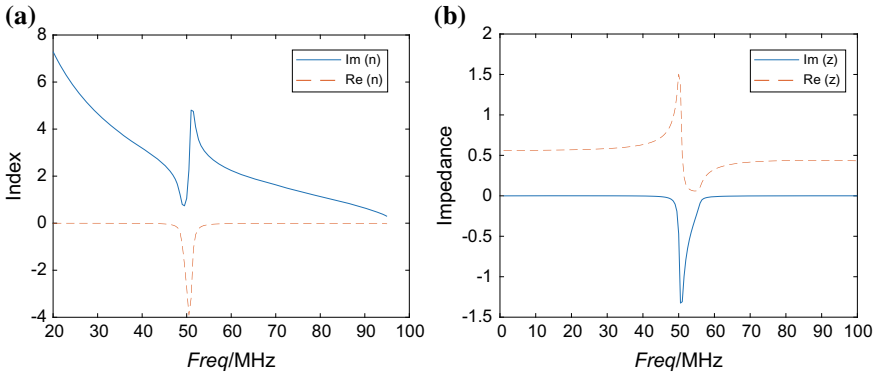


Fig. 4. **a** Refractive index curve; **b** characteristic impedance curve

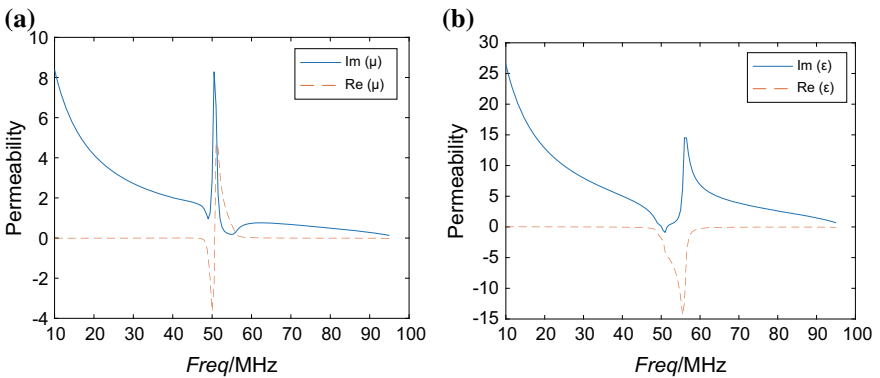


Fig. 5. **a** Permeability of a model structure; **b** permittivity of a model structure

4 Conclusion

In this paper, a new left-handed material model is designed based on the traditional SRR/Metal-wire structure. Through theoretical analysis and simulation, the equivalent index, characteristic impedance, permittivity and permeability of the model structure were extracted by the equivalent parameter extraction method. The results demonstrate that the structure of the model has double negative characteristics in the 46–52 MHz range, and has a great reference value for the wireless power transmission system [16–18].

Acknowledgements. This research was supported by the National Natural Science Foundation of China (61601329) and Applied Basic Research Program of Tianjin (15JCYBJC52300).

References

1. Gong, J.: The structure of left-handed material and its application in microwave engineering. Doctoral dissertation, Xi'an Electronic and Science University (2009)
2. Jing, S., Bin, W.G., Zhenyu, W., Jun, M.F.: Comparison of two methods for extracting electromagnetic parameters of double negative materials. *Comput. Simul.* **25**(5), 329–333 (2008)
3. Smith, D.R., Padilla, W.J., Vier, D.C., Nemat-Nasser, S.C., Schultz, S.: Composite medium with simultaneously negative permeability and permittivity. *Phys. Rev. Lett.* **84**(18), 4184 (2000)
4. Boopalan, G., Subramaniam, C.K.: Frequency dependence of magnetic flux profile in the presence of meta-materials for wireless power transfer. In: *IEEE International Symposium on Circuits and Systems*, pp. 2437–2440. IEEE (2014)
5. Choi, J., Seo, C.H.: High-efficiency wireless energy transmission using magnetic resonance based on negative refractive index meta-material. *Prog. Electromagn. Res.* **106**, 33–47 (2010)
6. Kim, S., Ho, J.S., Poon, A.S.: Midfield wireless powering of subwavelength autonomous devices. *Phys. Rev. Lett.* **110**(20), 203905 (2013)
7. Gong, J.Q., Chu, Q.X.: Effective electromagnetic parameters extraction method for 1D left-handed material. 1–4 (2007)
8. Smith, D.R., Vier, D.C., Koschny, T., Soukoulis, C.M.: Electromagnetic parameter retrieval from inhomogeneous meta-materials. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* **71**(3 Pt 2B), 036617
9. Hossain, M.I., Mri, F., Islam, M.T., Ullah, M.H.: A new wide-band double-negative meta-material for C- and S-band applications. *Materials* **8**(1), 57 (2014)
10. Fan, Y., Li, L.: Efficient wireless power transfer system by using highly sub-wavelength negative-index meta-materials. In: *Wireless Symposium*, pp. 1–4. IEEE (2013)
11. Huang, Y., Tang, H.J., Chen, E.C., Yao, C.: Effect on wireless power transmission with different layout of left-handed materials. *AIP Adv.* **3**(7), 142–146 (2013)
12. Ziolkowski, R.W.: Design, fabrication, and testing of double negative meta-materials. *IEEE Trans. Antennas Propag.* **51**(7), 1516–1529 (2003)
13. Smith, D.R., Schultz, S., Marko, P., Soukoulis, C.M.: Determination of effective permittivity and permeability of meta-materials from reflection and transmission coefficients. *Phys. Rev. B* **65**(19), 195104 (2001)
14. Wang, B., Teo, K.H., Nishino, T., Yerazunis, W., Barnwell, J., Zhang, J.: Experiments on wireless power transfer with meta-materials. *Appl. Phys. Lett.* **98**(25), 254101–254101-3 (2011)
15. Zhao, Y., Vutipongsatorn, V., Leelarasmee, E.: Improving the efficiency of wireless power transfer systems using meta-materials. In: *International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, pp. 1–4. IEEE (2013)
16. Zhang, X., Ho, S.L., Fu, W.N.: Quantitative design and analysis of relay resonators in wireless power transfer system. *IEEE Trans. Magn.* **48**(11), 4026–4029 (2012)
17. Zhang, X., Ho, S.L., Fu, W.N.: Quantitative analysis of a wireless power transfer cell with planar spiral structures. *IEEE Trans. Magn.* **47**(10), 3200–3203 (2011)
18. Zhang, X., Ho, S.L., Fu, W.N.: Analysis and optimization of magnetically coupled resonators for wireless power transfer. *IEEE Trans. Magn.* **48**(11), 4511–4513 (2012)



Evanescent-Mode Waveguide Filter with Transmission Zeroes Created by Shorted Waveguide Shunted in Coupling Region

Liu He^(✉), Xingjian Zhong, Zhendong Fan, Quan Zhang,
and Wei Zhang

College of Communications Engineering, The Army Engineering University
of PLA, Nanjing, China
15651766397@163.com

Abstract. This paper presents a direct method to generate a transmission zero in evanescent-mode waveguide filter by adding a shorted waveguide shunted in the coupling region between two resonators. First, a traditional evanescent-mode waveguide filter with series coupling topology is designed. To introduce a transmission zero to enhance the stopband performance, a shorted waveguide is shunted in the coupling region between two adjacent resonators. Adjusting length of the shunted waveguide can easily adjust the position of the transmission zero. To maintain the bandpass performance of the whole filter, the coupling region needs to be tuned to produce suitable coupling strength as the same as the original one. A fourth-order example is designed and measured, measured results show good accordance with the simulated ones, validating the proposed method in this paper.

Keywords: Transmission zero · Evanescent mode · Shorted waveguide · Fourth order

1 Introduction

In recent years, the performance of filters becomes more demanding than ever before, which requires the filters should possess small size, high selectivity, low insertion loss, and so on. In the late 1950s, Jaynes [1] and Edson [2] proposed resonators in evanescent-mode waveguide to design filters, which are called evanescent-mode waveguide filters now. Compared with general coupled resonant filters, the spurious response of evanescent-mode waveguide bandpass filters could be far away from the center frequency. Thus, these filters are widely used in satellite communication systems.

There are two common kinds of evanescent-mode waveguide filter, evanescent-mode E-plane finned waveguide bandpass filter [3] and evanescent-mode waveguide bandpass filter with non-touching-E-plane fins [4]. In this paper, by means of introducing capacitive posts in E-plane [5, 6], a new structure is proposed as shown in Fig. 1, the equivalent circuit model of the filter is shown in Fig. 2.

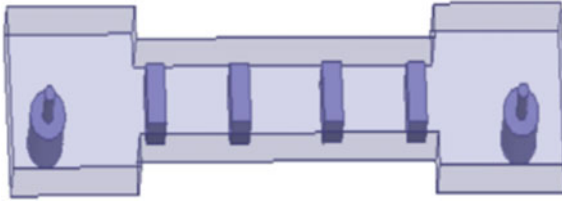


Fig. 1. Evanescent-mode waveguide filter

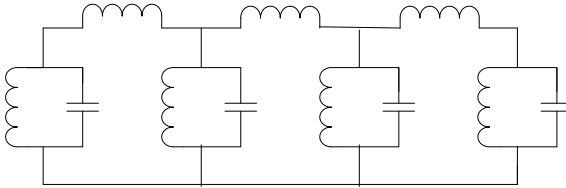


Fig. 2. Equivalent circuit model

2 Filter Design

Basic structure of evanescent-mode waveguide bandpass filter is shown in Fig. 1. Four capacitive posts are placed at suitable intervals along the evanescent waveguide to realize fourth order within the passband. The input and output of the filter are realized by coaxial feedings. Cut-off wavelength of the waveguide can be computed as

$$\lambda_c = c/f_c \tag{1}$$

The fundamental mode propagating in the waveguide is TE₁₀ mode whose cut-off wavelength is $2a$ (a is the width of the waveguide’s broadside), a can be computed as $\lambda_c/2$. The size of the capacitive posts and the intervals is determined by the coupling coefficient matrix. A 6×6 normalized coupling matrix corresponding to a fourth-order Chebyshev filter [7] is shown in the following:

$$M = M_{ij} = \begin{bmatrix} 0 & 1.1289 & 0 & 0 & 0 & 0 \\ 1.1289 & 0 & 1.013 & 0 & 0 & 0 \\ 0 & 1.013 & 0 & 0.75598 & 0 & 0 \\ 0 & 0 & 0.75598 & 0 & 1.013 & 0 \\ 0 & 0 & 0 & 1.013 & 0 & 1.1289 \\ 0 & 0 & 0 & 0 & 1.1289 & 0 \end{bmatrix} \tag{2}$$

Coupling coefficient $K_{i,i+1}$ can be computed by the following equations:

$$K_{i,i+1} = \frac{f_2 - f_1}{\sqrt{f_2 \times f_1}} \tag{3}$$

$$K_{i,i+1} = FBW \times m_{i,i+1} \tag{4}$$

where FBW is fractional bandwidth of the filter and $m_{i,i+1}$ is the corresponding elements in the normalized coupling matrix. Adjust the intervals between posts until $K_{i,i+1}$ in (3) equals or approaches closely to $K_{i,i+1}$ in (4). Then, a shorted waveguide is shunted in the coupling region between the second and the third resonators as shown in Fig. 3.

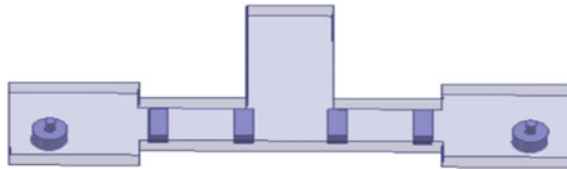


Fig. 3. Simulated model proposed in this letter by introducing a shunted waveguide

Length of the shorted waveguide is about $1/2 \lambda_g$, λ_g is waveguide length of the transmission zero's frequency and can be computed as

$$\lambda_g = \frac{\lambda}{\sqrt{1 - \left(\frac{\lambda}{\lambda_c}\right)^2}} \tag{5}$$

Due to the introduced shorted waveguide, the coupling between the second and the third resonators will change. Then tune the distance between the two resonators until obtaining the same bandpass performance.

An example is given to show the transmission zero created by the shorted waveguide as shown in Figs. 4 and 5. Width and height of the short waveguide's cross-section are 14 and 5 mm, length of the waveguide is set to 17.75 mm.

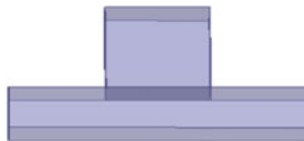


Fig. 4. A shorted waveguide shunted at the center of an evanescent-mode waveguide

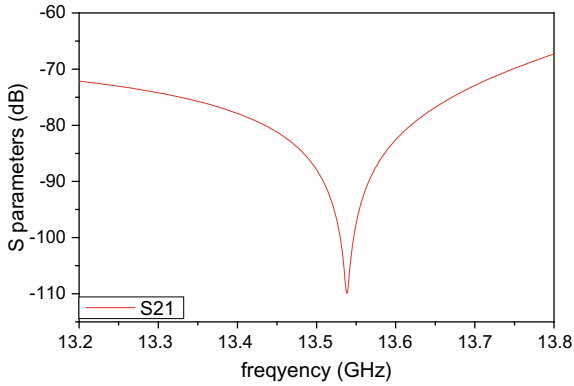


Fig. 5. Simulated results of Fig. 4

From Fig. 5, a transmission zero is created at a frequency of 13.54 GHz, half wavelength of this frequency is 18.1 mm, which is a little different from 17.75 mm. This small difference is caused by the influence between the two waveguides.

3 Result

After optimizing the filter by HFSS, parameters are determined and shown in Fig. 6, simulated results are shown in Fig. 7.

Figure 8 gives the magnetic field distribution at 13.58 GHz in the whole filter structure. From the figure it shows clearly that the signal at this frequency is almost all reflected to the input port.

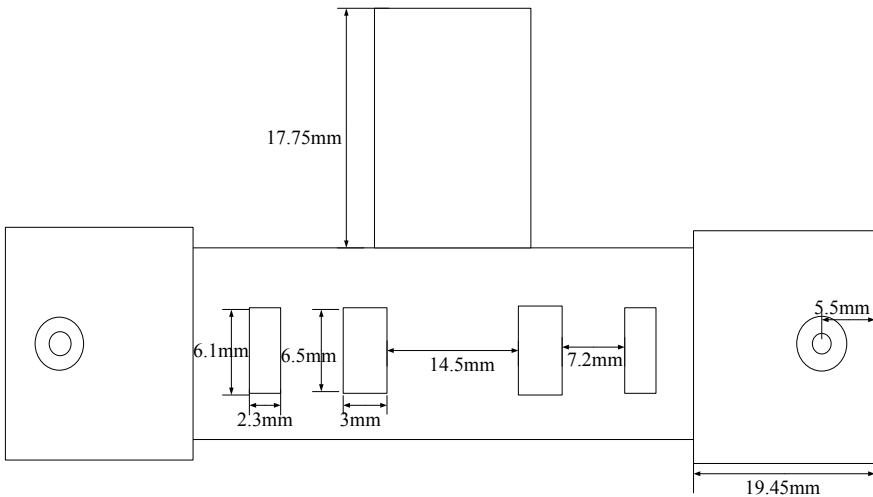


Fig. 6. Key parameters of the simulated model

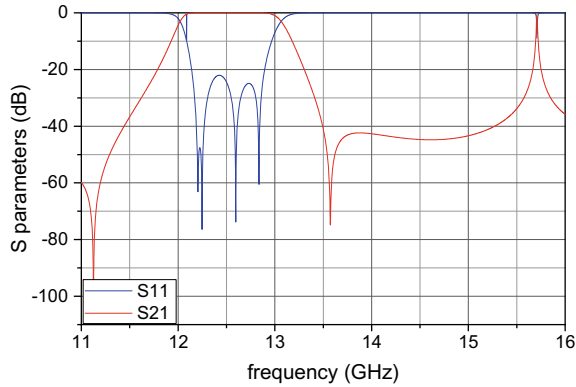


Fig. 7. Simulated results of the filter

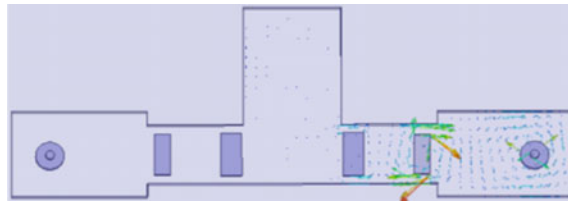


Fig. 8. Magnetic field analysis of solution frequency at 13.58 GHz

The designed filter is then fabricated and measured. Figure 9 shows the manufactured filter. Figure 10 gives the simulated and measured results. The measured results accord well with the simulated ones. The measured transmission zero appears at 13.6 GHz, shifting a little from the simulated one, making upper stopband steeper. Due to fabrication error, the measured insertion loss is about 1 dB, not as good as the simulated. The losses of two SMA connectors also contribute to the measured insertion loss.

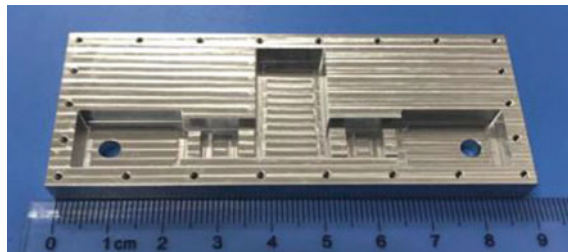


Fig. 9. Manufactured filter

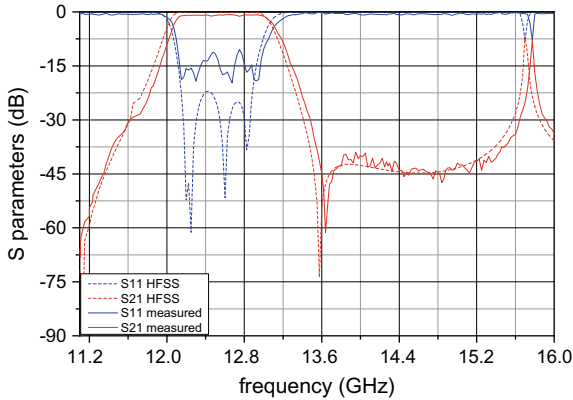


Fig. 10. Simulated and measured results of the filter

4 Conclusion

A simple and direct method to generate a transmission zero in evanescent-mode waveguide filter by adding a shorted waveguide shunted in the coupling region between two resonators has been proposed in this letter. No cross coupling is needed, the whole filter could be simply designed by two steps: a traditional evanescent waveguide filter design and introducing of a shorted waveguide to produce a transmission zero. The measured results are in good accordance with the simulated ones, validating the proposed method presented in this letter.

References

1. Jaynes, E.T.: Ghost modes in imperfect waveguides. In: Proceedings of IRE, vol. 46, pp. 416–418, Feb 1958
2. Edson, W.A.: Microwave filters using ghost-mode resonance. In: Proceedings of IRE Electronic Components Conference, vol. 19, p. 2 (1961)
3. Bornemann, J., Arndt, F.: Rigorous design of evanescent-mode E-plane finned waveguide bandpass filter. In: Proceedings of IEEE MTT-S Digest, pp. 603–606 (1989)
4. Zhang, Q., Itoh, T.: Computer-aided design of evanescent-mode waveguide filter with non-touching E-plane fins. *IEEE Trans. Microw. Theory Tech.* **MTT 36**, 404–412 (1988)
5. Snyder, R.V., Bastioli, S.: Broad passband, wide stopband, high power evanescent mode filters using capacitively-loaded ridges. In: 42nd European Microwave Conference, 2012, pp. 176–179
6. Kun, L., Dexin, Q., Xingjian, Z., et al.: Analysis of stopband performance of evanescent mode waveguide bandpass filter loaded with tiny post. *Appl. Electron. Tech.* **43**(8), 55–57, 65 (2017)
7. Caneron, R.J.: General coupling matrix synthesis methods for Chebyshev filtering functions. *IEEE Trans. MTT* **47**(4), 433–442 (1999)



Time-Aware Routing and Spectrum Assignment Assisted by 3D-Spectrum Auxiliary Graph in Elastic Optical Networks

Li Zhang, Cunqian Yu^(✉), and Rongxi He

College of Information Science and Technology, Dalian Maritime University,
Dalian 116026, China
yucunqian@163.com

Abstract. With setting up and tearing down of dynamic requests, there will inevitably be spectrum fragmentation in elastic optical networks, leading to a higher blocking ratio. However, conventional algorithms only consider current spectrum fragmentation in the 2D-spectrum auxiliary graph, which determines whether requests can be accepted when they arrive. To overcome this problem, we propose an innovative routing and spectrum assignment algorithm considering the future fragmentation, which determines whether requests can be accepted when other online requests leave. In this algorithm, we first introduce a time variance metric to measure the occupied holding time of both the current request and its adjacent online requests, and then propose a model with the minimal time variance in the 3D-spectrum auxiliary graph, which tries to ensure all the leaving time is similar around the current requests. Finally, we verify the advantages of the proposed algorithm in terms of blocking ratio, fragmentation ratio, and spectrum utilization ratio through simulation results.

Keywords: Elastic optical network · Dynamic routing and spectrum assignment · Spectrum fragmentation · 3D-spectrum auxiliary graph · Time variance

1 Introduction

Elastic optical networks (EONs) [1] have been regarded as the future development direction of optical transmission network because of orthogonal frequency division multiplexing (OFDM) technology, which can expand spectrum ranges based on orthogonal phase between subcarriers, and can support fine-grained service integration. The routing and spectrum assignment (RSA) [2–7] in EONs refers to establishing a connection for requests and allocating corresponding spectrum resources under the constraints of spectrum continuity, contiguity, and nonoverlapping. However, due to setting up and tearing down of requests as network traffic varies over time, there are many spectrum fragments formed. To address this issue, a lot of algorithms [6, 7] have been proposed, which are divided into two aspects: proactive and reactive, but those algorithms considered only current spectrum fragments but not the future spectrum fragments (FSFs) whether in the spectral domain or the spatial domain. This is because the focus of these algorithms was whether new requests can be accepted into the

network or not when they arrived. As for FSFs, it determines whether the online request's bandwidth can be occupied by other requests when they are released. The authors in [8, 9] proposed spectrum assignment algorithms in the time-spectral domain that avoid forming FSFs. Unfortunately, they did not consider time-spatial domain, that is, when the online requests left the network, they would generate spectrum fragments in their neighbor links.

To this end, when placing an incoming request, we must ensure its holding time is similar not only with the remaining time of other adjacent online requests in candidate links, but also with other online requests of the same place in neighbor links of candidate links. Therefore, we introduce a time variance metric to measure whether their occupied time is similar, and establish a model with the minimal time variance in the 3D-spectrum auxiliary graph. This model ensures all adjacent connections almost leave the network at the same time, and makes the necessary waiting time smaller between the adjacent frequency slots (FSs) when one of them has been released, which in turn causes less FSFs in the time-spectral and time-spatial domains. Thus, more and more requests will be accommodated in the future. Compared with the traditional algorithms shown in Fig. 1a, we can see that our algorithm shown in Fig. 1b makes the time distribution of online requests more uniform, where *Link1* belongs to the selected path for the current request, and *Link2* and *Link3* are the neighbor links of *Link1*. The vertical axis and horizontal axis are the occupied time and FSs of the online requests, respectively. Besides, the height of every FS is not only similar to its adjacent FSs in the same link, but also with its adjacent FSs in neighbor links. Only in this way, can we reduce FSFs formed.

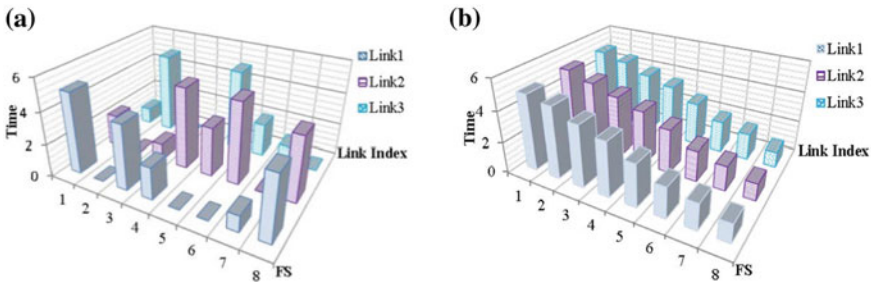


Fig. 1. a The traditional 3D auxiliary graph; b the time-aware 3D auxiliary graph

In this work, we investigate the time-aware routing and spectrum assignment (TRSA) assisted by the 3D-spectrum auxiliary graph in EONs. To the best of our knowledge, this paper is the first work proposing time variance metric and establishing RSA model with the minimal time variance, and we can reduce FSFs in the time-spectral and time-spatial domains effectively. Furthermore, we propose a novel dynamic algorithm, and then verify the advantages of the algorithm by simulation results. Compared with traditional algorithms, our proposed algorithm effectively reduces blocking ratio, fragmentation ratio and increases spectrum utilization ratio.

2 Minimal Time Variance RSA Model

The EONs topology can be represented as $G = \{V, E, C\}$. We use $R\{s_R, d_R, B_R, t_a^R, t_e^R\}$ to denote the request, which t_a^R is the arrival time and t_e^R is the ending time. Meanwhile, we define the holding time $H_R = t_a^R - t_e^R$. Before describing the problem in detail, we need to add the following notations listed in Table 1.

Table 1. Notations

Notations	
V : The set of bandwidth-variable switching nodes;	l_i^k : The i th link of p_k , and $l_i^k \in \phi_k$;
E : The set of bidirectional optical fiber links between nodes in V ;	ϕ_k : The neighbor link set of p_k , and $ \phi_k $ is the number of neighbor links of p_k ;
C : The set of FSs on each optical fiber link in E ;	p_k : The k th alternative path in the k -shortest paths;
ϕ_k : The link set of light path p_k , and $ \phi_k $ is the number of links in p_k ;	$V_{time-spec}^{B_{s,e}^k}$: The variance of $B_{s,e}^k$ in the time-spectrum domain;
L_x^k : The x th neighbor link of p_k , and $L_x^k \in \phi_k$;	$V_{sum}^{B_{s,e}^k}$: The variance sum of $B_{s,e}^k$;
H_{b+1}^k : The remaining time of FS # $b + 1$ in link l_i^k , and if FS is not occupied, $H_{b+1}^k = 0$;	H_{s-1}^k : The remaining time of FS # $s-1$ in link l_i^k , and if FS is not occupied, $H_{s-1}^k = 0$;
$B_{s,e}^k$: The free spectrum subcarrier of p_k , and the first index number and the last index number are s and e , respectively;	H_y^k : The remaining time of online request which occupies FS # y in the neighbor link L_x^k , and $y \in B_{s,e}^k$;
$AV_{l_i^k}^{s,e}$: The average time of the incoming request R and online requests on the both side of incoming request's position in link l_i^k ;	$AV_y^{s,e}$: The average time of the incoming request R and online requests which occupy the y th FS in ϕ_k , and $y \in B_{s,e}^k$;
$V_{min}^{B_{s,e}^k}$: The minimum time variance among free spectrum subcarriers	$V_{time-spa}^{B_{s,e}^k}$: The variance of $B_{s,e}^k$ in the time-spatial domain

In the dynamic scenario, the problem of fragmentation must be considered, especially FSFs, which cannot be occupied by other requests when online requests leave the network. In this section, we will establish an RSA model with the minimal time variance, which focuses on placing an incoming request whose holding time is similar with the remaining time of other online requests in the candidate path and its neighbor links, and reducing the FSFs in the time-spectrum domain and the time-spatial domain respectively.

In the time-spectral domain, we firstly aim at proposing a time variance model, which calculates the time variance between the holding time of the incoming request and the remaining time of its adjacent online requests in candidate links. The time variance equation can be achieved by Eqs. (1) and (2).

Also, we establish a time variance model about the time-spatial domain, which can calculate the time variance between the holding time of the incoming request and the remaining time of other online requests of the same place in neighbor links of candidate links, it can be achieved by Eqs. (3) and (4).

Moreover, the smaller the variance is, the more similar the occupied time of the adjacent FSs. In the 3D-spectrum auxiliary graph, we should calculate the variance sum between time-spectral domain and time-spatial domain through Eq. (5), and then select the minimum summary value through Eq. (6). In this way, we can achieve own objective that minimizes the time variance between the time-spectral and the time-spatial domains, and ensures the leaving time is similar around the position of the incoming request.

$$AV_{i_i}^{s,e} = H_{s-1}^{i_i} + H_R + H_{b+1}^{i_i} / 3 \quad (1)$$

$$V_{time-spec}^{B_{s,e}^k} = \sum_{i_i \in \phi_k} \left(H_{s-1}^{i_i} - AV_{i_i}^{s,e} \right)^2 + \left(H_R - AV_{i_i}^{s,e} \right)^2 + \left(H_{b+1}^{i_i} - AV_{i_i}^{s,e} \right)^2 / 3 |\phi_k| \quad (2)$$

$$AV_y^{s,e} = \left(H_R + \sum_{L_x^k \in \phi_k} H_y^{L_x^k} \right) / (|\phi_k| + 1) \quad (3)$$

$$V_{time-spa}^{B_{s,e}^k} = \sum_{y=s}^b \left(\left(H_R - AV_y^{s,e} \right)^2 + \sum_{L_x^k \in \phi_k} \left(H_y^{L_x^k} - AV_y^{s,e} \right)^2 \right) / B_R (|\phi_k| + 1) \quad (4)$$

$$V_{sum}^{B_{s,e}^k} = V_{time-spa}^{B_{s,e}^k} + V_{time-spec}^{B_{s,e}^k} \quad (5)$$

$$V_{min}^{B_{s,e}^k} = \text{Min} \left\{ V_{sum}^{B_{s,e}^1}, V_{sum}^{B_{s,e}^2}, \dots, V_{sum}^{B_{s,e}^k} \right\} \quad (6)$$

3 Heuristics

The TRSA is a dynamic algorithm that can handle the arrival events and departure events of requests randomly. We firstly initialize network resources, enter network topology $G = \{V, E, C\}$, and wait for the requests to arrive. When a request arrives at the system, we should determine whether the current request is an arrival event or a departure event. If it is a departure event, spectrum resources are released and the FSs status of the link is updated. If it is an arrival event, we should perform the following steps shown in Table 2. In TRSA algorithm, we will select the optimal subcarrier whose time variance sum is the least value in the 3D-spectrum auxiliary graph, and it can reduce the blocking ratio and fragmentation ratio effectively, thus the spectrum utilization ratio is improved greatly.

We first regard the remaining spectrum resources in each fiber link as the link cost, and use the k -shortest path algorithm to find the k best paths between the source and destination nodes as the alternative path set $P_K = \{p_k\}$. In the alternative paths, there may be multiple candidate subcarriers $B_{s,e}^k$, i.e., the continuously available spectrum segments. If P_k is not empty, we will select $B_{s,e}^k \left(B_{s,e}^k \geq B_R \right)$ and calculate its time variance sum of $B_{s,e}^k$, then put all eligible candidate subcarriers into the set λ . Otherwise, we will block R . If the set λ is not empty, we can select the least time variance sum $V_{\min}^{B_{s,e}^k}$ and assign FSs for R . Otherwise, we will block R .

Table 2. TRSA toward the arrival event

1. Calculate the k shortest paths for R and add them to the set p_k ;
2. **For** each candidate path in p_k **do**;
 For each candidate spectrum subcarrier $B_{b,e}^k \geq B_R$ in candidate path **do**;
 Count the variance sum $V_{sum}^{B_{b,e}^k}$ of $B_{b,e}^k$ and add them to the set λ ;
 End For
 End For
3. Select the minimum variance sum $V_{\min}^{B_{s,e}^k}$ from set λ for R ;

4 Performance Evaluation

In this section, we take platform Visual C, at the 64-bit Core i-7 (2.50 GHz) computer with 8 GB RAM to evaluate the performance of our TRSA assisted by the 3D-spectrum auxiliary graph, and the simulation is carried out in NSFNET with 14 nodes and 21 bidirectional fiber links with 358 FSs. We consider the traffic dynamical arrival set of 100,000 requests, which are generated according to a Poisson distribution process with an average arrival rate of λ and an average reservation holding time of $1/\mu$. The source and destination nodes are uniformly distributed throughout the network. The network load is measured in Erlang λ/μ , and we define $\mu = 1$ here. The allowed number of FSs per demand ranges from 1 to 10, and is selected randomly with equal probability, and the traffic load ranges [100, 400] Erlangs.

In order to evaluate the performance of TRSA algorithm, we consider two benchmarks: the minimal holding time difference (MHTD) algorithm [8] and holding time-aware RSA (HTA-RSA) algorithm [9]. MHTD algorithm tries to search for a path with the minimum holding time difference in time-spectral domain, and HTA-RSA algorithm attempts to select a spectrum subcarrier which has the minimum holding time difference sum between FSs and FS with the maximum holding time in their spectrum subcarrier.

We adopt blocking ratio, spectrum utilization ratio and fragmentation ratio as simulation metrics. We define blocking ratio as the total blocked requests divided by the total requests, and define spectrum utilization ratio as the FSs used by the total FSs in the network. Fragmentation ratio is calculated in the Eq. (7), where b_l is the number of spectrum subcarriers of every link, and S_l is the number of available FSs of every link.

$$F = \frac{1}{|E|} \sum_{l \in E} \frac{b_l}{S_l} \tag{7}$$

Figure 2 shows the blocking ratio among three algorithms, where blocking ratio values of the HTA-RSA and the TRSA algorithms are 0 when the traffic load is 100. We can see that the blocking ratio of the three algorithms increases with the increasing of load. The MHTD algorithm has the highest value, followed by the HTA-RSA, and the TRSA has the lowest one. This is because MHTD and HTA-RSA only consider whether the remaining time of the incoming request is similar to its adjacent FSs in candidate path. When the online requests leave the network, it will provide the next request with insufficient spectrum resource if this request is placed in neighbor links of the online requests. On the contrary, the TRSA algorithm not only considers whether the remaining time of the incoming request is similar with its adjacent FSs in candidate path, but also fully considers whether is similar with other online requests of the same place in neighbor links. So the TRSA algorithm has the lowest blocking ratio.

Figure 3 shows the fragmentation ratio among three algorithms, where the MHTD algorithm has the highest value, followed by the HTA-RSA, and the TRSA has the lowest value. This is because MHTD and HTA-RSA do not consider the fragments formed in time-spatial domain when assigning FSs.

Figure 4 shows spectrum utilization ratio among the three algorithms. We can see that the MHTD algorithm performs better than the TRSA algorithm, and the HTA-RSA is lowest at the light traffic loads, but the TRSA algorithm has the highest value, followed by the MHTD algorithm, and the HTA-RSA algorithm has the lowest value at the high traffic load. This is because more fragments are formed with the increasing of load in the MHTD algorithm, and blocking ratio becomes more and more higher. The TRSA algorithm has lower blocking ratio and fragmentation ratio, so the spectrum utilization ratio is higher than the other algorithms.

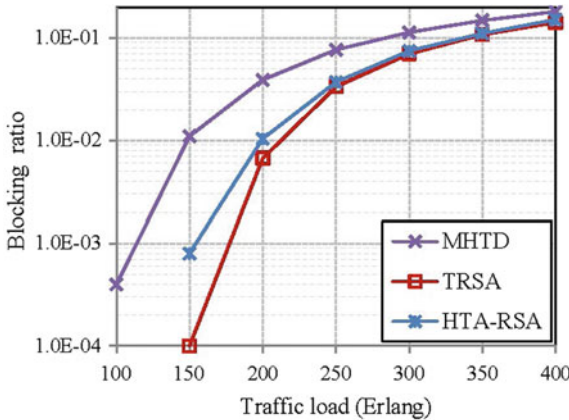


Fig. 2. The blocking ratio in NSFNET

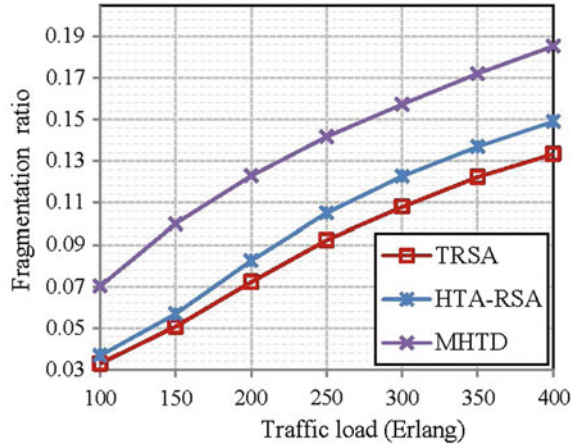


Fig. 3. The fragmentation ratio in NSFNET

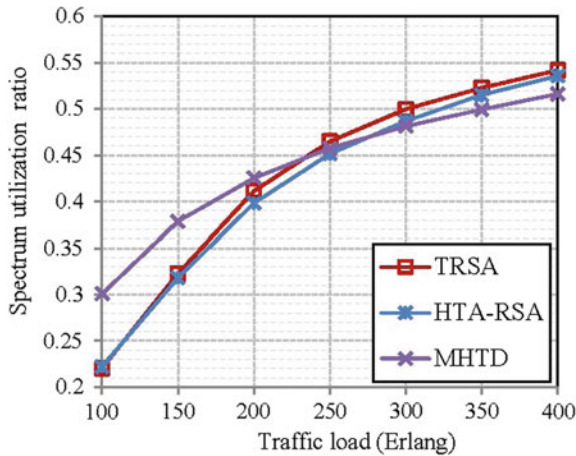


Fig. 4. The spectrum utilization ratio in NSFNET

5 Conclusions

In this paper, a new time variance metric is designed for measuring the occupied holding time of adjacent FSs, and a model with the minimal time variance in the 3D-spectrum auxiliary graph is further established. Based on the model, TRSA algorithm assisted by the 3D-spectrum auxiliary graph in EONs is proposed and simulated. The simulation results show that the proposed algorithm outperforms obviously than the other algorithms in terms of blocking ratio, spectrum utilization ratio, and fragmentation ratio.

Acknowledgements. This work was supported in part by the National Natural Science Foundation of China under Grant 61371091, the “13th five-year” Key Research Project of DMU under Grant 3132016318, and the Fundamental Research Funds for Central Universities under Grant 3132017078.

References

1. Abkenar, F.S., Rahbar, A.G.: Study and analysis of routing and spectrum allocation (RSA) and routing, modulation and spectrum allocation (RMSA) algorithms in elastic optical networks (EONS). *Opt. Switch. Netw.* **23**, 5–39 (2017)
2. Xuan, H., Wang, Y., Xu, Z., Hao, S., Wang, X.: New bi-level programming model for routing and spectrum assignment in elastic optical network. *Opt. Quantum Electron.* **49**(5), 186 (2017)
3. Chen, H., Zhao, Y., Zhang, J., He, R., Wang, W., Wu, J.: Time-spectrum consecutiveness based scheduling with advance reservation in elastic optical networks. *IEEE Commun. Lett.* **19**(1), 70–73 (2015)
4. Zhang, G., Chen, H., Zhang, J., Yu, X., Li, Y., Zhao, Y.: Demonstration of multi-domain spectrum defragmentation with minimum controller-participation degree in elastic optical networks. *IEEE/OSA J. Opt. Commun. Netw.* **9**(9), 782–791 (2017)
5. Wang, N., Jue, J.P.: Holding-time-aware routing, modulation, and spectrum assignment for elastic optical networks. In: *Global Communications Conference*, pp. 2180–2185. IEEE (2015)
6. Liu, H.L., Lv, L., Chen, Y., Wei, C.: Fragmentation-avoiding spectrum assignment strategy based on spectrum partition for elastic optical networks. *IEEE Photonics J.* **9**(5), 1–13 (2017)
7. Zhang, H., Xia, M., Zhang, M., Yoo, S.J.B., Dahlfors, S., Yin, Y.: Spectral and spatial 2D fragmentation-aware routing and spectrum assignment algorithms in elastic optical networks [invited]. *IEEE/OSA J. Opt. Commun. Netw.* **5**(10), A100–A106 (2013)
8. Yang, F., Wang, L., Chen, X., Zhao, Y., Zhang, J.: A holding-time-aware routing and spectrum allocation algorithm in elastic optical network. In: *Opto-Electronics and Communications Conference*, pp. 1–3 (2017)
9. Singh, S.K., Jukan, A.: Non-disruptive spectrum defragmentation with holding-time awareness in optical networks. In: *International Conference on Optical Network Design and Modeling*. IEEE (2016)



Routing and Spectrum Assignment for Software-Defined Elastic Optical Networks

Ziwei Lin, Rongxi He^(✉), and Tongtong Liu

College of Information Science and Technology, Dalian Maritime University,
Dalian 116026, China
hrx@dmlu.edu.cn

Abstract. Software-defined Elastic Optical Networks (SD-EONs) combine software-defined network (SDN) with elastic optical networks (EONs) to further improve the spectrum utilization of routing and spectrum assignment (RSA) algorithms. However, some existing RSA algorithms usually only consider the hops when solving route subproblem, whereas the link spectrum status is also very important. At the same time, they usually neglect the vertical fragmentation between neighbor links during the spectrum assignment. Therefore, we propose an innovative hop and consecutiveness-based routing and fragmentation-aware spectrum assignment (HCR-FSA) algorithm with a joint consideration of hops and spectrum consecutiveness in routing selection and pays attention to both horizontal and vertical fragmentations during spectrum assignment. We implement our HCR-RSA algorithm in RYU controller to evaluate its performance. The experimental results show that our algorithm achieves better blocking probability and fragment degree compared with the previous works.

Keywords: Elastic optical networks (EONs) · Software-defined network (SDN) · Routing and spectrum assignment (RSA) · Horizontal fragmentation · Vertical fragmentation

1 Introduction

With the rapid growth of the diversified traffic demands, the traditional network is confronted with enormous challenges generated by carrying capacity. Elastic optical networks (EONs) [1] have emerged as a promising solution, which can maximize spectral utilization by employing fine granularity frequency slots (FSs). While the finer granularity of spectrum allocation is the key feature that makes EONs attractive, it also introduces new challenges. In a dynamic traffic scenario, the setting up and tearing down of the dynamic path will inevitably lead to spectral fragmentation which results in poor blocking probability (BP) and fragment degree (FD).

The dynamical and global resource allocation is essential for spectrum defragmentation in EONs. Thus, centralized control works more efficient than a distributed way. As is known to all, software-defined network (SDN) [2] has the ability to perceive and control the network in a global view by using a centralized controller. The authors in [3] explained that combining the SDN and EONs, software-defined

elastic optical networks (SD-EONs) provide an easier way to manage spectrum resource. SD-EONs architecture is shown in Fig. 1. Bandwidth variable wavelength selective switches (BV-WSSs) and edge routers (ERs) working in underlying data plane are interconnected by an optical fiber. The control plane consists of a centralized controller and several OpenFlow Agents (OF-AGs) that communicate with each other by sending extended OpenFlow [4] messages.

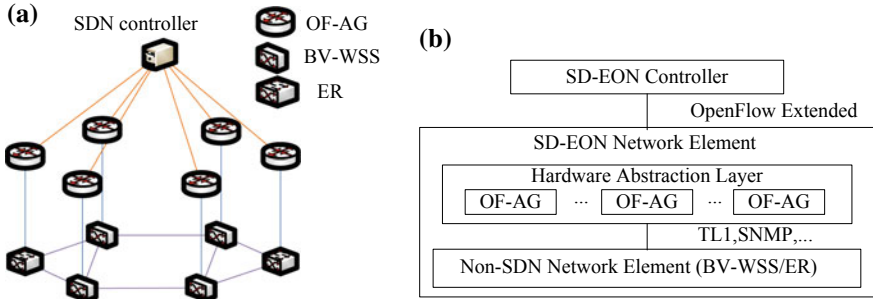


Fig. 1. **a** Network architecture of SD-EONs. **b** Detail of the network architecture of SD-EONs

The RSA problems can be divided into routing subproblem and spectrum assignment subproblem. Several algorithms have been raised to address the RSA problem [5–7]. Most previous investigations in [5, 6] use Shortest Path (SP) algorithms to solve the routing problem, and they usually just get candidate paths weighted by hops or distance. However, the link resources status is also very important actually. If we choose a candidate path that only depends on hops, the chosen path might not have enough available FSs for the requests. Therefore, it is better to consider both hops and link spectrum consecutiveness in routing selection.

In addition, most existing investigations have drawn little attention to the influence on BP and FD caused by fragmentations. Shortest path first-fit (SP-FF) algorithm [6] simply selects the first available frequency block (FB) by considering the time complexity of the algorithm. Nonetheless, out of the consideration of the fragmentations, SP-FF may disperse the available FBs into nonuniform small pieces and lead to poor performance in terms of BP and FD. The authors in [7] first analyzed the fragmentation problems in EONs in two dimensions and categorized it as the spectral fragmentation and spatial fragmentation, and proposed a two-dimensional fragmentation awareness (2D-FA) algorithm to reduce BP. However, they emphasize on the spectral fragmentation and only consider spatial fragmentation when there is more than one FB with minimal spectral fragmentation. Actually, fragmentations can be classified into horizontal fragmentation and vertical fragmentation.

In order to further improve the spectral utilization, we propose a hop and consecutiveness based routing and fragmentation-aware spectrum assignment (HCR-FSA) algorithm for SD-EONs. Unlike the works in [6, 7], we pay attention to both hops and link spectrum consecutiveness to select routes and take horizontal and vertical fragmentations into consideration in a more comprehensive way for spectrum assignments.

We also give the mathematical definition of these two types of fragmentations, which can globally represent the spectrum utilization status.

The rest of this paper is organized as follows. Section 2 analyses the RSA problems in SD-EONs. Our HCR-FSA algorithm is given in Sect. 3. Simulation results are analyzed in Sect. 4. Finally, we summarize the paper in Sect. 5.

2 Problem Analysis

RSA problem takes a significant part in SD-EONs, in which an available FB is required for each new arrival traffic request. We prefer our HCR-FSA algorithm serving the maximum number of traffic requests and alleviate BP and FD. In the routing problem, we aim to select several candidate paths between source–destination node pairs for each traffic request, and our objective is to minimize the hops of the transmission path and maximize the spectral consecutiveness.

In terms of hops, we desire to allocate the request on the shortest path, since we have to assign continuous spectrum among all links on the forwarding path and the shortest path can minimize the total amount of FSs. We take the topology in Fig. 2a for an example. Assume that there is a request r between source node v_3 and destination node v_6 and its bandwidth requirement is 2 FSs. Obviously, there are several candidate paths can be routed to request r , such as $v_3-v_5-v_6$, $v_3-v_2-v_4-v_6$, $v_3-v_5-v_4-v_6$, and $v_3-v_1-v_2-v_4-v_6$. Since the hops on those paths are 2, 3, 3, and 4, we have to allocate 4, 6, 6, and 8 FSs in total for request r respectively. Therefore, in order to reduce the total amount of occupied FSs, we prefer the paths with fewer hops. On the other hand, the overall spectrum consecutiveness on the forwarding path is also an important parameter during the selection of the candidate paths. The path with high spectrum consecutiveness can potentially transmit more traffic requests, which lead to lower BP.

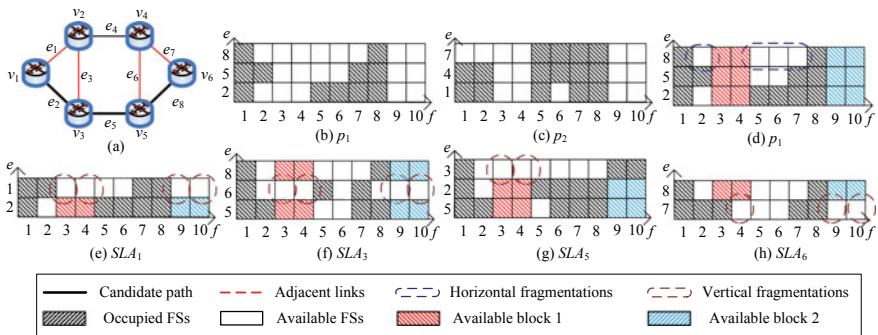


Fig. 2. **a** An example of network topology. **b, c** An example of spectral status on candidate paths p_1 and p_2 . **d** Horizontal fragmentations on candidate path p_1 . **e-h** Vertical fragmentations on adjacent links of $v_1, v_3, v_5,$ and v_6

We discuss the spectrum assignment problems in two dimensions, and describe two kinds of fragmentations namely as horizontal and vertical fragmentations. On one hand, since there is spectrum contiguity constraint in EONs, the dispersal of available spectral resources might generate useless spectrum fragmentations. We define horizontal fragmentation as those useless spectrum resources caused by irrational allocations, which dispersed available spectral slots into pieces. The vertical fragmentation, on the other hand, indicates those usable FSs that cannot assign to traffic requests, due to the FSs on the neighbor links are not available along the candidate paths, featured by spectrum continuity constraint in EONs.

3 HCR-FSA Algorithm Description

We define SD-EONs as a graph $G(V, E)$, in which $V = \{v_1, v_2, \dots, v_n\}$ represents the set of network nodes and $E = \{e_1, e_2, \dots, e_l\}$ represents the set of bidirectional edges between nodes in V . $|V| = N$, $|E| = L$ denotes the numbers of network nodes and links, respectively.

To better describe our algorithm, we introduce the following notations.

$r(v_s, v_d, b)$: The traffic request r , where v_s and v_d are the source and destination nodes, b represents the request bandwidth.

F_s : The number of FSs on each link.

P_r, Hop_p : The sets of candidate lightpaths p for request r , and the hops of p .

U_F^e : The sets of u_f^e , where u_f^e describe the status of the FS on link e , taking 0 if the FS f on link e is occupied, otherwise taking 1.

B_F^e : The number of available FB on link e , formulated as Eq. (1).

C_F^e : The consecutiveness of link e indicates the level of the spectrum utilization status on each link, formulated as Eq. (2). The first multiplier in denominator represents the number of continuous available FSs, and the other multiplier in denominator indicates the total number of available FSs.

C_p : The average consecutiveness of links on the candidate path p , given in Eq. (3).

SLT_p, SN_p : The sets of links and nodes in the candidate lightpath p , respectively.

SLA_v : The sets of all adjacent links for each node v in SN_p .

HF_b : The degree of the horizontal fragmentations on links in set SLT_p , formulated as Eq. (4). Especially, we define $u_{f-1}^e = 1$ and $u_{f_{FS+1}}^e = 1$ where f_b is the boundaries ($start = 1$ or $end = F_s$), since it will not introduce new fragmentations.

AF_{f_i}, OF_{f_i} : The number of available FSs and occupied FSs on slot f_i among the adjacent links, defined as Eqs. (5–6).

VE_b : The vertical efficiency of FB b . It indicates the level of the vertical fragmentations on adjacent links of node v in SLA_n , for each node v belongs in set SN_p , given in Eq. (7).

$$B_F^e = \sum_{i=1}^{F_s} u_{f_i}^e - \sum_{i=1}^{F_s-1} u_{f_i}^e \cdot u_{f_{i+1}}^e \quad (1)$$

$$C_F^e = \sum_{i=1}^{F_s-1} u_{f_i}^e \cdot u_{f_{i+1}}^e \times \sum_{i=1}^{F_s} u_{f_i}^e / B_F^e \times F_s \quad (2)$$

$$C_p = \sum_{e \in SLT_p} \frac{C_F^e}{|SLT_p|} \quad (3)$$

$$HF_b = \sum_{e \in SLT_p} u_{f_{start-1}}^e \cdot u_{f_{end+1}}^e / |SLT_p| \quad (4)$$

$$AF_{f_i} = \sum_{v \in SN_p} \sum_{e \in SLA_v} u_{f_i}^e \quad (5)$$

$$OF_{f_i} = \sum_{v \in SN_p} \left(|SLA_v| - \sum_{e \in SLA_v} u_{f_i}^e \right) \quad (6)$$

$$VE_b = \sum_{i=start}^{end} |AF_{f_i} - OF_{f_i}| / \sum_{v \in SN_p} |SLA_v| \cdot b \quad (7)$$

In SD-EONs, we can easily get the status of network resources by using a centralized controller. To support SD-EONs, we have made an extension to both Packet-in and Flow-mod messages. The main procedure of our HCR-FSA algorithm is shown in Table 1. Firstly, when a traffic flow arrives at the source ER, it starts by performing a table lookup in the flow table. If the flow can be matched, ER should implement the flow-matching actions. Otherwise, the OF-AG connecting with the source ER generates an extended Packet-in message, which contains the source and destination nodes and the bandwidth of the traffic requests. Then, managing the network resources, OF-controller obtains the network status in a global view, and after a Packet-in message is received, the controller calculates candidate lightpaths based on the spectrum resources status. Here, we use KSP algorithm [8] to get k candidate lightpaths, and choose the best path with the biggest $weight_p$ formulated as Eq. (8).

$$weight_p = C_p / Hop_p \quad (8)$$

$$cost_b = \alpha \cdot HF_b + (1 - \alpha) \cdot (1 - VE_b) \quad (9)$$

After the routing selection, the controller calculates the optimal FB with lowest $cost_b$ for each request, defined as Eq. (9). Then, the controller informs the optimal path and spectrum resources by sending Flow-mod messages to the OF-AGs, and those OF-AGs along the transmission path insert the optical flow entries. Finally, if there are available spectrum resources for the request, then r could be transmitted through the established lightpath between source node and destination node. According to the pseudocode given in Table 1, the worst time complexity of our HCR-FSA algorithm is approximately $O(KLF)$.

Taking Fig. 2 for an example, we assume that $k = 2$, and there is a request $r(v_1, v_6, 2)$. It is easy to see that the transmission paths between n_1 and n_6 are $SLT_{p_1} = \{e_2, e_5, e_8\}$ and $SLT_{p_2} = \{e_1, e_4, e_7\}$. In addition, we can get the consecutiveness of p_1 and p_2 , according to Eq. (3), i.e., $C_{p_1} = 1.45$ and $C_{p_2} \approx 0.644$. Since $Hop_{p_1} = Hop_{p_2} = 3$ and $weight_{p_1} > weight_{p_2}$, the optimal transmission path between v_1 and v_6 is p_1 . According to the resources utilization shown in Fig. 2d–h, there are two available FBs for request r in p_1 , b_1 (colored in red) and b_2 (colored in blue). We can calculate $HF_{b_1} \approx 0.333$, $HF_{b_2} = 0$, $VF_{b_1} = 0.3$ and $VF_{b_2} = 0.4$. According to Eq. (9), $cost_{b_1} > cost_{b_2}$. Thus b_2 is the optimal available FB for request r .

Table 1. HCR-FSA algorithm

1.	Calculate the k shortest paths for each source and destination nodes in $G(V, E)$;
2.	When traffic request $r(v_s, v_d, b)$ arrives:
3.	Pick up the k shortest paths for r and add them to the set P_r ;
4.	If $P_r = \text{null}$:
5.	block request r ;
6.	Else :
7.	For path p in P_r do :
8.	Get the best transmission path p with maximum $weight_p$;
9.	Search the available FBs ($FB \geq b$) and storage them into set B_F ;
10.	If $B_F = \text{null}$:
11.	block request r ;
12.	Else :
13.	For FB in B_F do :
14.	Get the optimal block B with minimum $cost_b$;
15.	End For
16.	End For
17.	Allocate the resources B for request r .

4 Experimental Results

In this section, to evaluate the performance of our RSA algorithm, we set up a SD-EONs testbed, in which we use the RYU controller to implement our RSA algorithm, and use Mininet simulating OF-AGs. Each OF-AG communicates with the RYU by using our extended OpenFlow protocol. We built our experimental topology based on the NSTNET topology in Mininet. The network topology shown in Fig. 3a is discovered by RYU controller through OpenFlow messages which contain Link Layer Discovery Protocol (LLDP) information. In our RYU controller, when it receives extended Packet-in messages from OF-AGs, it performs our HCR-FSA algorithm. A capture of our RSA result has been given in Fig. 3b. After routing selection and spectrum assignment it sends extended Flow-mod messages to the OF-AGs among the transmission path.

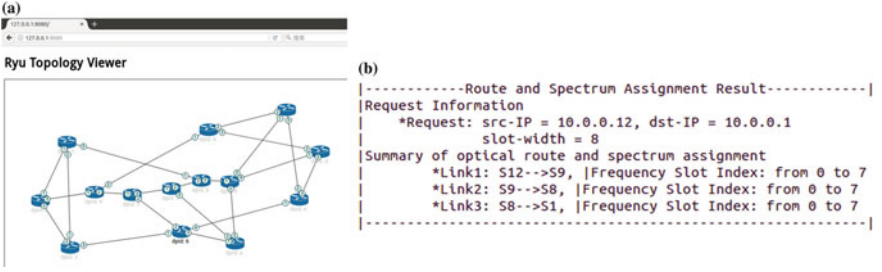


Fig. 3. **a** NSFNET topology shown in RYU controller. **b** A capture of HCR-FSA result

Our experiment is based on a dynamic network scenario, where the source and destination node pairs of the requests are randomly generated between 14 nodes in NSTNET. The traffic request model is randomly according to a Poisson process, and the holding time of each request follows a negative exponential distribution. In addition, in our experiment, we assume each fiber link has 400 spectral slots, and the traffic bandwidth is randomly distributed from 1 to 15 FSs.

We conducted the simulations of our RSA algorithm, and compare it with two traditional algorithms, namely, SP-FF algorithm in [6] and 2D-FA algorithm in [7] in terms of BP and FD. Simulation results are shown in Fig. 4.

Figure 4a compares the BP of three algorithms in our dynamic network scenario. It indicates that the BP of all algorithms rises with the increase of traffic loads. Moreover, three curves have nearly no difference in the traffic loads of 100 and 200 Erlang. However, the BP of SP-FF increases significantly when the traffic loads higher than 200 Erlang. The main reason is that the SP-FF algorithm simply assigns the available FSs with minimum index, which might result in spectrum fragmentations to a large extent. The BP of 2D-FA is better than SP-FF, since 2D-FA algorithm aims to avoid fragmentation during spectrum assignment. Considering the network spectrum status both in routing selection and spectrum assignment selection, our HCR-FSA algorithm can provide better solutions, which reduce BP significantly comparing to both SP-FF algorithm and 2D-FA algorithm.

$$FD = \sum_{r \in R} \sum_{e \in E} \left(\sum_{i=1}^{F_s-1} \left(\frac{u_{f_i}^e \wedge u_{f_{i+1}}^e + 1}{2} \right) \right) / \sum_{j=1}^{F_s} u_{f_j}^e \times |R| \times L \quad (10)$$

Figure 4b shows the fragment degree versus traffic load for three algorithms. The FD can be calculated according to Eq. (10), where the denominator represents the number of the available FB, and the first multiplier in the numerator is the number of the available FSs on link e . In terms of we prefer the smaller one, which means less fragmentations. The figure shows that our HCR-FSA has the best performance, followed by 2D-FA algorithm and SP-FF algorithm in sequence. The reason is that our HCR-FSA aims to avoid fragmentation during both routing selection and spectrum assignment. Whereas 2D-FA algorithm only considers reducing the fragmentation during the spectrum assignment and the SP-FF does not consider the fragmentation at all.

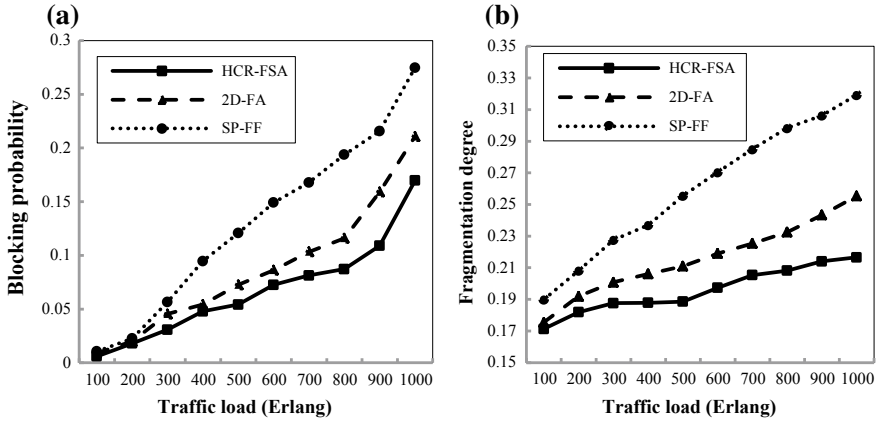


Fig. 4. Performance comparison of different algorithms in terms of **a** blocking probability and **b** fragment degree

5 Conclusions

In this paper, we proposed a novel HCR-FSA algorithm that considers both hops and spectrum consecutiveness in routing selection and pays attention to both horizontal and vertical fragmentations during spectrum assignment. We implement our RSA algorithm in an SD-EONs testbed, and the simulation results show that the HCR-FSA algorithm can achieve better blocking probability and fragment degree.

Acknowledgements. This work was supported in part by the National Natural Science Foundation of China under Grant 61371091, the “13th five-year” Key Research Project of DMU under Grant 3132016318, and the Fundamental Research Funds for Central Universities under Grant 3132017078.

References

1. Zhang, G., Leenheer, M.D., Morea, A., Mukherjee, B.: A survey on OFDM-based elastic core optical networking. *IEEE Commun. Surv. Tutor.* **15**(1), 65–87 (2013)
2. Nunes, B.A.A., Mendonca, M., Nguyen, X.N., Obraczka, K., Turletti, T.: A survey of software-defined networking: Past, present, and future of programmable networks. *IEEE Commun. Surv. Tutor.* **16**(3), 1617–1634 (2014)
3. Zhu, Z., Chen, X., Chen, C., Ma, S., Zhang, M., Liu, L., Yoo, S.J.B.: Openflow-assisted online defragmentation in single-/multi-domain software-defined elastic optical networks [invited]. *IEEE/OSA J. Opt. Commun. Netw.* **7**(1), A7–A15 (2015)
4. Mckeown, N., Anderson, T., Balakrishnan, H., Parulkar, G., Peterson, L., Rexford, J., Shenker, S., Turner, J.: Openflow: enabling innovation in campus networks. *ACM SIGCOMM Comput. Commun. Rev.* **38**(2), 69–74 (2008)

5. Wang, X., Zhang, Q., Kim, I., Palacharla, P.: Utilization entropy for assessing resource fragmentation in optical networks. In: Optical Fiber Communication Conference and Exposition, pp. 1–3 (2012)
6. Wan, X., Wang, L., Hua, N., Zhang, H., Zheng, X.: Dynamic routing and spectrum assignment in flexible optical path networks. In: Optical Fiber Communication Conference and Exposition, pp. 1–3 (2011)
7. Zhang, H., Xia, M., Zhang, M., Yoo, S.J.B., Dahlfort, S., Yin, Y., Zhu, Z.: Spectral and spatial 2D fragmentation-aware routing and spectrum assignment algorithms in elastic optical networks [invited]. *IEEE/OSA J. Opt. Commun. Netw.* **5**(10), A100–A106 (2013)
8. Yen, J.Y.: Finding the k shortest loopless paths in a network. *Manag. Sci.* **17**(11), 712–716 (1971)



Numerical Implementation of a Wideband Chaotic Light Based Ring-and-Spur Long-Reach Passive Optical Network: Architectures and Real-Time Secure Communications

Xinyu Dou^{1,2}, Hongxi Yin²(✉), and Bin Wu²

¹ School of Information Science and Technology, Dalian Maritime University, Dalian 116026, Liaoning, China

² Lab of Optical Communications and Photonic Technology, School of Information and Communication Engineering, Dalian University of Technology, Dalian 116023, Liaoning, China
hxyin@dlut.edu.cn

Abstract. In this paper, a wideband chaotic light based ring-and-spur long-reach passive optical network (ring-and-spur LR-PON) is proposed and numerically demonstrated. The bidirectional long-haul secure communication between the optical line terminal (OLT) and the optical network unit (ONU) is realized, in which the bandwidth of the chaos carrier is more than 20 GHz. As a result, the data rate can reach up to 10 Gb/s for each wavelength channel. The proposed network has potential practical values, owing to its advantages such as providing a higher bit rate, achieving a higher level of real-time security, survivability, etc.

Keywords: Fiber optics communications · Chaos · Ring-and-spur LR-PON · Real-time security

1 Introduction

Next-generation passive optical network (NG-PON) is a promising technique to solve the growing demand for bandwidth in the access area [1–3]. Additionally, the security of network communication is a crucial issue. Due to simple realization, real-time encryption, and compatibility with the existing optical communication system, the chaotic light based secure communication has been widely investigated [4, 5]. Up to now, the chaos-based secure communication in PON with the tree-and-branch topology has been reported [6]. However, the span of the tree-and-branch topology is only 20 km [7]. Thus, there has been a growing interest in the ring-and-spur long-reach passive optical network (ring-and-spur LR-PON) recently. The ring topology can extend the coverage to more than 60 km, and can provide a better anti-fault property, because the direction of the traffic signal can be flexibly changed to avoid the fiber fault

[8]. Meanwhile, the wavelength-division multiplexing (WDM) technique is exploited in the ring-and-spur PON, and the bit rate for each wavelength channel can reach up to 10 Gb/s [9]. To the best of our knowledge, the chaotic light based secure communication in the ring-and-spur LR-PON is rarely investigated.

In this paper, a wideband chaotic light based ring-and-spur LR-PON is proposed. The bidirectional long-haul secure communication between the OLT and the ONU is numerically demonstrated. The rest of this paper is arranged as follows. The architecture of the ring-and-spur LR-PON is shown in Sect. 2. In Sect. 3, the performance of the bidirectional chaos-based secure communication system, which is suitable for the proposed ring-and-spur LR-PON, is numerically investigated. Finally, a conclusion is presented in Sect. 4.

2 Architecture of the Ring-and-Spur LR-PON

The scheme of the ring-and-spur LR-PON is shown in Fig. 1. Take three remote nodes (RN) as an example to illustrate the structure of the network. The OLT and the three RNs constitute the ring topology. Each RN connects to several ONUs. For wideband chaotic light based secure communication, two fiber links in the inner ring (black in the picture) are deployed. The other two fiber links (red in the picture) in the outer ring are the back up of those in the inner ring. They are activated when a fiber fault occurs in the inner ring. The WDM technique is exploited in the ring. A specific wavelength (or waveband) is dropped and added at each RN, e.g., λ_1 at RN-1, λ_2 at RN-2, and so on. Therefore, the OLT must contain the transmitters for all wavelengths, and as a result, wavelength-division multiplexers and de-multiplexers are deployed in each RN, as well as the power amplification and the dispersion compensation devices to restore the distortion induced

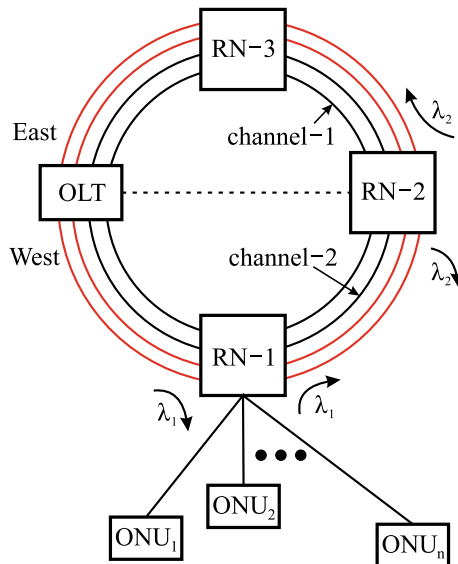


Fig. 1. Architecture of the ring-and-spur LR-PON

by the fiber link. Each ONU contains a pair of chaotic receiver and transmitter for secure communication. The network is divided into two parts. The upper part is defined as “East”, while the lower part is “West”. Normally, both the downstream and upstream data are transmitted in the counterclockwise direction (from West to East). When the protection link is activated, the data is transmitted in the clockwise direction.

3 Wideband Chaotic Light Based Secure Communication System Suitable for the Ring-and-Spur LR-PON

3.1 Architecture of the Wideband Chaotic Light Based Real-Time Secure Communication System

The wideband bidirectional chaotic light based secure communication system, which is suitable for the ring-and-spur LR-PON is shown in Fig. 2. The system enables broadcasting in the downlink and multichannel reception in the uplink. The master laser M in the OLT is subjected to optical feedback and mutually coupled with the slave laser S. Thus, both lasers are able to emit the wideband chaotic light. The downstream data m_1 is hidden in the chaotic light emitted by the slave laser. After transmission in the fiber link, the two chaotic signals are launched into the chaos receiver at ONU. First, both the two chaotic signals are divided into two parts by optical couplers, respectively. One part of the chaotic signal in channel-1 is injected to the laser S'-1, which enables the synchronization between the laser S in the OLT and the S'-1 in the ONU. Then, the output of S'-1 is also divided into two parts. One path is used to recover the downstream data m_1' by subtracting the chaotic signal from channel-2 and filtering, while the other path is used to mask the upstream data m_2 and transmit backward to the OLT through channel-1. The structure of S'-2 is symmetric to that of S'-1, while the upstream chaos signal (chaos + m_3) is launched into the channel-2. In the OLT, both the upstream data

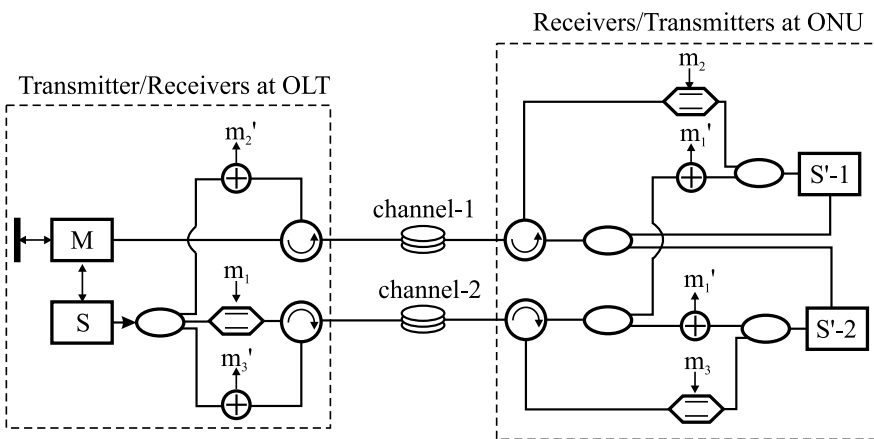


Fig. 2. Scheme of the wideband bidirectional chaotic light based secure communication system

m_2' and m_3' are recovered by subtraction between the backward chaos signals and the output of laser S. Therefore, the bidirectional secure communication is completed.

3.2 Performance of the Real-Time Secure Communication

The dynamic behavior of the transmitters and receivers can be described by the Lang–Kobayashi rate equation as follows:

$$\begin{aligned}
 \frac{dE_M(t)}{dt} &= \frac{1}{2}(1+i\alpha) \left[\frac{G[N_1(t) - N_0]}{1 + \varepsilon|E_M(t)|^2} - \frac{1}{\tau_P} \right] E_M(t) + \kappa_r E_M(t - \tau) \exp(-i\omega\tau) + \kappa_c E_S(t), \\
 \frac{dN_M(t)}{dt} &= \frac{I_M}{qV} - \frac{1}{\tau_n} N_M(t) - \frac{G[N_M(t) - N_0]}{1 + \varepsilon|E_M(t)|^2} |E_M(t)|^2, \\
 \frac{dE_S(t)}{dt} &= \frac{1}{2}(1+i\alpha) \left[\frac{G[N_S(t) - N_0]}{1 + \varepsilon|E_S(t)|^2} - \frac{1}{\tau_P} \right] E_S(t) + \kappa_c E_M(t), \\
 \frac{dN_S(t)}{dt} &= \frac{I_S}{qV} - \frac{1}{\tau_n} N_S(t) - \frac{G[N_S(t) - N_0]}{1 + \varepsilon|E_S(t)|^2} |E_S(t)|^2,
 \end{aligned} \tag{3.1}$$

$$\begin{aligned}
 \frac{dE_{S'-1}(t)}{dt} &= \frac{1}{2}(1+i\alpha) \left[\frac{G[N_{S'-1}(t) - N_0]}{1 + \varepsilon|E_{S'-1}(t)|^2} - \frac{1}{\tau_P} \right] E_{S'-1}(t) + \kappa_c E_S(t), \\
 \frac{dN_{S'-1}(t)}{dt} &= \frac{I_{S'-1}}{qV} - \frac{1}{\tau_n} N_{S'-1}(t) - \frac{G[N_{S'-1}(t) - N_0]}{1 + \varepsilon|E_{S'-1}(t)|^2} |E_{S'-1}(t)|^2, \\
 \frac{dE_{S'-2}(t)}{dt} &= \frac{1}{2}(1+i\alpha) \left[\frac{G[N_{S'-2}(t) - N_0]}{1 + \varepsilon|E_{S'-2}(t)|^2} - \frac{1}{\tau_P} \right] E_{S'-2}(t) + \kappa_c E_S(t), \\
 \frac{dN_{S'-2}(t)}{dt} &= \frac{I_{S'-2}}{qV} - \frac{1}{\tau_n} N_{S'-2}(t) - \frac{G[N_{S'-2}(t) - N_0]}{1 + \varepsilon|E_{S'-2}(t)|^2} |E_{S'-2}(t)|^2,
 \end{aligned} \tag{3.2}$$

where the $E_M(t)$, $E_S(t)$, $E_{S'-1}(t)$, and $E_{S'-2}(t)$ represent the slowly varying complex electronic field amplitude of the master laser and the slave laser in the OLT and that of the laser S'-1 and S'-2 in the ONU, respectively. Other parameters are as the same as those in Ref. [10].

The fiber link is described by the nonlinear Schrodinger equation, and the parameters are as same as those in Ref. [11]. Assuming the length of the fiber link to be 20 km, the simulation results based on the theoretical models are shown in Fig. 3.

As can be seen from Fig. 3, the bandwidth of the carrier is 26.2 GHz, which enables the carrier to mask the 10 Gb/s data sequence. The calculation of the correlation coefficient (CC) is similar to Ref. [11]. Then, the CC of forward and backward transmission is 0.97 and 0.94, respectively. This indicates that the secure communication for both directions can be achieved. However, the quality of backward synchronization is worse than that of the forward transmission. This is because the incomplete synchronization induced by the forward transmission is overlaid on the

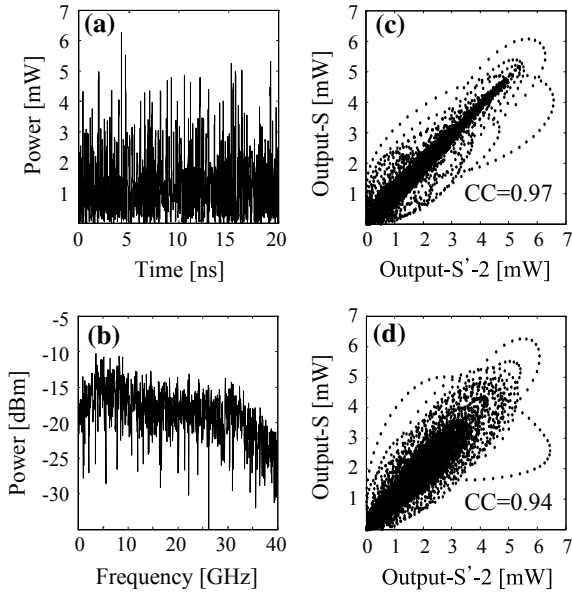


Fig. 3. Performance of the chaos generation and chaos correlation. **a** Waveform of the chaotic carrier, **b** RF spectrum of the chaotic carrier, **c** correlation plot of the forward transmission, **d** correlation plot of the backward transmission

backward signal. Hence, after the transmission in the fiber link, the backward signal is distorted more severely. As shown in Fig. 4, this phenomenon becomes more apparent when the distance increases. When the distance is more than 60 km, the CC of both directions falls rapidly. Especially, the CC of the backward transmission is lower than 0.7, which means the synchronization is failed. Therefore, dispersion compensation fiber (DCF) is required for long-haul bidirectional transportation. As can be seen in Fig. 4, high-quality synchronization (0.99) can be obtained after DC. For instance, the length of the DCF is 2.1 km when the distance is nearly 80 km.

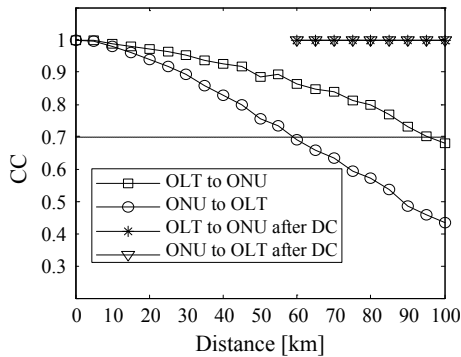


Fig. 4. Relationship between the CC and the transmission distance before and after dispersion compensation

Based on the high-quality chaos synchronization, secure communication can be realized. Let the distance between each RN be 20 km, and the downstream data of each wavelength channel to be 10 Gb/s, the downlink communication results of arbitrary ONU of each wavelength channel is shown in Fig. 5. During the simulation, the wavelengths of the three channels are 1554.94 nm (λ_1), 1553.33 nm (λ_2), and 1551.72 nm (λ_3), respectively. All ONUs can successfully recover the message. The calculated Q-factors are all above 3.5. According to Ref. [11], the BER can be 10^{-13} after FEC, which satisfies the requirement of the communication system. Thanks to the dispersion compensation, the performance of the transmission of the upstream data is nearly the same as the downstream scenario.

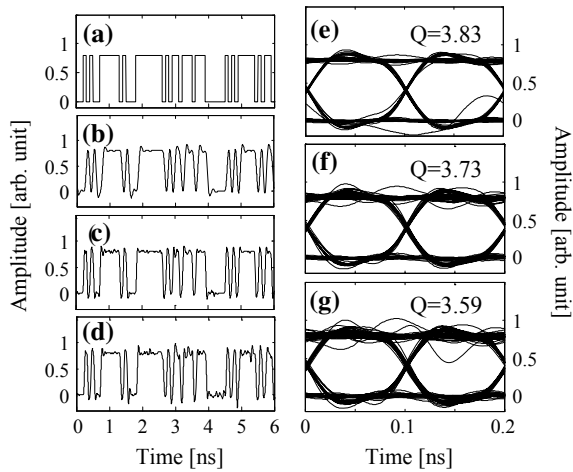


Fig. 5. Communication results of the downstream data. **a** The original data sequence, **b–d** the recovered data sequence of arbitrary ONU of channel λ_1 , λ_2 , and λ_3 , respectively, **e–f** the corresponding eye diagrams

4 Conclusion

The wideband chaotic light based ring-and-spur LR-PON is numerically demonstrated in this paper. The architecture of the network and its key element, such as the OLT and the RN, as well as the bidirectional wideband chaos-based secure communication system are designed. In addition, real-time online fiber fault detection and location are discussed. The bandwidth of the chaotic carrier is more than 20 GHz, which enables the carrier to mask the 10 Gb/s data sequence. Through dispersion compensation, the span of the network is extended to 60 km, which satisfies the requirement of the bit rate and the coverage of the NG-PON. Thus, the proposed network has potential practical values, due to the higher communication capacity, higher security, and the lower cost of operation, administration, and maintenance.

Acknowledgements. This work is supported in part by the National Natural Science Foundation of China (NSFC) under Grants 61071123 and 61172059, and the Fundamental Research Funds for Central Universities (Grant No. 3132018181, 3132016318 and 3132017018).

References

1. Wong, E.: Next-generation broadband access networks and technologies. *J. Lightwave Technol.* **30**(4), 597–608 (2012)
2. Esmail, M.A., Fathallah, H.: Physical layer monitoring techniques for TDM-passive optical networks: a survey. *IEEE Commun. Surv. Tut.* **15**(2), 943–958 (second quarter 2013)
3. Kantarci, B., Mouftah, H.T.: Bandwidth distribution solutions for performance enhancement in long-reach passive optical networks. *IEEE Commun. Surv. Tut.* **14**(3), 714–733 (third quarter 2012)
4. Grgyris, A., Synridis, D., Larger, L., Annovazzi-Lodi, V., Colet, P., Fischer, I., Garcia-Ojalvo, J., Mirasso, C.R., Pesquera, L., Shore, K.A.: Chaos-based communications at high bit rates using commercial fibre-optic links. *Nature* **438**(17), 343–346 (2005)
5. Argyris, A., Grivas, E., Bogris, A., Syvridis, D.: Transmission effects in wavelength division multiplexed chaotic optical communication systems. *J. Lightwave Technol.* **28**(21), 3107–3114 (2010)
6. Jiang, N., Liu, D., Zhang, C., Qiu, K.: Modeling and simulation of chaos-based security-enhanced WDM-PON. *IEEE Photon. Technol. Lett.* **25**(19), 1912–1915 (2013)
7. Rad, M.M., Fluli, K., Fathallah, H.A., Rusch, L.A., Maier, M.: Passive optical network monitoring: challenges and requirements. *IEEE Commun. Mag.* **49**(2), s45–s52 (2011)
8. Song, H., Kim, B.W., Mukherjee, B.: Long-reach optical access networks: a survey of research challenges, demonstrations, and bandwidth assignment mechanisms. *IEEE Commun. Surv. Tut.* **12**(1), 112–123 (first quarter 2010)
9. Chen, X., Zhang, Z., Hu, X.: The evolution trends of PON and key techniques for NG-PON. In: International Conference on Information and Signal Processing, Dec 2013
10. Dou, X., Wu, C., Chen, X., Yin, H., Zhao, Q., Hao, Y., Zhao, N.: Broadband chaotic light transmitter. *Chinese Optics Lett.* **12**(suppl. 1), s10610-1–s10610-4 (2014)
11. Zhao, Q., Yin, H.: Performance analysis of dense wavelength division multiplexing secure communications with multiple chaotic optical channels. *Opt. Commun.* **285**, 693–698 (2012)



A Theoretical Method for Constructing the Boundary of the Color Domain of the System

Sile Liu^(✉), Yan Li, Yali Huang, and Dan Zhen

Electronic and Communication Engineering College, Tianjin Normal University,
Tianjin, China
1400399217@qq.com

Abstract. This paper proposes a method to determine the color gamut boundary of the color system by theoretical calculation based on the colorimetric parameters of three primary colors and the reference color white. Gamut mapping is the fundamental way to improve the chromatic aberration between different color systems, and the determination of gamut boundaries is the prerequisite to achieve gamut mapping. In this paper, the frame of a three-dimensional color gamut of the system is established by using the chromaticity parameters of the three basic color of the system and the base white, and then the theoretical construction of the three-dimensional color gamut of the system is realized by the interpolation method. In this paper, the 3D gamut boundaries of the conventional gamut and the Rec.2020 WCG gamut are completed by using this method. The calculation results show that the construction method in this paper has a small amount of computation, short time required, high accuracy, and universality.

Keywords: Display technology · 3D color field boundary · Linear interpolation fitting

1 Introduction

Different color systems have different ranges of color gamut, so the color difference can be produced when the same color is displayed in different systems. A reasonable gamut mapping scheme is one of the solutions, and the determination of the color domain boundaries of two color systems to be mapped is the basis for color gamut mapping. At present, the existing color gamut boundary definition methods are mostly based on the image gamut boundary description. The convex hull algorithm, such as Kress and Stevens [2], describes the color gamut, which is only an approximate algorithm. All data points are treated as convex packets. The results are larger than the actual color gamut, and the color gamut boundary cannot be described accurately, and Cholewo and Love [3] proposed an algorithm to build the color gamut boundary based on alpha-shape, and build a relatively accurate gamut boundary by controlling the triangulation of the color gamut, but because of the introduction of the parameters, it affects the

accuracy. Then, Morovic and Luo [4, 5] put forward the partition maximization algorithm. By partitioning the color space first and using the points in the partition to construct the color domain surface, the real color gamut boundary points can be obtained by this method, but the construction of the color domain is restricted by the partition, and the accuracy can also be affected.

In this paper, the three-dimensional color gamut boundary of the system is calculated by testing the three basic color parameters of the system and the chromaticity parameters of the base white, and then the three-dimensional color gamut of the color system is constructed by the interpolation method.

2 Methods for Constructing 3D Color Domain

2.1 Establish the Color Space Conversion Matrix

Color space is divided into device-dependent color space and device-independent color space. The visual uniformity of the device-independent color space is more consistent with the human eye characteristics, so the gamut mapping algorithm should be implemented in the device independent color space. In this paper, the CIE 1976 LAB color space is chosen as the parameter computing space. In this section, the calculation matrix for the conversion of the conventional color gamut and the Rec.2020 color gamut to the CIE 1976 LAB color space will be established, respectively. Table 1 shows the colorimetric parameters of the conventional gamut and Rec.2020 color space. Figure 1 is the CIE 1931 color map of the conventional gamut and Rec.2020.

Table 1. RGB tricolor and reference whiteness parameters of Rec.709 and Rec.2020

Color space	White (D65)		Red		Green		Blue	
	x_C	y_C	x_1	y_1	x_2	y_2	x_3	y_3
Rec.709	0.3127	0.3290	0.640	0.330	0.300	0.600	0.150	0.060
Rec.2020	0.3127	0.3290	0.708	0.292	0.170	0.797	0.131	0.046

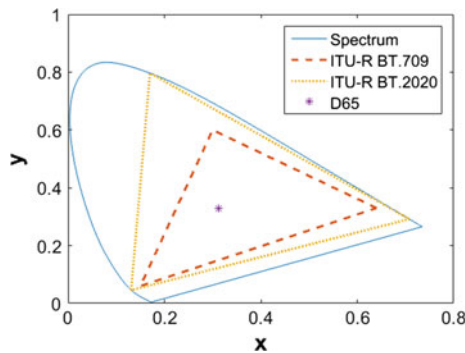


Fig. 1. CIE 1931 chromaticity map of Rec.2020 and conventional color fields

So you can see that

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4125 & 0.3576 & 0.1804 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9503 \end{bmatrix} \begin{bmatrix} R_e \\ G_e \\ B_e \end{bmatrix} \quad (1)$$

In Eq. (1) $\begin{bmatrix} 0.4125 & 0.3576 & 0.1804 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9503 \end{bmatrix}$ To REC.709 [6]. The equipment related to the color matrix.

According to formula (1), the relationship between X, Y, and Z of the three stimulus values and the unit quantity of the three base colors can be obtained:

$$\begin{cases} X = 0.4125R_e + 0.3576G_e + 0.1804B_e \\ Y = 0.2126R_e + 0.7152G_e + 0.0722B_e \\ Z = 0.0193R_e + 0.1192G_e + 0.9503B_e \end{cases} \quad (2)$$

Among them $Y = 0.2126R_e + 0.7152G_e + 0.0722B_e$ is the luminance equation in the color domain of Rec.709.

So you get Rec.2020 [7]. The matrix transformation of the tristimulus values X, Y, Z and trichromatic units is

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.6370 & 0.1446 & 0.1689 \\ 0.2627 & 0.6780 & 0.0593 \\ 0.0000 & 0.0281 & 1.0609 \end{bmatrix} \begin{bmatrix} R_e \\ G_e \\ B_e \end{bmatrix} \quad (3)$$

According to formula (3), the relationship between X, Y, Z and trichromatic unit quantities of the three stimulus values is

$$\begin{cases} X = 0.6370R_e + 0.1446G_e + 0.1689B_e \\ Y = 0.2627R_e + 0.6780G_e + 0.0593B_e \\ Z = 0.0000R_e + 0.0281G_e + 1.0609B_e \end{cases} \quad (4)$$

$Y = 0.2627R_e + 0.6780G_e + 0.0593B_e$ is the luminance equation of the Rec.2020 color gamut.

2.2 Framework for Determining Color Domain Boundaries

According to the luminance equation, the characteristic luminance plane of the display device can be established. Among the three base colors of Rec.709, the brightness of green is the highest, that of red is the second, and that of blue is the lowest, that is $Y_g > Y_r > Y_b$. And it satisfies $0 < Y_b < Y_r < Y_{b+r} < Y_g < Y_{g+b} < Y_{g+r} < Y_{g+r+b}$, so the brightness of Rec.709 can be divided into 7 brightness ranges $[0, Y_b]$, $[Y_b, Y_r]$, $[Y_r, Y_{b+r}]$, $[Y_{b+r}, Y_g]$, $[Y_g, Y_{g+b}]$, $[Y_{g+b}, Y_{g+r}]$, $[Y_{g+r}, Y_{g+r+b}]$, eight luminance planes can be obtained, which are $0, Y_b, Y_r, Y_{b+r}, Y_g, Y_{g+b}, Y_{g+r}, Y_{g+r+b}$. Through calculation, the boundary of each luminance plane can be obtained, and then the

resulting luminance plane boundary can be interpolated to obtain the three-dimensional color gamut boundaries of Rec.709 and rec.2020.

This paper takes the Rec.709 system as an example to illustrate. The luminance equation of Rec.709 is $Y = 0.2126R_e + 0.7152G_e + 0.0722B_e$, and the corresponding brightness plane is $Y = 0, Y = 0.0722, Y = 0.2126, Y = 0.2848, Y = 0.7152, Y = 0.7874, Y = 0.9278, Y = 1$, respectively.

The three basic colors of red, green, and blue can be mixed in a certain proportion by adding a mixed color principle to get any color in nature, so the sum brightness of the red, green and blue is equal to the brightness of the mixed color. First, select a luminance plane, then according to the luminance equation, the ratio of three basic colors of imaging to meet the requirements is determined. The proportion of the three primary colors of the imaging computed is substituted into the formula (2), the corresponding three stimulus value XYZ can be obtained, and then the three stimulus value XYZ is replaced in the formula (5), and the Lab values of the CIELAB color space is obtained. The three-dimensional gamut boundaries of REC.709 in CIELAB space can also be obtained.

$$\begin{cases} L^* = 116f\left(\frac{Y}{Y_n}\right) - 16 \\ a^* = 500 \left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right) \right] \\ b^* = 200 \left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right) \right] \end{cases} \quad (5)$$

In Eq. (5),

$$f\left(\frac{X}{X_n}\right) = \begin{cases} \sqrt[3]{X/X_n} & X/X_n > 216/24389 \\ 841/108(X/X_n) + \frac{16}{116} & X/X_n \leq 216/24389 \end{cases}$$

$$f\left(\frac{Y}{Y_n}\right) = \begin{cases} \sqrt[3]{Y/Y_n} & Y/Y_n > 216/24389 \\ 841/108(Y/Y_n) + \frac{16}{116} & Y/Y_n \leq 216/24389 \end{cases}$$

$$f\left(\frac{Z}{Z_n}\right) = \begin{cases} \sqrt[3]{Z/Z_n} & Z/Z_n > 216/24389 \\ 841/108(Z/Z_n) + \frac{16}{116} & Z/Z_n \leq 216/24389 \end{cases}$$

The color gamut boundary of each plane is determined by calculation on the eight luminance planes of Rec.709 selected in this paper. The display brightness required by the display device is set as $Y = C_r \times Y_r + C_g \times Y_g + C_b \times Y_b$. C_r, C_g, C_b represent the ratio of the three base color luminance required to display a certain luminance, and Y_r, Y_g, Y_b present the brightness of the three base colors of the image, respectively, that is, the coefficient in the luminance equation of Rec.709. Thus, the boundary point of the luminance plane $Y = 0.0722$ is obtained, as shown in Fig. 2.

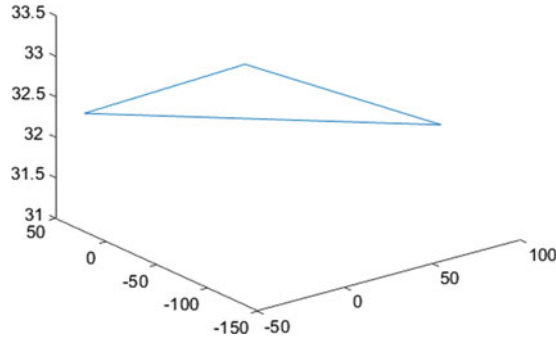


Fig. 2. $Y = 0.0722$ The color field boundary of the luminance plane

2.3 Interpolation of Color Domain Boundary

In this paper, according to these data, the three-dimensional color stereo boundary interface is fitted by linear interpolation in $L^*a^*b^*$ space. The interpolation interval is selected as 164,527 boundary points data are obtained by interpolating in Rec.709 gamut, and 113,949 boundary points data are obtained in Rec.2020 gamut, then these points are connected to form the three-dimensional color gamut boundaries of two color gamut. The result is shown in Fig. 3. In this paper, linear interpolation is used to fit the three-dimensional gamut boundary, which has the advantages of simple calculation and fast speed.

Finally, these boundary points are output in tabular form $I_{709} = [L_1^*, a_1^*, b_1^*]$, $I_{2020} = [L_2^*, a_2^*, b_2^*]$. A central point is selected in two 3D color stereos, respectively. In this paper, it is selected in Rec.709 3D color stereos $(L_1^*, a_1^*, b_1^*) = (50, 0, 0)$, also choose from Rec.2020 3D color stereo $(L_2^*, a_2^*, b_2^*) = (50, 0, 0)$, and then, according to the space position of each point, the hue angle θ_i and the space angle φ_i , $i = 1, 2$, of each point are determined according to the formulas (6) and (7), and the two angles can determine the only point on the boundary. The two tables can be expanded to $n \times 5$, $I'_{709} = [L_1^*, a_1^*, b_1^*, \theta_1, \varphi_1]$, $I'_{2020} = [L_2^*, a_2^*, b_2^*, \theta_2, \varphi_2]$, hue Angle θ_i And the spatial angles φ_i are accurate to 0.0001.

$$\theta_i = \frac{b_i}{a_i} \tag{6}$$

$$\varphi_i = \frac{\sqrt{(L_i - 50)^2 + (a_i - 0)^2 + (b_i - 0)^2}}{\sqrt{(a_i - 0)^2 + (b_i - 0)^2}} \tag{7}$$

3 Experimental Results

Through calculation, we get the three-dimensional gamut boundary data of Rec.709 and Rec.2020, as shown in Tables 2 and 3. The data are given in the uniform color space CIE 1976 LAB, which is convenient for color gamut mapping.

Table 2. LAB spatial data of eight luminance planes in the color domain of Rec.709

Surface brightness	Y = 0	Y = 0.0722			Y = 0.2126			
L*	0	32.30	32.30	32.30	53.23	53.23	53.23	53.23
a.*	0	55.93	40.13	79.15	80.15	57.51	92.10	88.63
b*	0	46.25	38.77	107.85	9.52	55.54	72.61	73.32
Surface brightness	Y = 0.2848				Y = 0.7152			
L*	60.32	60.32	60.32	60.32	87.74	87.74	87.74	87.74
a.*	98.26	60.34	63.40	6.37	3.52	26.54	86.18	42.99
b*	60.83	70.32	61.18	61.88	87.73	18.19	83.18	19.20
Surface brightness	Y = 0.7874				Y = 0.9278			Y = 1
L*	91.12	91.12	91.12	91.12	97.14	97.14	97.14	100.00
a.*	10.19	13.61	48.10	58.69	21.52	5.94	13.53	0.02
b*	90.12	12.54	14.12	87.28	94.48	4.19	4.52	0.01

Table 3 LAB spatial data of eight luminance planes in the color domain of Rec.2020

Surface brightness	Y = 0	Y = 0.0593			Y = 0.2627			
L*	0	29.24	29.24	29.24	58.29	58.29	58.29	58.29
a.*	0	71.45	76.33	86.13	117.34	125.66	123.02	16.85
b*	0	50.41	46.83	120.28	100.50	84.99	70.18	70.70
Surface brightness	Y = 0.322				Y = 0.678			
L*	63.51	63.51	63.51	63.51	85.91	85.91	85.91	85.91
a.*	97.75	130.55	134.42	33.44	17.73	48.42	172.33	98.51
b*	105.92	61.19	91.02	61.87	125.52	23.49	116.60	24.15
Surface brightness	Y = 0.7373				Y = 0.9407			Y = 1
L*	88.79	88.79	88.79	88.79	97.66	97.66	97.66	100.00
a.*	7.91	38.30	115.71	106.23	21.47	7.87	16.50	0.02
b*	128.18	18.66	121.58	19.33	136.86	3.89	4.04	0.01

Based on the linear interpolation method, two color gamut boundaries of the three systems are obtained, as shown in Fig. 3.

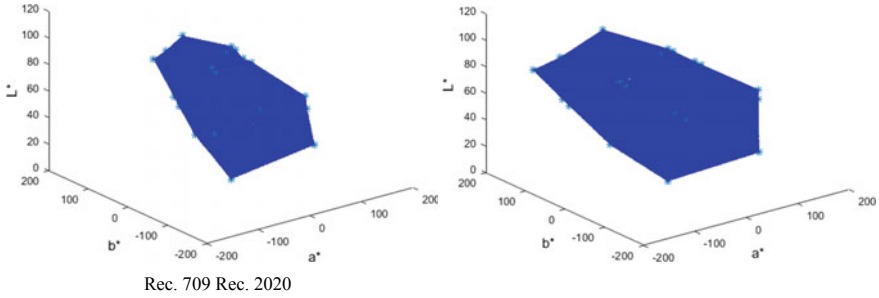


Fig. 3. 3D color domain of Rec.709 and Rec.2020

In this paper, a total of four sets of data are obtained by testing the chromaticity parameters of the three basic color and base white of the system. The three-dimensional color gamut framework of the system is constructed by theoretical analysis and calculation, and then the three-dimensional color gamut of the system is finally obtained by linear interpolation. This method avoids a large number of chromaticity test experiments and complex mathematical calculation, and has the advantages of shortcut, convenience, and universality.

Acknowledgements. This work was supported by Student's Platform for Innovation and Entrepreneurship Training Program (No: 201810065410).

References

1. Liu, Q.: Coverage of color domain and color domain. *Telev. Technol.* (11), 49–51 (2004)
2. Kress, W., Stevens, M.: Derivation of 3-dimensional gamut descriptors for graphic arts output devices. In: *Proceedings of TAGA, 1994*, p. 199
3. Cholewo, T.J., Love, S.: Gamut boundary determination using alpha-shapes. In: *Color and Imaging Conference, 1999*, pp. 200–204
4. Morovic, J., Luo, M.R.: Gamut mapping algorithms based on psychophysical experiments. In: *Color and Imaging Conference, 1997*, pp. 44–49
5. Morovic, J., Luo, M.R.: Calculating medium and image gamut boundaries for gamut mapping, *Color Res. Appl.* **252**, 394–401
6. Rec.709-4. Parameter values for the HDTV standards for production and international programme exchange (2000)
7. Rec.2020(08/2012). Parameter values for ultra-high definition television systems for production and international programme exchange (2012)
8. Ponomarenko, N., Lukin, V., Zelensky, A., Egiazarian, K., Carli, M.



Research and Evaluation of Color Gamut Extension Algorithm Based on Image

Yali Huang^(✉), Yan Li, Sile Liu, and Dan Zhen

Electronic and Communication Engineering College, Tianjin Normal University,
Tianjin, China
1400399217@qq.com

Abstract. In this paper, a new color gamut extension mapping algorithm based on the image is proposed by using the theoretical three-dimensional color domain boundary. The algorithm can be compatible with the existing conventional gamut video system and give full play to the reappearance advantage of WCG Rec.2020 display devices. In this paper, the mapping relationship between the two color gamut boundaries of Rec.709 and Rec.2020 is determined by the tone angle and the space position angle, and the extension coefficient lookup table of the conventional color gamut extends to the Rec.2020 WCG gamut is obtained. Then the color gamut expansion between the two system gamut is realized by the bilateral filtering and the lookup table assignment. In addition, this paper selects two images from the TID 2008 image library, and carries out data analysis and index evaluation on the color gamut mapping algorithm proposed in this paper. The results show that because the relative position between pixels of the source image is taken into account, the algorithm preserves the color features of the image better.

Keywords: Color gamut · Wide color gamut · Color gamut extension · Bilateral filtering

With the development of new technologies such as laser display [1], quantum dot technology [2] and so on, the fast development of wide gamut display technology has been promoted. Nowadays, it is very common that the color gamut of the source image is less than the wide color gamut system, so it is necessary to map the narrow color domain of the image to the wide color domain in some way, that is, the gamut expansion.

Kang et al. [3, 4] propose a multi-anchor gamut extension mapping algorithm. The algorithm uses brightness to determine the mapping direction for color expansion based on the equal hue. Liu [5] proposed a new color gamut boundary description method, realized the linear extension of the Rec.709 color gamut to the laser display color gamut (WCG). Jiang [6] use tetrahedron interpolation to propose a fast lookup table structure based on real-time applications to achieve gamut mapping, and complete the hardware implementation. Li [7] proposed a variable hue nonlinear gamut extension algorithm, which completed the gamut extension according to the color hue expansion function. Huang [8] proposed a color gamut expansion algorithm based on natural elasticity. The algorithm uses the tensile characteristics of nonuniform spring to determine the extended stretching parameters and realize the nonlinear expansion of the

color gamut. Han [9] proposed the variable hue skin color protection algorithm and the skin color protection algorithm based on the probability model, the former makes the extension mapping between the linear and the nonlinear by adjusting the angle. Before the latter is extended, the skin color area is judged, the non-skin part is extended normally, and the skin color is extended with a smaller coefficient, and the natural color is maintained well. In this paper, the gamut extension mapping of conventional gamut Rec.709 and wide gamut Rec.2020 is realized through bilateral filtering and lookup table assignment.

1 Bilateral Filtering

Bilateral filter is a nonlinear, noniterative, edge-preserving image smoothing filter. A definition domain (spatial domain) low-pass filter and a range (luminance domain) low-pass filter are combined together. It uses the intensity weighted mean value of pixels in certain proximity instead of the intensity of each pixel. This weight is based on the Gauss distribution, not only depending on the Euclidean distance of the pixels, but also on the pixel similarity, and is an improved algorithm based on the Gauss filter [10–12].

When the image is processed by Gauss filter, the pixel value of a point on the image is usually determined by the value of a small local neighborhood pixel around its location. That is, the pixel values in a certain neighborhood around the image are assigned to different Gauss weights, respectively, and the weighted average of the Gauss weight is assigned to this point. The Gauss weighting factor is generated by the spatial distance relation between two pixels, and its formula description is shown in Eqs. (1) and (2).

$$h(x) = k_d^{-1}(x) \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi)c(\xi - x)d\xi \quad (1)$$

$$k_d(x) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} c(\xi - x) d\xi \quad (2)$$

where c represents gaussian weight based on spatial distance and $k_d(x)$ is used to unit the result.

2 Color Gamut Extension Algorithm

Because the CIE $L^*a^*b^*$ color model (Lab) is based on the sense of human color and describes the way of color display, Lab is regarded as a device-independent color model. Therefore, we do color gamut extension mapping in Lab space. In this paper, we choose the image in the ITU-R BT.709 color gamut as the experimental image, and the image is extended to wide color gamut, and the image is presented in the ITU-R BT.2020 gamut. When the image is input, the RGB value of the pixel is converted to

the XYZ value according to the formula (3), and then the XYZ value is converted to $L^*a^*b^*$ according to the formula (4).

$$\begin{cases} X = 0.4125R_e + 0.3576G_e + 0.1804B_e \\ Y = 0.2126R_e + 0.7152G_e + 0.0722B_e \\ Z = 0.0193R_e + 0.1192G_e + 0.9503B_e \end{cases} \quad (3)$$

$$\begin{cases} L^* = 116f\left(\frac{Y}{Y_n}\right) - 16 \\ a^* = 500\left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right] \\ b^* = 200\left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right] \end{cases} \quad (4)$$

Then, according to the formulas (5) and (6), the hue angle θ and space angle φ of each pixel in the image is calculated, as exactly as 0.0001, the data form of the image is obtained, $I_{img} = [L_{img}, a_{img}, b_{img}, \theta_{img}, \varphi_{img}]$, and then the corresponding items are found in the ITU-R BT.709 color gamut and the ITU-R BT.2020 color domain.

$$\theta_i = \frac{b_i}{a_i} \quad (5)$$

$$\varphi_i = \frac{\sqrt{(L_i - 50)^2 + (a_i - 0)^2 + (b_i - 0)^2}}{\sqrt{(a_i - 0)^2 + (b_i - 0)^2}} \quad (6)$$

If ITU-R BT.709 color gamut and ITU-R BT.2020 the two color gamut boundary point hue angle θ and space angle φ are the same, the mathematical model can be determined according to the boundary points at the angle of the obtained ITU-R BT.709 and the ITU-R BT.2020 two gamut and the selected center point (50, 0, 0), as shown in Fig. 1.

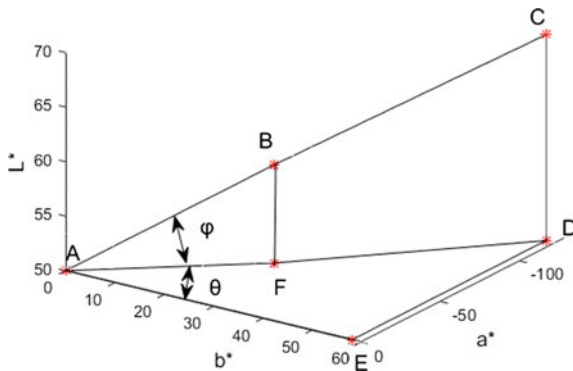


Fig. 1. Mathematical models of two color gamut boundary points and central points

The A point in the graph represents the center point of the selection (50, 0, 0), and the B point represents the boundary point of the ITU-R BT.709 gamut on this angle. The C point represents the boundary point of the ITU-R BT.2020 gamut on this angle, D is the projection of the C point in the a^*b^* plane, and the F is the projection of point B in the $a^* b^*$ plane, and the point E is the projection of point D on the b^* axis. In the color gamut extension mapping, the points B and C are, respectively, set as (L_{1j}, a_{1j}, b_{1j}) , (L_{1k}, a_{1k}, b_{1k}) B and C must satisfy $\theta_{img}(m) = \theta_{1j}$ and $\varphi_{img}(m) = \varphi_{1j}$ and $\theta_{img}(m) = \theta_{2k}$ and $\varphi_{img}(m) = \varphi_{2k}$, $m = 1, 2, 3, 4, 5, \dots, j = 1, 23, \dots, k = 1, 2, 3, \dots$ then the color gamut extension $(L'_{img}(m), a'_{img}(m), b'_{img}(m))$ can be obtained by formulas (7), (8) and (9).

$$AB = (L_{1j} - 50)^2 + (a_{1j} - 0)^2 + (b_{1j} - 0)^2 \tag{7}$$

$$AC = (L_{1k} - 50)^2 + (a_{1k} - 0)^2 + (b_{1k} - 0)^2 \tag{8}$$

$$(L'_{img}(m), a'_{img}(m), b'_{img}(m)) = \frac{\sqrt{AC}}{\sqrt{AB}} (L_{img}(m), a_{img}(m), b_{img}(m)) \tag{9}$$

If $\theta_{img}(m) = \theta_{1j}$ and $\varphi_{img}(m) = \varphi_{1j}$ and $\theta_{img}(m) = \theta_{2k}$ and $\varphi_{img}(m) = \varphi_{2k}$ is not satisfied, according to the formulas (10) and (11), respectively, in the ITU-R BT.709 color gamut and ITU-R BT.2020 color gamut boundary surface minimum error in the data table lookup and the target pixel points, the data selected by the two tables are expressed as $G(L_{1p}, a_{1p}, b_{1p}, \theta_{1p}, \varphi_{1p})$, and $H(L_{1q}, a_{1q}, b_{1q}, \theta_{1q}, \varphi_{1q})$, and then according to the formulas (12)–(14) to calculate the extended color gamut $(L'_{img}(m), a'_{img}(m), b'_{img}(m))$.

$$E = \sqrt{(\theta_{1j} - \theta_{image}(m))^2 + (\varphi_{1j} - \varphi_{image}(m))^2} \tag{10}$$

$$F = \sqrt{(\theta_{1k} - \theta_{image}(m))^2 + (\varphi_{1k} - \varphi_{image}(m))^2} \tag{11}$$

$$AG = (L_{1p} - 50)^2 + (a_{1p} - 0)^2 + (b_{1p} - 0)^2 \tag{12}$$

$$AH = (L_{1q} - 50)^2 + (a_{1q} - 0)^2 + (b_{1q} - 0)^2 \tag{13}$$

$$(L'_{img}(m), a'_{img}(m), b'_{img}(m)) = \frac{\sqrt{AH}}{\sqrt{AG}} (L_{img}(m), a_{img}(m), b_{img}(m)) \tag{14}$$

$$\begin{cases} f_y = (L + 16)/116 \\ f_x = \frac{a}{500} + f_y \\ f_z = f_y - \frac{b}{200} \end{cases} \tag{15}$$

$$X = \begin{cases} X_r \times f_x^3 f_x^3 > \epsilon \\ X_r \times \left(\frac{116f_x - 16}{k}\right) f_x^3 \leq \epsilon \end{cases} \quad (16)$$

$$Y = \begin{cases} Y_r \times f_y^3 & L > k\epsilon \\ Y_r \times \frac{L}{k} & L \leq k\epsilon \end{cases} \quad (17)$$

$$Z = \begin{cases} Z_r \times f_z^3 f_z^3 > \epsilon \\ Z_r \times \left(\frac{116f_z - 16}{k}\right) f_z^3 \leq \epsilon \end{cases} \quad (18)$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.7165 & -0.3556 & -0.2534 \\ -0.6666 & 1.6164 & 0.0158 \\ 0.0177 & -0.0428 & 0.9422 \end{bmatrix} \times \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (19)$$

After extending the color gamut, $(L'_{img}(m), a'_{img}(m), b'_{img}(m))$ is obtained. Then converted the data to the XYZ space by formulas (15)–(18), and then converted back to the RGB space according to the formula (19), and the image is reconstructed based on the RGB data to obtain the image after the extended mapping.

3 Experimental Results And Analysis

In this paper, two images from the color gamut map library TID 2008 were selected for the experiment, as shown in Fig. 2I1 and I2. The experimental environment was MATLAB R2015b.

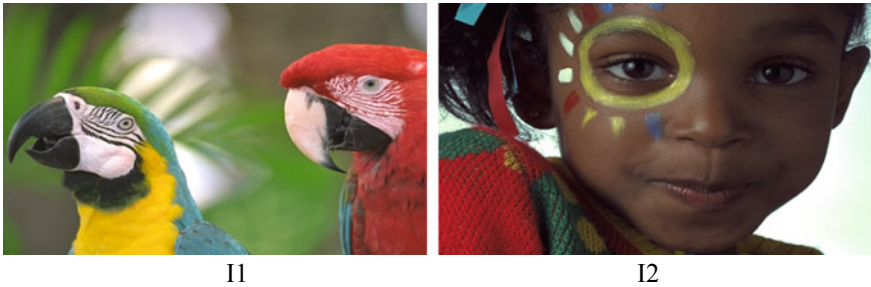


Fig. 2. Test images used in the experiment

The algorithm is used to extend the gamut of conventional gamut to Rec.2020 wide gamut of two test images, respectively. It is necessary to explain that because there is no Rec.2020 standard display to display the image after the mapping, the algorithm evaluation link in this paper cannot take the way of subjective experiment, and cannot display the image after the extended mapping.

However, the image's chromaticity diagram can restore the image information truthfully, and its composition, distribution, and boundary extension can be used to describe the expressive force of the image. Therefore, the source image (conventional gamut image) and the extended image (the image displayed on the Rec.2020 device) are projected to the CIE XY color plane, and their two-dimensional chromaticity distributions are obtained, respectively, as shown in Figs. 3 and 4. At the same time, this section also draws a comparison map of the 3D color distribution of the extended image through proper selection, as shown in Figs. 5 and 6.

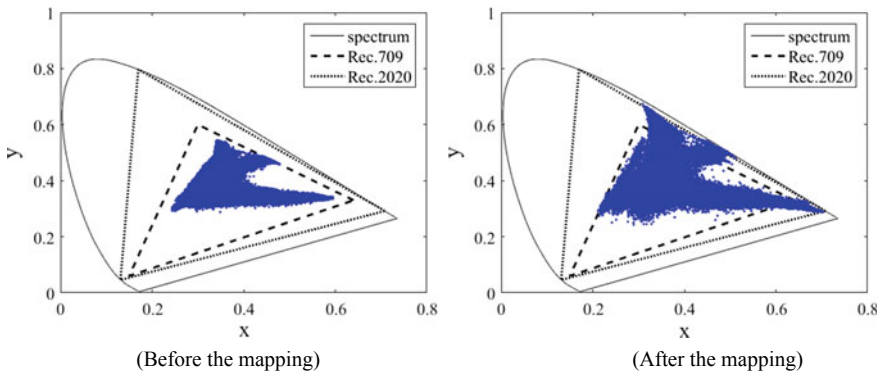


Fig. 3. Color contrast before and after image I1 expansion

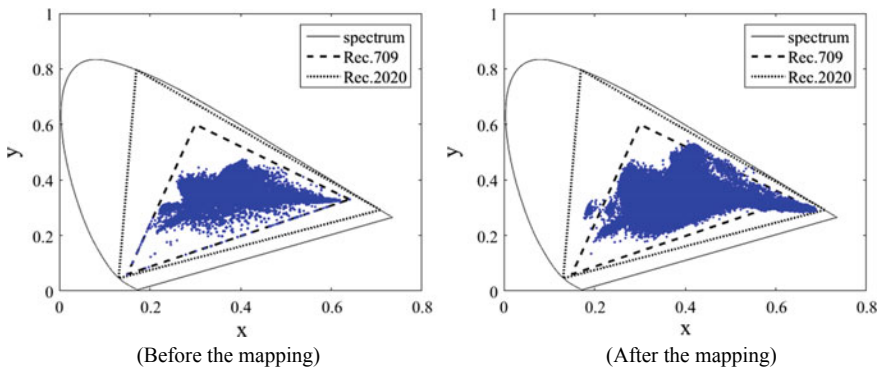


Fig. 4. Color contrast before and after image I2 expansion

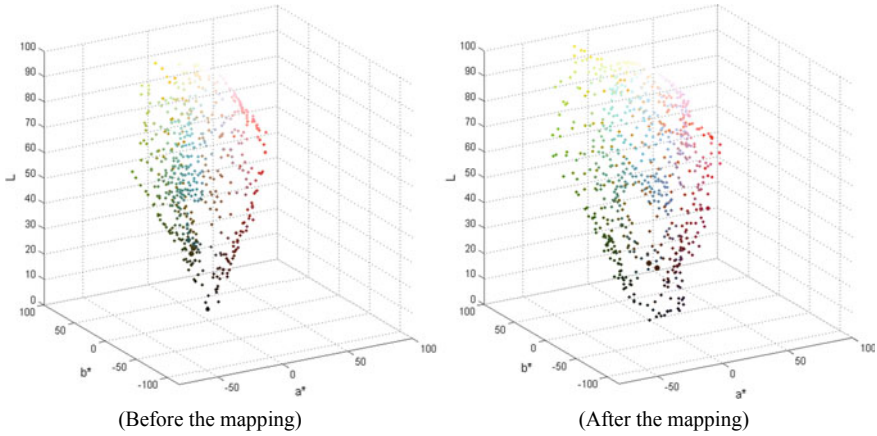


Fig. 5. Image I1 before and after the expansion of 3D color gamut comparison

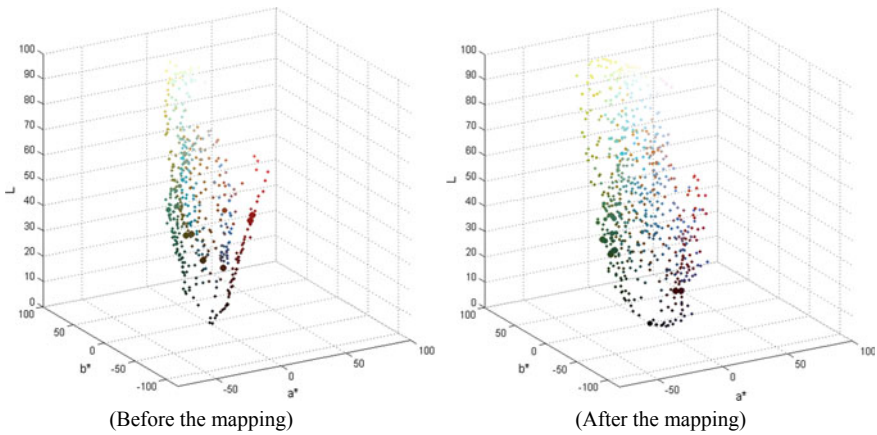


Fig. 6. Image I2 before and after the expansion of 3D color gamut comparison

The number of pixels in each image in Fig. 2 is 196608, reading the R, G, and B values of each pixel point, converting it to CIE 1931 XYZ chromaticity values, and drawing two-dimensional (x, y) data to obtain Figs. 3 and 4, respectively, to compare the distribution of the color gamut before and after the two test images in Fig. 2.

In Figs. 3 and 4, the most peripheral tongue shape curve is a spectral path, and the large triangle connected by a long dotted line is a wide target color gamut Rec.2020 in this paper. The point dashed line is connected with a smaller range as the conventional gamut range, and the source color domain used in the color gamut expansion of this article. This section takes I1 images with large difference before and after gamut expansion as an example to illustrate the extension effect of gamut mapping in this paper.

The I1 image is rich in red and yellow. Rec.2020 wide gamut has a more obvious gamut extension in red, yellow, green and other color regions. That is to say, compared with the conventional gamut display, Rec.2020 displays can display higher saturation and more red and yellow colors. In fact, these two colors also play an important role in improving the visual perception of the whole image.

In addition to the general improvement of color saturation, Figs. 3 and 4 show a commonality: the color gamut mapping is very similar before and after gamut mapping. That is to say, the color gamut expansion does not have a great influence on the relative position of each color point in the image, which helps to maintain the color of the original image, so that the image after the mapping is still realistic.

Figures 5 and 6 are three-dimensional gamut of two test images before and after gamut extension, respectively. CIE 1976 LAB uniform color space is used in color space. The same conclusion can be obtained from the two-dimensional data analysis.

In this paper, through bilateral filtering, the conventional color gamut images are divided into two layers of the basic layer and detail layer. Then, based on the basic layer data, the color gamut mapping is given by the established color gamut extension mapping. Finally, the detail layer is compensated to the image after the mapping, and the extended mapping of the color domain is finished. The experimental results show that the color gamut extension mapping algorithm can extend the color gamut from the conventional color gamut to the Rec.2020 WCG. The algorithm keeps the color features of the source image better.

Acknowledgements. This work was supported by the National Natural Science Foundation of China (Grant No: 61701344) and Student's Platform for Innovation and Entrepreneurship Training Program (Grant No: 201810065410).

References

1. La Torre, J.P., Mayes, N., Riza, N.A.: Laser display system for multi-depth screen projection scenarios. *Appl. Opt.* **56**(32), 9023–9029 (2017)
2. Su, X.-B., Ding, Y., Ma, B., Niu, Z.C.: Elimination of bimodal size in InAs/GaAs quantum dots for preparation of 1.3- μm quantum dot lasers. *Nanoscale Res. Lett.* **13**(1), 59 (2018)
3. Kang, B.-H., Morovic, J., Luo, R., Cho, M.-S.: Gamut compression and extension algorithms based on observer experimental data. *ETRI J.* **25**(3), 156–170 (2003)
4. Kang, B.-H., Cho, M.-K., Choh, H.-K., Kim, C.-Y.: Perceptual gamut mapping on the basis of image quality and preference factors. In: *Proceedings of SPIE 6058, Color Imaging XI: Processing, Hardcopy, and Applications*, p. 605806, 16 Jan 2006
5. Liu, Y., Song, G., Li, H.: A hue-preserving Gamut expansion algorithm in CIELUV color space for wide Gamut displays. In: *2010 3rd International Congress on Image and Signal Processing*, vol. 5, pp. 2401–2405 (2010)
6. Jiang, H.: Research and hardware implementation of color gamut extension algorithm. Master's thesis, Shanghai Jiaotong University (2007)
7. Li, Y.: Study on color range extension of laser display. Master's thesis, Shandong University (2012)
8. Huang, H.: Study on color gamut mapping based on wide color domain display equipment. Master's thesis, Shandong University (2014)

9. Han, Y.: Study on color gamut extension mapping based on wide color gamut display. Master's thesis, Shandong University (2015)
10. Wang, N., Zhu, M., Chen, G.: A spatial image color gamut mapping algorithm based on double filtering. *Electron. J.* **44**(8), 0372–2112 (2016)
11. Sun, K.: Research and implementation of high dynamic range image mapping algorithm based on bilateral filtering. Master's thesis, Xi'an University of Electronic Science and Technology (2015)
12. Wu, K.: Study on color image enhancement algorithm based on edge preserving filter. Doctoral dissertation, Chinese Academy of Sciences University (2017)



A Novel Hinge Structure of Fiber Bragg Grating Acceleration Sensor

Yuliang Li, Wei Wang^(✉), and Libo Qiao

Tianjin Key Laboratory of Wireless Mobile Communications and Power
Transmission, Tianjin Normal University, Tianjin 300387, China
weiwang@tjnu.edu.cn

Abstract. In view of the fact that existing Fiber Bragg Grating (FBG) accelerometers cannot achieve high-precision signal and high vibration frequency measurement at the same time, this paper proposes a novel type of FBG accelerometer based on the hinge structure. A novel mechanical structure of the sensor is designed, the elastic system is composed of a mass block and a flexible hinge, and the strain of the hinge amplified by the mass block is reflected by the fiber grating. Based on the analysis of the working principle of the new sensor, the influence factors of the resonant frequency and sensitivity of the sensor are studied. The mathematical model of the elliptical flexible hinge is used to optimize the design parameters of the sensor structure. The finite element analysis results show that the new structure sensor achieves the desired effect and it has a high resonance frequency (356 Hz) and sensitivity (284 pm/g).

Keywords: Fiber Bragg grating · Flexible hinge · Acceleration sensor · Finite element analysis

1 Introduction

Since the first fiber grating was invented, with the increasing development of fiber optic sensing technology, the use of fiber Bragg gratings in the field of sensing has become more and more widespread. In the twenty-first century, grating optical fiber sensing has experienced rapid development. Especially in the measure of acceleration, this technology has become a new hot spot for optical fiber sensing. Compared with traditional acceleration sensors, fiber Bragg gratings acceleration sensors belong to wavelength-coded optical sensors. Compared with the electrical sensors, there were many advantages such as anti-electromagnetic interference, the signal can be transmitted far away, intrinsic protection against electricity, and so on [1].

Based on this characteristic, people developed temperature sensors, stress sensors, strain sensors, displacement sensors, and so on, which have achieved a better engineering effect. In the field of navigation, because there are too many situations of uncertainties in the sea, the fiber Bragg gratings accelerometers adapted to hulls cannot be able to satisfy the actual needs. The cantilever beam shape is a common structure of most acceleration sensors, and it is also a common structure of FBG acceleration sensors developed by some scholars or research institutions [2]. Although this structure

can achieve the frequency requirements, the sensitivity is always limited by the frequency and it is vulnerable to horizontal interference.

2 Sensor Design

2.1 The Basic Structure of the Sensor

The structure of the hinge structure FBG acceleration sensor is shown in Fig. 1. It consists of a base fixed end, flexible hinge and a mass block. The mass is connected with the base through a flexible hinge, and the base is directly connected to the measured structure, both base and mass block have grooves; The sensor is made by a whole piece of beryllium bronze which is cut and processed linearly, and it is a holistic model. Two grating areas of the FBG are pasted between the upper and lower ends of the mass block and the base, respectively. Using the epoxy adhesive to paste the grating fiber and the sensor together, the fiber has a 2 mm pre-drawn.

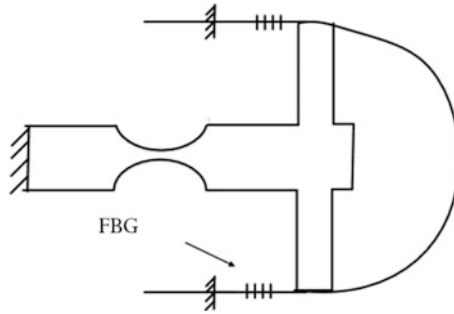


Fig. 1. Structure diagram

2.2 The Working Principle of the Sensors

The new sensor of this structure is an inertial elastic system consisting of a hinge and a mass, which responds to the mechanical vibration of the sensor. Using hinges as elastomers and grating fibers as sensing elements, when the hinge is subjected to an external force, strain is generated at the hinge portion, and the optical fiber grating attached to the structure is also displaced. At this time, the sensor converts the mechanical vibration signal into a lightwave signal.

The test structure is connected to the base of the sensor, when measured structural vibration and sensor will vary at the same frequency. At the same time, since the mass has a large inertia, the relative movement will produce between the mass and the base. Because it is connected with the optical fiber, the mass of the fiber grating will undergo slight expansion while the mass moves. According to the fiber coupling mode theory, when the fiber Bragg grating is affected by external influences, the center reflection wavelength will drift. Therefore, the grating fiber wavelength will change accordingly.

3 Mathematical Model and Calculation of Sensor

The elastic flexible hinge is the core of the structure and is the most critical part of the entire structural design. This paper establishes a mathematical model of the elliptical flexible hinge. Figure 2 is a schematic view of an elliptical flexible hinge, and this paper proposed a way to solve its stiffness.

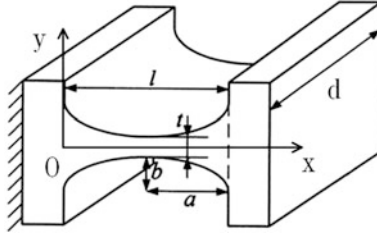


Fig. 2. Elliptical flexible

The neutral surface radius of curvature of the flexible hinge is [3, 4]

$$\frac{1}{\rho} = \frac{M(x)}{EL(x)} \tag{1}$$

where E is the elastic modulus of the structure, M(x) is the bending moment acting on a small distance dx, L(x) is the moment of inertia of the cross-section of the small segment distance dx facing the center axis, so the curvature at any point on the surface is

$$\frac{1}{\rho} = \frac{\frac{d^2y}{dx^2}}{\left[1 + \left(\frac{dy}{dx}\right)^2\right]^{\frac{3}{2}}} \tag{2}$$

Generally speaking, the change of the hinge bending moment is small, and M(x) can be regarded as a constant M. The deflection curve differential equation is expressed as follows:

$$\frac{d^2y}{dx^2} = \frac{M}{EL(x)} \tag{3}$$

The moment of inertia of the section to the center axis is

$$L(x) = \frac{d \times [h(x)]^3}{12} \tag{4}$$

where

$$h(x) = 2b + t - 2b\sqrt{1 - \frac{(a-x)^2}{a^2}} \tag{5}$$

where t is the minimum thickness of the hinge in the Y-axis, $h(x)$ is the thickness of each point on the Y-axis, and d is the thickness of the hinge on the Z-axis, so the angle θ of the elliptical flexible hinge is

$$\theta = \frac{dy}{dx} = \int \frac{d^2y}{dx^2} dx = \int \frac{12M}{Ed[h(x)]^3} dx \tag{6}$$

After the calculus operation, solve the rotational stiffness of elliptical flexible hinges:

$$K = \frac{M}{\theta} = \frac{Ed^2t}{24a\lambda} \tag{7}$$

We can draw a conclusion from the formula.

The rotational stiffness is affected by many factors, and it is inversely proportional to the long axis of the ellipse and nonlinearly increasing with the minor axis of the ellipse. It has the largest relationship with the minimum thickness of the hinge. In addition, it is linearly related to the elastic modulus of the material.

When the sensor structure vibrates with acceleration a in the vertical direction, according to the moment balance [5]:

$$MaC - k\Delta lh - K\theta = 0 \tag{8}$$

where C is the distance from the center of mass of the block to the center of the hinge, h is the vertical distance from the highest point of the mass to the mass center of the mass, k is the elastic coefficient of the fiber, and K is the rotational stiffness.

Therefore, the sensor resonant frequency can be obtained:

$$f_0 = \frac{\omega_0}{2\pi} = \frac{1}{2\pi} \sqrt{\frac{K_f + K_t}{M^*}} \tag{9}$$

Generally, the sensitivity of the FBG acceleration sensor is expressed by the ratio of the wavelength variation and the acceleration of the optical fiber.

The wavelength change is [6]

$$\Delta\lambda_B = (1 - P_e)\lambda_B\varepsilon_B \tag{10}$$

Sensitivity is

$$P = \frac{(1 - P_e)\lambda_B\varepsilon_B}{a} \tag{11}$$

The strain in the formula is

$$\varepsilon_B = \frac{\Delta L}{L} \quad (12)$$

In summary, natural frequency and sensitivity are key parameters in the sensor design process. The measurement range of the sensor is positively correlated with the natural frequency. The measurement precise will improve with the increase of the sensitivity. Therefore, natural frequency and sensitivity are important parameters for evaluating sensor performance. When the structure is made of beryllium bronze, the sensor has good elastic effect and it is easy to process. However, in the structural design, the sensitivity of the sensor will become smaller due to the increase of the resonant frequency. In order to improve the sensitivity of the acceleration sensor, the strain, and the resonant frequency, using a fiber grating with a gate length of 8 mm, the distance L between the two points on both sides of the paste is 20 mm. Using SMF28-C Acrylate fiber, Acrylate coating ensures that the fiber has enough strain.

4 Finite Element Analysis

The mechanical structure model of grating fiber acceleration sensor designed in this paper is relatively simple and can be modeled and analyzed directly in ANSYS15.0 software. Since the modulus of elasticity and density of the grating fiber are much smaller than beryllium bronze, it can be neglected.

The structural model diagram built in ANSYS is shown in Fig. 3. Before performing a series of operations, the structure needs to be meshed. The mesh is divided into several types, considering the complex meshing of the hinge part, and this paper uses the most effective method intelligent meshing [7]. Due to the large deformation at the hinge, smart size selects the optimal level and divides the result into 10,966 nodes and 6576 units. After meshing, impose constraints on the model. For this model, the displacement of the sensor on the base is constrained, so that the degree of freedom of displacement of the base in the XYZ direction is zero, and a gravitational acceleration is applied to the entire sensor structure. The effect is shown in Fig. 4.

4.1 Static Analysis

Static analysis ignores the effect of time on the load. Under conditions where the structure is constant, the load is applied to the structure and produces a response. Therefore, this paper calculates the effect of gravity acceleration on the mechanical structure. Figure 5 is the deformation of the structure in the X-axis direction after adding two gravitational accelerations in the negative direction of the Y-axis.

According to the finite element analysis, the displacement of the upper and lower ends of the mass is the largest, which is $\pm 4.71 \times 10^{-6}$ m. The distance between the two ends of the gate is 20 mm. According to formula (12), the fiber grating strain

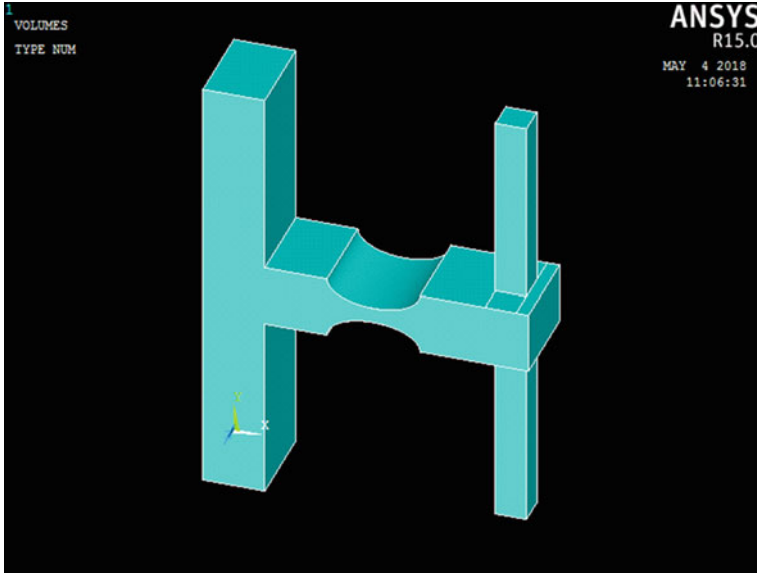


Fig. 3. Structural model

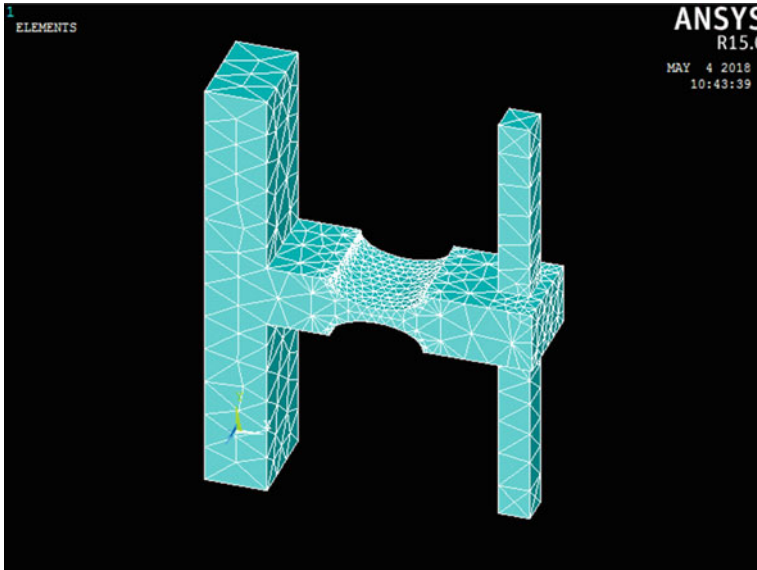


Fig. 4. Finite element mesh

is $\pm 235.5 \mu\epsilon$. The reflection wavelengths of the two gratings of the fiber grating are 1550 nm and 1555 nm, respectively. In the formula (11), the sensor sensitivity is 284 pm/g.

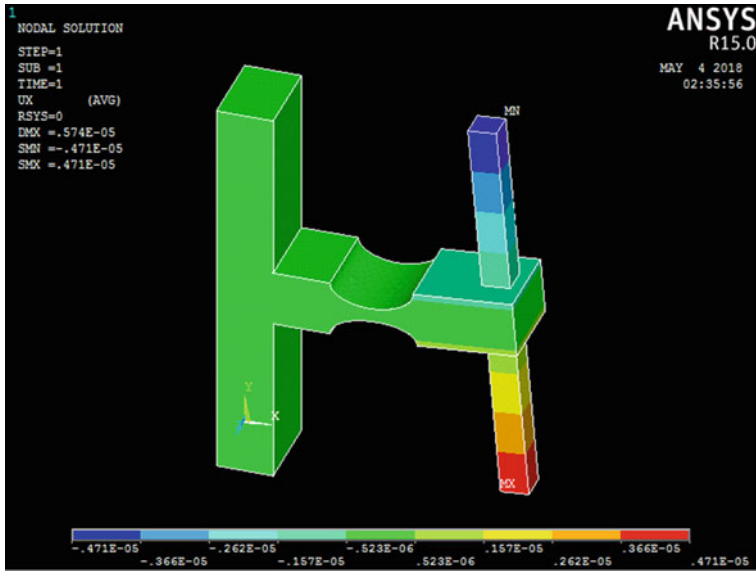


Fig. 5. Structural deformation

4.2 Modal Analysis

Modal analysis is used to determine the resonant frequency and mode shape of the sensor structure [8]. Figure 6 shows the first-order mode shape, which results in the vibration of the sensor in the Y-axis direction when the structure is fixed on the pedestal. Figures 7, 8 and 9 show the results of the second-order mode shape to the fourth-order mode shape, respectively.

From the modal analysis of the vibration pattern, the natural frequency of the FBG acceleration sensor is 356.448 Hz. The difference between the first-order modal frequency and the second-order to the fourth-order modal frequency is larger. Because the natural frequency of the sensor is related to the structural stiffness, the larger the modal frequency of the same structure, the greater the modal stiffness. Therefore, the difference between the first-order modal stiffness and the second to fourth orders is also large. It shows that the cross-coupling of this structure is very small, and it can meet the actual demand of sensor measurement accuracy.

4.3 Harmonic Response Analysis

Figure 10 is the harmonic response curve of the sensor from 0 to 500 Hz. It can be clearly seen from the figure that the obvious resonance phenomenon occurs at 355 Hz. Generally, in order to prevent the grating from being damaged due to the resonance response, the operating frequency of the sensor does not exceed one-fourth of the first-order modal frequency in the undamped state. In summary, the operating frequency range of the FBG acceleration sensor designed in this paper is around 0–89 Hz.

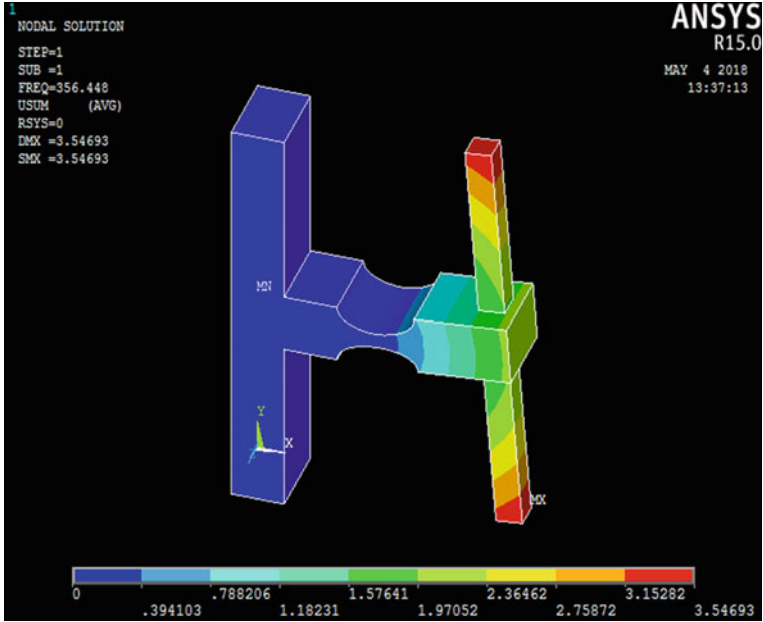


Fig. 6. First-order mode shape

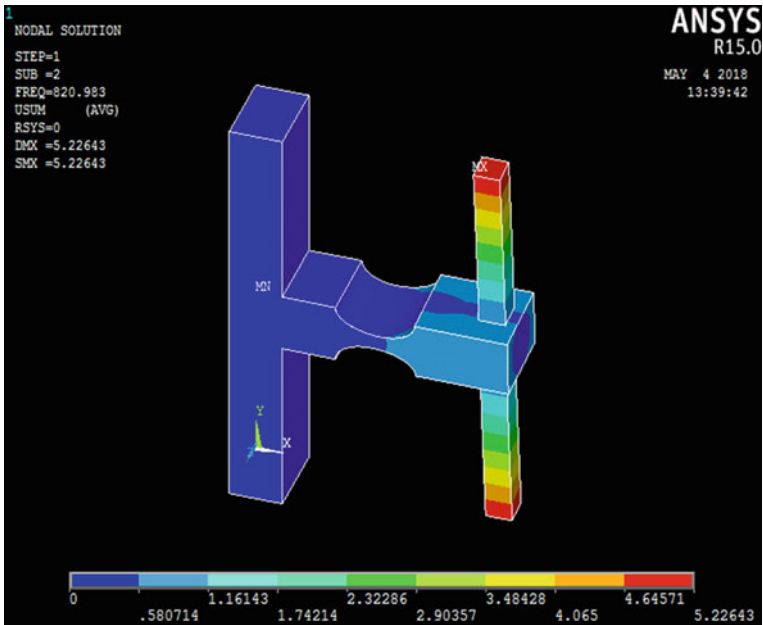


Fig. 7. Second-order mode shape

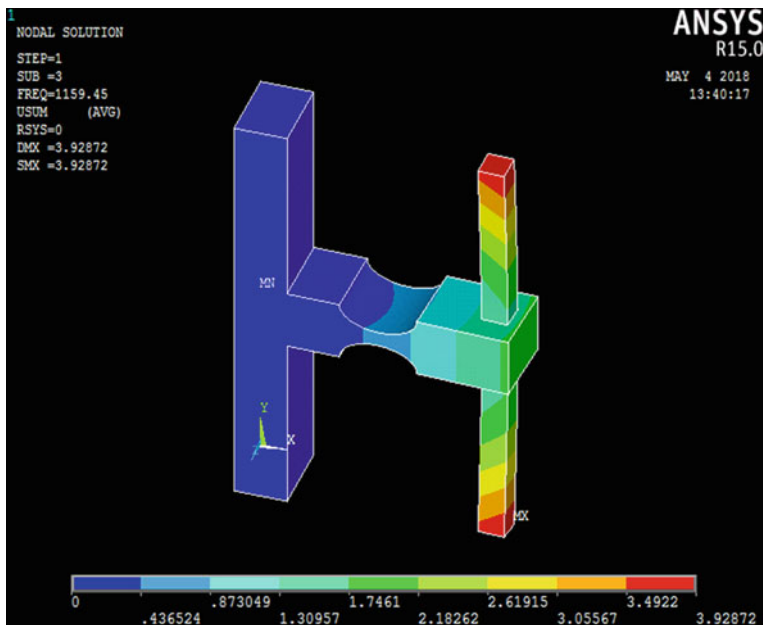


Fig. 8. Third-order mode shape

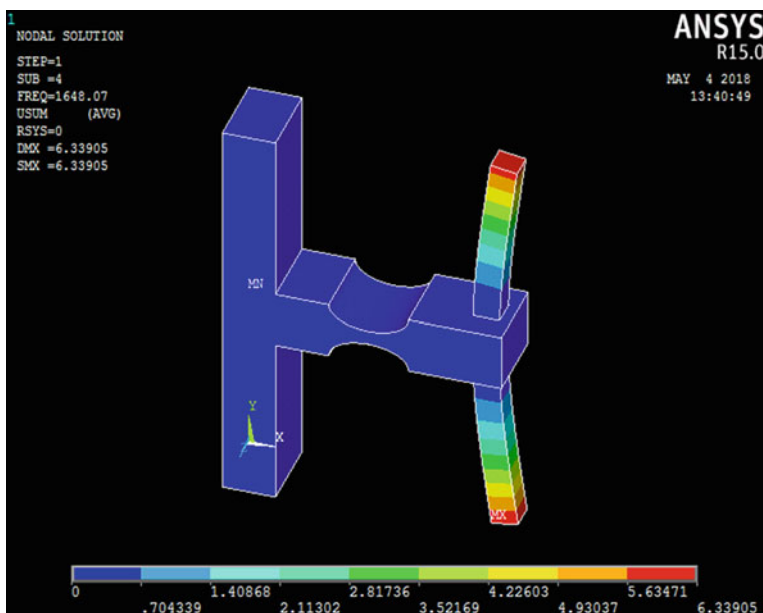


Fig. 9. Fourth-order mode shape

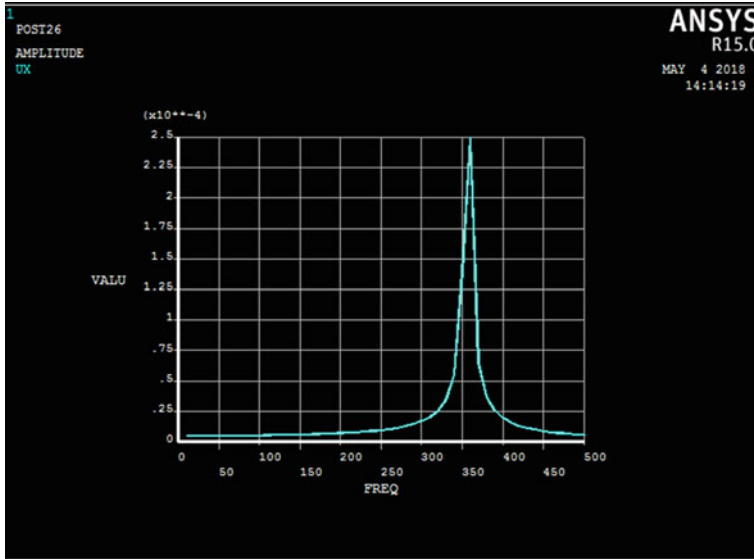


Fig. 10. Harmonic response curve

5 Conclusions

In this paper, a grating optical fiber acceleration sensor with a hinge structure is designed. The sensor is made of beryllium bronze and an elliptical flexible hinge connects the mass to the base. Through mathematical modeling analysis and optimization of structural parameters, an FBG accelerometer with high sensitivity and large measurement range is designed. Through finite element analysis, the natural frequency of the sensor is 356 Hz, and the measurement frequency is 0–89 Hz, and the sensitivity is 284 pm/g.

Acknowledgements. This paper is supported by Natural Science Foundation of China (61271411), Natural Youth Science Foundation of China (61501326, 61401310). It also supported by Tianjin Research Program of Application Foundation and Advanced Technology (15JCZDJC31500), and Tianjin Science Foundation (16JCYBJC16500).

References

1. Chen, D.: Research on dynamic characteristics of fiber Bragg grating acceleration sensor. Wuhan University of Technology, Wuhan (2004)
2. Guo, Y., Zhang, D., Zhou, Z., et al.: Research progress of fiber Bragg grating acceleration sensors. *Laser Optoelectron. Prog.* (2013)
3. Liang, L., Li, D., Qiu, L., Xu, G.: An FBG acceleration sensor based on flexible hinge. *J. Optoelectron. Laser* **4**, 347–352 (2016)
4. Wang, W., Yu, Y.: Dynamic modeling and analysis of elliptical flexible hinge compliant mechanism. *Mech. Des. Res.* **6**, 54–73 (2007)

5. Zhang, F., Jiang, M., Yan, Q., et al.: High-sensitivity low-frequency fiber grating acceleration sensor based on flexible hinge structure. *Infrared Laser Eng.* (2017)
6. Wang, W.: Study of key technology on ship hull structure health monitoring with fiber Bragg grating. Tianjin University, Tianjin (2007)
7. ANSYS15.0 finite element analysis from getting started to mastering. Tsinghua University Press
8. ANSYS14.5 finite element analysis. Tsinghua University Press

Circuit System Design



Study of Giant Magnetostrictive Thin Film Pressure Sensor Based on Villari Effect

Liyuan Dong^{1(✉)}, Shaopeng Yu^{1,2}, Tingting Han³, Bowen Wang²,
and Xinxin Cui²

¹ Department of Information Engineering, Renai College of Tianjin University,
Tianjin 301636, China

dongliyuan1987@126.com

² National-Ministry Joint Key Laboratory of Electromagnetic Field and
Electrical Apparatus Reliability, Hebei University of Technology, Tianjin
300130, China

³ Tianjin Key Laboratory of Wireless Mobile Communications and Power
Transmission, College of Electronic and Communication Engineering, Tianjin
Normal University, Tianjin 300387, China

Abstract. Pressure sensor can be used from industrial production to modern life with fast-growing technology and high-pursuit performance. Because of the advantages of strong magnetostrictive effect, high response speed, noncontact drive, and high electromechanical coupling, giant magnetostrictive thin film are not only used to miniaturize the sensor dimension, but also makes the sensitivity improved significantly. The induced voltage value of sensor detection coil can be determined by the intensity of external pressure and magnetic field. Meanwhile, the variation rate of magnetization depends on the differential magnetic susceptibility and the magneto-mechanical effect change rate by Villari effect, so a hysteretic nonlinear magneto-mechanical coupling model is established based on the J-A model and the law of energy conservation in this paper. The COMSOL simulation results show that the magnetization curve of the model can describe the magnetization trend and hysteresis characteristics better, and the relationship between input pressure and output voltage can be predicted, which verifies the correctness and accuracy.

Keywords: Pressure sensor · Villari effect · Giant magnetostrictive thin film · Coupling model

1 Introduction

The pressure sensor plays a very important role in the modern measurement and control system [1–4], which can be used in aerospace, robotics and intelligent home. Giant magnetostrictive thin film can be used to design sensor with the advantages of small size, high sensitivity, easy integration, etc. The paper studies the expression of induced voltage from J-A model and magneto-mechanical effect method, and then a dynamic coupling model is derived about magnetic field strength and external pressure which is based on Villari effect. As compared the simulation results with fact, the correct accuracy, wide practicability, and better hysteretic nonlinearity are verified. A pressure

sensor is designed by this model, the magnetic property change of giant magnetostrictive thin film is converted into the output voltage change of detection coil, so as to achieve the purpose of pressure detection. This sensor has passive and wireless transmission characteristics because there is no need to provide power and arrange wires, so it can achieve the internal signal measurement in complex structure.

2 Coupling Model

The voltage of sensor is determined by the magnetic flux change of coil, according to Faraday Law of Electromagnetic Induction:

$$u(t) = -\frac{d\phi}{dt} = -\frac{dNBS}{dt} \quad (1)$$

N is turn number of coil, S is cross-sectional area, B is magnetic induction intensity.

$$B = \mu_0(M + H) \quad (2)$$

μ_0 is permeability of free space, M is magnetization, and H is magnetic field strength.

Substituting Eq. (2) in Eq. (1):

$$u(t) = -NS\mu_0\left(\frac{dM}{dt} + \frac{dH}{dt}\right) \quad (3)$$

Magnetic field strength is only related to the current through the excitation coil. Magnetization is nonlinear magneto-mechanical coupling term related to magnetic field and external pressure, which is the function of excitation magnetic field H , stress σ and temperature of environment T :

$$\frac{dM}{dt} = \frac{\partial M}{\partial H} \frac{dH}{dt} + \frac{\partial M}{\partial \sigma} \frac{d\sigma}{dt} + \frac{\partial M}{\partial T} \frac{dT}{dt} \quad (4)$$

$\partial M/\partial H$ is differential susceptibility, expressing the magnetization induced by magnetic field. $\partial M/\partial \sigma$ is rate of magneto-mechanical effect, expressing the magnetization induced by stress. The eddy current loss and temperature effect of thin film can be neglected, so $(\partial M/\partial H) \cdot (dT/dt)$ is zero. It can be seen that $\partial M/\partial H$ and $\partial M/\partial \sigma$ are the key to build the model, $\partial M/\partial H$ can be determined by J-A model, and $\partial M/\partial \sigma$ can be determined by magneto-mechanical effect method based on the Law of Conservation of Energy.

2.1 Differential Susceptibility Based on J-A Model

J-A model is established on Weiss molecular field theory and physical principle and based on the mathematical description of the domain wall, so it is not strict derivation and experimental data. It quantifies the non-hysteretic, reversible and irreversible

magnetization components, and uses them to characterize the total magnetization generated by the input magnetic field [5, 6]. From the initial magnetization to saturation magnetization, atomic magnetic moment inside ferromagnetic material has strong molecular field, which can make the moment transferred to mutual parallel trend, so the internal magnetic field of material is superposition of external magnetic field and molecular field inside the atom and is given by

$$H_e = H + \alpha M \quad (5)$$

$$M_{an} = M_s \left(\coth \frac{H_e}{a} - \frac{a}{H_e} \right) \quad (6)$$

$$\frac{dM_{irr}}{dH} = \frac{M_{an} - M_{irr}}{\delta k - \alpha(M_{an} - M_{irr})} \quad (7)$$

$$M_{rev} = c(M_{an} - M_{irr}) \quad (8)$$

$$M = M_{rev} + M_{irr} \quad (9)$$

H_e is effective magnetic field, M_{an} is magnetization without hysteresis, M_{irr} is irreversible magnetization, M_{rev} is reversible magnetization, δ is direction parameter, a is shape factor, α is domain wall interaction coefficient, k is irreversible loss coefficient, c is reversible coefficient.

The formula can be worked out by combining Eqs. (5)–(9):

$$\frac{\partial M}{\partial H} = \frac{\frac{M_{an}-M}{\delta k - \frac{\alpha}{c}(M_{an}-M)} + \frac{cM_s}{a} \left[1 - \coth^2 \frac{H_e}{a} + \left(\frac{a}{H_e} \right)^2 \right]}{1 - \frac{\alpha c M_s}{a} \left[1 - \coth^2 \frac{H_e}{a} + \left(\frac{a}{H_e} \right)^2 \right]} \quad (10)$$

Equation (10) is the function of the magnetization changed with magnetic field strength. With no or constant external pressure, the magnetization is only related to the applied magnetic field. In order to facilitate the simulation, the formula can be simplified in the case of non-distortion by Langevin Function. So Eq. (6) can be expanded as

$$\frac{M_{an}}{M_s} = L(0) + \frac{L'(0)}{1!} \frac{H'_e}{a} + \frac{L^{(n)}(0)}{n!} \left(\frac{H'_e}{a} \right)^n + o \left[\left(\frac{H'_e}{a} \right)^{n+1} \right] \quad (11)$$

$n = 2$ is usually taken and higher order terms is ignored as

$$M'_{an} = \frac{H'_e}{3a} M_s \quad (12)$$

Simplifying Eq. (10) as

$$\frac{\partial M}{\partial H} = \frac{3\alpha(1-c)(M_{an}-M)}{(1-c)\delta k - \alpha(M_{an}-M)} + cM_s \tag{13}$$

The relationship between magnetization and magnetic field strength is simulated as Fig. 1. It can be seen that the numerical simulation result is not very smooth, but the trend of curve is basically consistent and linear, which is because the temperature and other environmental factors are not considered, external magnetic field, however, the model still can well describe the magnetization trend, which can prove the correctness of basic numerical solution.

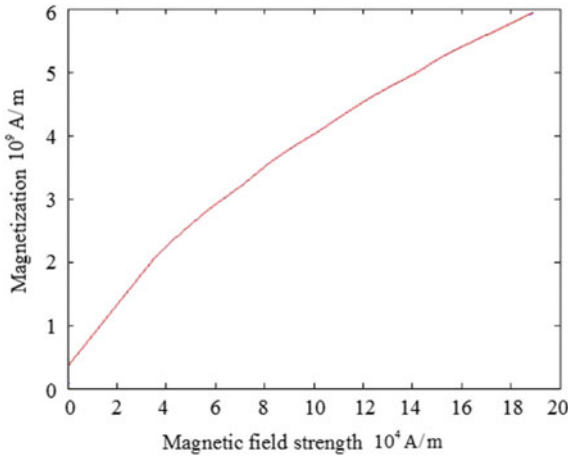


Fig. 1. Relationship between magnetization and magnetic field strength

2.2 Rate of Magneto-mechanical Effect Based on Magneto-mechanical Effect Method

Jiles and Sablik believe that non-hysteresis magnetization has the relationship with stress, meanwhile, the experimental results show that the magnetization of ferromagnetic material can be affected by stress in the form of strain, so the interior magnetization is related to the applied magnetic field, molecular field, and stress. Introducing stress equivalent magnetic field H_σ which is the magnetic effect of external pressure on material [7], given by

$$H'_e = H + \alpha M + H_\sigma \tag{14}$$

Stress equivalent magnetic field can be derived by the Law of Thermodynamics as

$$H_\sigma = \frac{3\sigma}{2\mu_0} \frac{\partial \lambda}{\partial M} \tag{15}$$

Here, the magnetostrictive coefficient λ has the power function with magnetization as

$$\lambda = \sum_{i=0}^{\infty} \gamma_i M^{2i} \approx \gamma_1 M^2 + \gamma_2 M^4 \quad (16)$$

When the external pressure is applied to ferromagnetic material, the elastic energy and magnetic energy can be converted into each other. The change rate of total magnetization to the unit volume energy is proportional to the displacement of total magnetization to the non-hysteresis magnetization, so the relationship between the irreversible magnetization and elastic energy can be expressed as

$$\frac{dM_{irr}}{dW} = \frac{1}{\xi} (M_{an} - M_{irr}) \quad (17)$$

For isotropic material, the elastic energy can be expressed as

$$W = \frac{\sigma^2}{2E} \quad (18)$$

ξ is unit volume energy factor, related to the load form, E is Young's modulus, and derived as

$$\frac{dM_{irr}}{d\sigma} = \frac{\sigma (M'_{an} - M_{irr})}{E\xi} \quad (19)$$

The total magnetization, reversible magnetization, and irreversible magnetization are consistent with the J-A model without stress as

$$M'_{an} = M_s \left(\coth \frac{H'_e}{a} - \frac{a}{H'_e} \right) \quad (20)$$

$$M_{rev} = c' (M'_{an} - M_{irr}) \quad (21)$$

$$M = M_{rev} + M_{irr} \quad (22)$$

The formula can be worked out by combining Eq. (14) to Eq. (22):

$$\frac{\partial M}{\partial \sigma} = \frac{\frac{\sigma (M'_{an} - M)}{E\xi} + \frac{3c' M_s}{2a\mu_0} \left[1 - \coth^2 \frac{H'_e}{a} + \left(\frac{a}{H'_e} \right)^2 \right] \frac{\partial \lambda}{\partial M}}{1 - \frac{\alpha c' M_s}{a} \left[1 - \coth^2 \frac{H'_e}{a} + \left(\frac{a}{H'_e} \right)^2 \right]} \quad (23)$$

Equation (23) is the function about magnetization changed with stress. With no or constant magnetic field strength, the magnetization is only related to the applied pressure. Same as Eqs. (13) and (23) can be simplified as

$$\frac{\partial M}{\partial \sigma} = \frac{\frac{a\mu_0\sigma(M_{an}-M)}{E\xi} + cM_s(\gamma_1 M + 2\gamma_2 M^3)}{a\mu_0 - c\sigma M_s(\gamma_1 + 6\gamma_2 M^2)} \tag{24}$$

The numerical result can be easily seen from Fig. 2 that the change of magnetization obtained from the Eq. (24). The external pressure and the magnetic intensity are almost linear, so that the design of the sensor is convenient. With no or constant external magnetic field, the magnetization is only related to the applied pressure. The change of magnetization intensity is obvious, so the generated voltage signal can be detected by the amplification circuit with perfect performance.

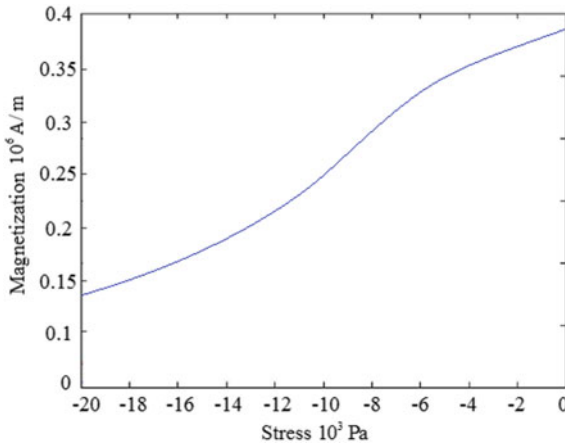


Fig. 2. Relationship between magnetization and stress

3 Structure of Sensor

The sensor is composed of three parts: giant magnetostrictive thin film is on the top which can perceive the strain of the substrate, and transmit the strain into the change of the magnetic induction strength. Binder is in the middle which can maintain the magnetostrictive material and substrate change synchronously. Substrate is at the bottom which can make the sensor bear the greater external force and extend the measurement range. Due to the small thickness of giant magnetostrictive thin film, while the thickness of substrate is larger, so the change of sensor stiffness can be neglected that is caused by the magnetostrictive material. Pressure sensor structure is shown in Fig. 3.

The signal generator produces sinusoidal signal, which is amplified by amplifier and input excitation coil to provide alternating excitation signals required by the work of sensor. The detection coil detects the output of sensor and inputs the signal to spectrum analyzer in order to get the valid signal. Finally, the computer calculates and analyzes the valid signal to obtain the external pressure. The experimental testing system is shown in Fig. 4.

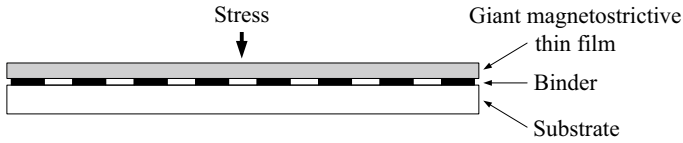


Fig. 3. Pressure sensor structure

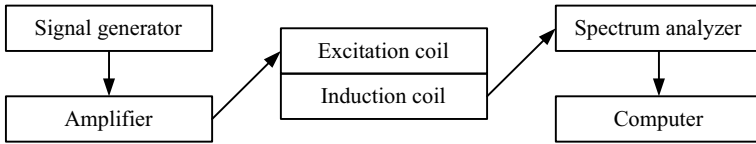


Fig. 4. Experimental testing system

4 Simulation and Analysis

The sensor model is regular structure, so two-dimensional axisymmetric simulation is selected in the COMSOL, so the finite element model can be established. The film thickness of the sensor is established with 0.05 mm as interval, so the simulation model with different height range from 0.1 to 0.5 mm can be obtained. It is found that the displacement of calculation point reaches the maximum and the current–displacement curve has the best linear effect when the distance is 0.35 mm, therefore, the simulation results of the 0.35 mm film type is shown. The strain intensity can be seen in Fig. 5, the substrate is set as fixed boundary condition, so the concentration effect of stress and strain is reflected on the bottom of film.

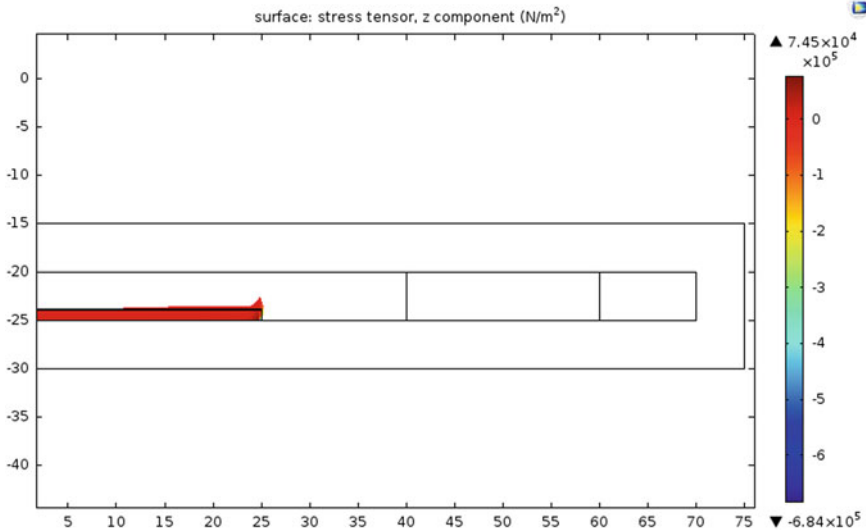


Fig. 5. Strain distribution

The relationship curve between current and displacement can be seen in Fig. 6, displacement of giant magnetostrictive thin film increases nearly linear with the current change of coil, which can achieve the function of micro-displacement sensor.

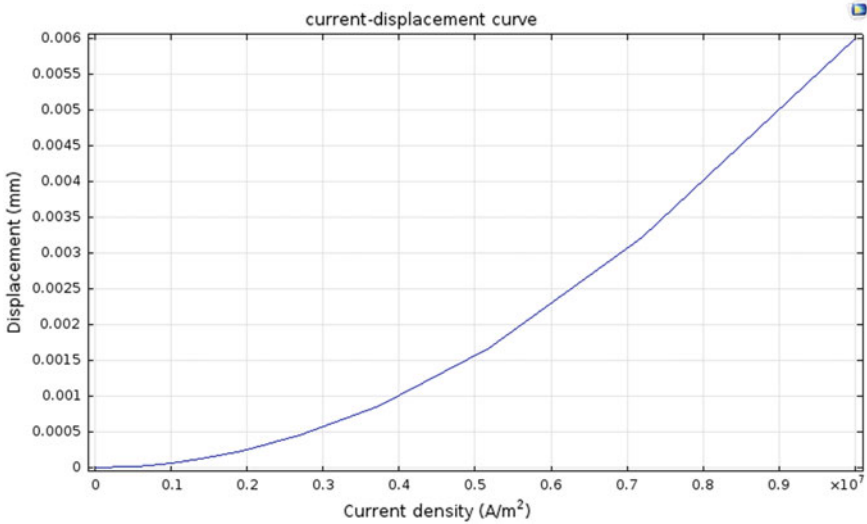


Fig. 6. Current–displacement curve

5 Conclusion

The dynamic magneto-mechanical model coupling is derived according to J-A model and magneto-mechanical effect method, and the relationship between induced voltage with magnetic field strength and external pressure based on Villari effect is shown. COMSOL is used to build the finite element calculation model of the giant magnetostrictive thin film sensor, the distribution of strain and the current–displacement curve is obtained. From the simulation results, it can be seen that the model has a better capability of accuracy and hysteretic nonlinearity. Therefore, it is proved that the magneto-mechanical coupling model can describe the electromechanical coupling and output response of pressure sensor effectively.

References

1. Claeysen, F., Lhermet, N., Leletty, R.: Actuators, transducers and motors based on giant magnetostrictive materials. *J. Alloy. Compd.* **258**, 61–73 (1997)
2. Lu, Q.G., Chen, M., Chen, D.F.: Research on hysteresis nonlinearity of GMA based on Preisach modified model. In: *7th International Conference on Computer-Aided Industrial and Conceptual Design*, pp. 829–833 (2006)

3. Li, H., Preidikman, S., Balachandran, B.: Nonlinear free and forced oscillations of piezoelectric microresonators. *J. Micromech. Microeng.* **16**(2), 188–196 (2006)
4. Benbouzid, M.E.H., Body, C., Reyne, G.: Finite element modeling of giant magnetostriction in thin films. *IEEE Trans. Magn.* **31**(6), 3563–3565 (1995)
5. Chakrabarti, S., Dapino, M.J.: Coupled axisymmetric finite element model of a hydraulically amplified magnetostrictive actuator for active powertrain mounts. *Finite Elem. Anal. Des.* **60**, 25–34 (2012)
6. Jiles, D.C., Atherton, D.L.: Theory of the magnetization process in ferromagnets and its application to the magnet mechanical effect. *J. Phys. D Appl. Phys.* **17**, 1265–1281 (1984)
7. Sablik, M.J., Chen, Y., Jiles, D.C.: Modified law of approach for the magnetomechanical model. *Rev. Progr. Quant. Nondestruct. Eval.* **19**, 1565–1572 (2000)



Research on ATML-Oriented Test Application Development Platform

Jin Luo, Hua Yang^(✉), and Jing Huang

Army Infantry College, Nanchang, China
zoe_luojin@163.com, 434639978@qq.com

Abstract. To enhance the development efficiency of test programs in Automatic Test System (ATS), a framework of test application development platform based on ATML (Automatic test description language) standard is built in this paper. ATML-oriented test application development platform contains configuration tools, test flow description, ATS resource description, and runtime system. ATML is used to describe ATS-related information and the test flow based on signal definition. By leveraging the ATML description documents, runtime system achieved the automatic matching of the test resource and test requirements, and generated the test programs automatically. The platform is extensible, and the test program sets can be transplanted to other platforms.

Keywords: ATS · ATML · Test function service · Test program automatic generation

1 Introduction

Currently, ATS engineers developed the test program sets in an inefficient and error-prone process as follows: the test requirements and test flow for a product are created by product development engineers, when it is time to develop the tests for these units, the software engineers convert the test requirements into test programs manually. To improve the programming efficiency and realize the transplantation of the test software in ATS, a new automatic generation technology for test programs is needed.

ATML provides a standard for test information description in ATS, it defines classes that represent test system components and the interoperability between classes. This paper will discuss a new framework of test application development platform based on ATML standard. By leveraging the ATML test description standard for creating test requirement documents, the work of generating test program from test requirements can be completed by the test application development platform, so as to achieve the unambiguous transplantation of test requirements among different platforms.

2 The Framework of Test Application Development Platform

To enhance the transplantation and development efficiency of test program in ATS, this paper proposes a new framework of test application development platform based on ATML, which is shown in Fig. 1.

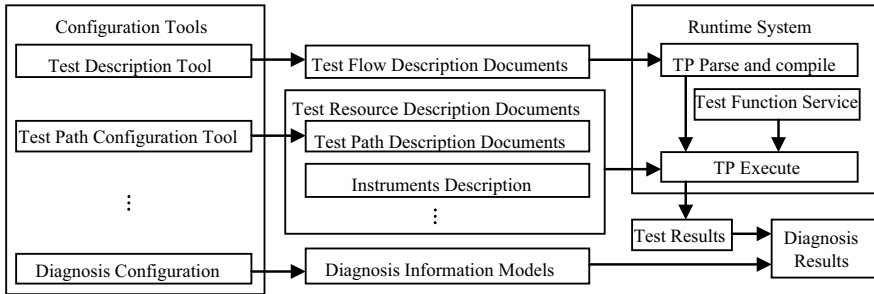


Fig. 1. Framework of test application development platform

Test application development platform consists of the configuration tools, test flow description documents, ATS resource description documents, and runtime system. Three main configuration tools in the platform are test description tool, test adapter configuration tool, and diagnosis configuration tool. By utilizing the test description tools, test flow description XML documents are generated automatically according to ATML standards. Test flow description document describes the test requirement and test execution order for UUT. Test flow description document includes UUT information description, interface information, test resource requirements, signal requirements, etc. Runtime system is the center of the test application development platform. In runtime system, test flow description documents are converted into C++ codes as the intermediate program, then the C++ codes are compiled and linked into executable test program by using the commercial compiler. The test function server components in runtime system support the automatic matching of test capabilities and test requirements based on the test resource description information and test function interface definition. Test function server components classified the signal, integrated different instrument drivers into the same test function interface, and executed test function by calling the instrument drivers according to the test path. Runtime system based on test function server realizes the expandability and interoperability, which support higher developing efficiency of the TPSs.

3 Description of Test Information

Test description documents based on ATML standard describe the test requirement information, such as UUT description, performance characteristics, interface requirements, signal requirements, fault data, and detailed test information, which is necessary to verify the correct functioning of a UUT. One of the important elements in an ATML

test flow description document is a DetailedTestInformation element, which describes the input condition, test requirements, and test flow for UUT test. The logical model view for the structure of element DetailedTestInformation is shown as Fig. 2. This paper modified the ATML test description Schema by reducing some optional data items since too much data items would reduce the parsing rate of the XML documents. DetailedTestInformation element comprises a variety of detailed test information of the entry points, testgroups, and actions. Test flow description is described by testgroups and actions elements.

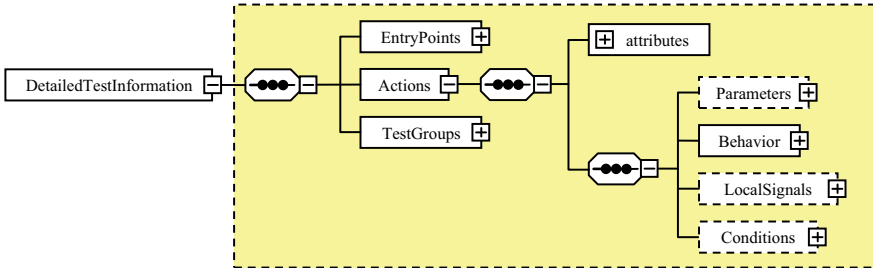


Fig. 2. Structure of element DetailedTestInformation

Actions element are used to describe a group of test or session action necessary for test execution. Actions element contains parameters, behavior, local signals, and conditions. Parameters identify the action parameters and specify its values. In current actions, local signal may be shared by different operations, we can use LocalSignals child element to define the local signals in current actions. Different conditions verify the correct execution of the action. There are two conditions to support the action execution, one is precondition, which verifies the action be executed when the conditions are satisfied, the other is post-conditions, which ensure the action can achieve after successful execution. Conditions child element specifies both conditions for the action. Behavior element describes the detailed operation procedure and tasks for the current action executing, it is the most important element in action. Behavior element can be a call to the test group by specifying its ID, or a sequence of predefined operation according to the test definition.

TestGroups are used to describe a group of actions that can be referenced from multiple places in a test description document. TestGroups element comprises one or more TestGroup elements, each TestGroup element contains one or more actions, each action contains one or more behaviors, and each behavior in an action is executed sequentially. TestGroup ensures the correct execution order of the actions by specifying its conditions, initial operation, and terminal operation. This mode of description for test order is similar to a fault tree. Different TestGroup elements specify test flow of the UUT test program like the control structure in programming language.

4 Translation of Test Description Document and Runtime Service

In the test application development platform, the test program set is replaced by the test requirement set, and the automatic generation of the executable test program from test requirements is completed by runtime system. The test procedure information contained in test description document is converted into C++ codes as the intermediate program. Runtime system provided the test function service to realize the automatic matching of test resources and test requirements and make the software with excellent expandability.

4.1 Static Data Conversion in Test Description Documents

In test description documents, static data includes data type, parameters definition, and other description information except test flow description. Mostly data type definitions contain common elements and signal definitions defined in STD standard. Some primitive data types can be converted into the data types in C++ language directly, such as integer and float; some complex data type can be converted into customized data type, such as struct and union. The abstract data type can be represented by using struct, abstract class, or union. Struct data type is used to represent the derived data type, and union is used to represent the abstract type which needs to be passed as a parameter in the COM interface function. Runtime system converts the static data in XML-based test description documents into data type and parameters definition which is indicated in C formal, and store the conversion result in *.c and *.h files.

4.2 Translation of Test Flow Description to Test Sequences

Test flow description is described by TestGroups and Actions elements. Different TestGroups elements specify the test flow, Actions elements specify a group of test or session action for test execution. Therefore, we can convert the test flow description into the control structure in C++ language.

Each TestGroup element will be translated into a corresponding test sequence, since a test sequence in test program has similar constructs as the test group in test description document, such as passing the test result parameters, specifying their names, setup or clean up step groups, and so on. Every step defined in TestGroup element will be translated into a test step in the main test program sequence with a next action expression according to its execution result.

Each Actions element can be translated into functions in C++ language since it has similar constructs as functions. The parameters of Actions will be converted to function parameters, the LogicSignals element will be converted to signal definition, and the Conditions element will be converted to “switch” structure or “if...else...” structure according to its outcome.

Take a project for the DC Voltage output in the test description document as an example, which is described as Fig. 3. The name of Action in this project specifies the C++ function name “Apply_DC_Volt”. This Behavior of Action includes four

operations to apply DC voltage to UUT, namely OperationSetup, OperationConnect, OperationDisconnect, and OperationReset. OperationSetup defined the source signal for DC power, OperationConnect connected the signal source to UUT power pins, OperationDisconnect disconnect the source and UUT, and OperationReset reset the Power. Translator translates the operations to a corresponding test function call according to the sequence defined in TestGroup. Signal characters description of the Operation is used to assign value to the member variables of signal characteristics, such as “dc_ampl = 12.0V”. According to the principle and method, test description of “Apply_DC_Volt” project and its converted code is shown in Fig. 3.

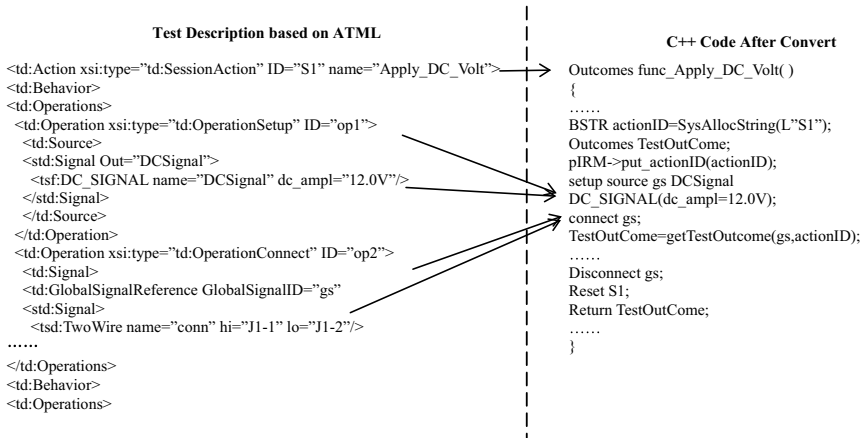


Fig. 3. Test description code and C++ Code

4.3 Resource Mapping and Test Function Service

Test configuration documents provided the test configuration information for UUT measurement, such as instruments, test adapters, test paths, test software, and test requirement, etc. Based on the test configuration information, runtime system can calculate switch paths based on the test adapters description, acquire the test resource information required to execute a test, achieve the correct matching of the test resource and test requirement according to the test paths, call the instrument drivers based on the instrument information, and generate the executable test programs automatically.

Test is defined as behavior by using operations, and the operations are interpreted into the test function service call. Runtime system classified the signal test types based on the test function, and defined the test function interface with the same test function service name for different signal classes, such as Apply_DC_Volt, Measure_Resistance, etc. The test function of the instrument resources is abstracted, and the instrument driver is encapsulated in the dynamic linking library for test programs calling based on its test function. Therefore, test function interface separated the instrument control from the test program, shielded the differences of the specific instrument, and made the software with excellent expandability.

5 Conclusion

To enhance the programming efficiency of ATS, a new framework of test application development platform based on ATML is built. It uses the test requirement set to replace the test program set, and generates test program from test requirements automatically. This paper discusses the method for translating test requirement document to a C++ test executive program, examines how test description document constructs such as test group, test action, and behavior can be mapped to constructs in the test program, and expatiates the mechanism of the test function service provided by runtime system.

References

1. IEEE Std 1671-2010.: Standard for Automatic Test Markup Language for Exchanging Automatic Test Equipment and Test Information via XML (2011)
2. Fan, S.: Study of test resources and test requirements automatic matching method in ATML. *Computer Meas. Control* **24**(6), 8–11 (2016)
3. Jain, A., Delgado, S.: Automatic ATML test description translation to a COTS test executive. In: *IEEE AutoTestCon*, pp. 190–194 (2009)
4. Ordonez Camacho, D., Mens, K.: Apparel: a tool for building automated program translators using annotated grammars. In: *IEEE ICASE*, pp. 489–491 (2008)
5. Zhu, S.: Research and application of test application development platform on STD standard. *Comput. Technol. Autom.* **29**(4), 111–114 (2010)



A School Violence Detection Algorithm Based on a Single MEMS Sensor

Jifu Shi¹, Liang Ye^{1,2(✉)}, Hany Ferdinando^{2,3}, Tapio Seppänen⁴,
and Esko Alasaarela²

¹ School of Electronic and Information Engineering, Harbin Institute of Technology, Harbin, China

shijiful993@126.com, yeliang@hit.edu.cn

² OPEM Unit, Health and Wellness Measurement Research Group, University of Oulu, Oulu, Finland

{hferdina, esko.alasaarela}@ee.oulu.fi

³ Department of Electrical Engineering, Petra Christian University, Surabaya, Indonesia

⁴ Physiological Signal Analysis Team, University of Oulu, Oulu, Finland
tapio@ee.oulu.fi

Abstract. School violence has become more and more frequent in today's school life and caused great harm to the social and educational development in many countries. This paper used a MEMS sensor which is fixed on the waist to collect data and performed feature extraction on the acceleration and gyro data of the sensors. Altogether nine kinds of activities were recorded, including six daily-life kinds and three violence kinds. A filter-based Relief-F feature selection algorithm was used and Radial Basis Function (RBF) neural network classifier was applied on them. The results showed that the algorithm could distinguish physical violence movements from daily-life movements with an accuracy of 90%.

Keywords: Activity recognition · School violence · Relief-F · MEMS accelerometer · RBF neural network

1 Introduction

School violence is a worldwide problem, and relevant survey data showed that 50% of the students admitted to the existence of bullying behavior, whereas 47% of the students confirmed that they have been hurt by verbal bullying or physical violence. It has caused many students to fear campus life and seriously affects students' physical and mental health as well as their studies. Throughout recent years, there is a common problem of many campus violence that bullied students are inflicting psychological shadows on their young hearts because they are afraid or threatened not to inform parents or teachers about their being bullied. Traditional campus monitoring has the problem of blind area because of geographical location restriction, and cannot realize active detection so as to prevent violence efficiently.

With the development and popularization of various smart wearable devices, smartphones, and bracelets have become the most common wearable sensor devices. These devices have built-in sensors, including optical sensors, inertial sensors (accelerometers and gyroscopes), GPS locators, and so on. Many scholars use these technologies to identify human daily behaviors and recognize complex behaviors. In this situation, the study of active detecting of school violence has come into being. We use active sensors which are built into wearable devices on the students to actively detect campus violence in real-time and transmit the results to teachers or parents in an active and timely manner. This active detection avoids being bullied and undiscovered, and can solve school violence problems in a timely and effective manner.

The remainder of this paper is organized as follows: In Sect. 2, we review the related work on activity recognition; in Sect. 3, we describe the entire system framework and physical bullying detection algorithm; in Sect. 4, we present the classification results; and the conclusion is given in Sect. 5.

2 Activity Recognition

Study of school bullying was originally initiated by some European scholars for passive methods. Along with the popularity of smartphones on campus, there have been a number of school bullying detection schemes which are based on the intelligent mobile phone. For example, “Stop Bullies”: there is a shortcut key on the mobile phone, we can press this key to send photos of the incident location and geographical location to the intended recipients when suffered bullying. However, a common drawback of these detection schemes is passive detection, the victims are often too late to operate or dare not to operate when they are suffering from bullying, which greatly reduces the detection rate and does not solve school bullying in a timely and effective manner.

With the development of pattern recognition technology, great progress has been made in the field of physical activity recognition. The physical activity recognition system using single or multi-sensors is used in real life. Some scholars have used the accelerometer to obtain good results in many fields. Reference [1] used a three-axis accelerometer to propose an accurate physical activity recognition system for medical monitoring and rehabilitation. A patient wore a wireless acceleration sensor to identify six daily movements and transition events. Feature selection was performed using SFFS and the naive Bayes and KNN classification algorithms were adopted. The recognition accuracy rate was 95%. In addition, PA (Physical Activity) recognition systems based on handheld mobile terminals have become a research hotspot [2–5]. In [6], accelerometers and gyroscopes in smartphones were used to identify daily activities, and the recognition accuracy could reach about 85%. However, considering the battery capacity and memory size of handheld mobile terminals, reducing the complexity of the algorithm has also become the optimization goal for many scholars [7]. Multi-sensor data fusion enables a more comprehensive capture of human movements compared with a single sensor [8–13]. This paper proposes a posterior-adapted class-based fusion of multi-accelerometers data algorithm by using waist and leg sensor data. The results show that the recognition accuracy is improved compared with a single sensor.

3 Methods

The system framework includes data collection and preprocessing, feature extraction, normalization, feature selection, and classifier design. This study considers time efficiency and resource utilization. Algorithm complexity is a priority for the design of feature selection algorithms and classifiers in this paper.

3.1 Data Acquisition and Preprocessing

The background of this research project is for the campus violence scene. In order to simulate campus violence scenes more realistically, data acquisition was performed indoors with protective gears and motion sensors. Data acquisition consisted of ten subjects who were asked to perform six types of daily-life activities (jump, play, run, walk, stand, and fall-down) and three bullying activities (beat, push, and push-down). The position of the motion sensor on the waist is shown in Fig. 1.

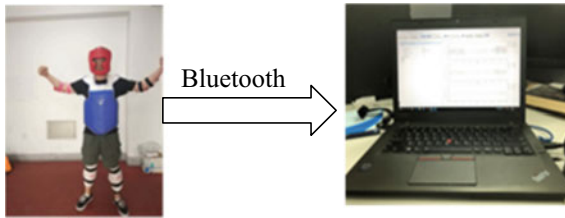


Fig. 1. The process of data acquisition

The sensor sampled data at a sampling rate of 50 Hz and included tri-axial accelerometer and tri-axial gyroscope signals. In the process of data acquisition, we recorded video to remove nonexperimental data to establish a database of bullying. In addition, the raw data was smoothed by using wavelet filtering because it contained random noise and jitter noise. The y-axis of the tri-axial accelerometer and the gyroscope is the vertical direction, and the x-axis and the z-axis are the horizontal directions. For the acceleration, the x-axis and the z-axis are combined in the horizontal direction, i.e.,

$$ACC_{Hori} = \sqrt{(ACC_{x-axis}(i)^2 + ACC_{z-axis}(i)^2)} \quad (1)$$

For the gyroscope data, it is combined as follows:

$$Gyro = \sqrt{(Gyro_{x-axis}(i)^2 + Gyro_{y-axis}(i)^2 + Gyro_{z-axis}(i)^2)} \quad (2)$$

After the preprocessing of these data, these signals are segmented into windows with a 50% overlap between two consecutive windows. The hamming window was adopted and the length of the sliding window was set to be 256 points. As the sensor

sampling rate is 50 Hz, the corresponding time length of the data in each window is about 5 s, and an action can be completed according to the actual observation. In this case, every decision made on the activity is for the duration of the window.

3.2 Feature Extraction

After data preprocessing, a large set of features were extracted from the each window for the *Acc_{y-axis}*, *Acc_{Hori}*, *Gyro*. First, feature extraction is performed within each window of the data in the time domain, and then FFT is carried out to extract the features in the frequency domain.

Table 1 lists the extracted features from each window. These features were based on previous PA research [14–16].

Table 1. List of features extracted

No	Features	Count
1	Mean	3
2	Median difference	3
3	Maximum value	3
4	Minimum value	3
5	Differential maximum	3
6	Differential mean	3
7	Standard deviation	3
8	Kurtosis	3
9	Skewness	3
10	Cross-correlation of ACC axis	3
11	25th percentile	3
12	75th percentile	3
13	Zero-crossing rate of ACC axis	3
14	Energy of frequency domain	3
15	Mean of frequency domain	3
16	Median difference of frequency	3
17	Maximum value of frequency	3
	Total number of features extracted	51

3.3 Normalization and Feature Selection

Data normalization processing is a basic work of data mining. Different evaluation indicators often have different dimensions and dimension units. In order to eliminate the dimensional impact among different indicators, we need to standardize data to solve the data comparability of indicators. Normalization is required to limit the feature in a certain range. The Z-score normalization was applied in this paper. For example, feature y can be normalized by the following formula,

$$\hat{y}_i = \frac{y_i - \bar{y}}{\sigma} \quad (3)$$

where \bar{y} and σ are mean and standard deviation, respectively.

Feature selection algorithm is one of the most critical steps that affect the recognition accuracy of the entire system. Selecting good features can increase the recognition accuracy, but useless features can cause over-fitting and lower recognition accuracy, and increase algorithm complexity. Therefore, it is necessary to select the best feature subsets for all the extracted features to classify. The filter-based Relief-F feature selection algorithm is applied to select the most useful feature set in this study. This selection algorithm is efficient, simple, and feature selection can be performed without building a classifier. In this article, the data were used to calculate the correlations between feature and each labeled activity, and then a threshold was set to select feature subsets with greater correlation for training and testing.

3.4 Classification Algorithms

The design of the classifier needs to consider practical applications and classification performance. Research work shows that the BP neural network has very limited learning speed due to its global approaching nature, and it has great limitations in some cases with high real-time requirements. The Convolutional Neural Network (CNN) based on deep learning has the advantages of no need to manually select features and excellent classification performance, but training requires a large amount of sample data and high-performance equipment. Compared with BP and CNN, RBF has the advantages of simple structure, fast convergence, and ability to approximate any nonlinear function. The generalized RBF classifier was adopted based on this project for real-time and data resources. According to the analysis of experimental results (see Sect. 4 for details), the RBF has a very high classification performance and can accurately distinguish between school bullying activity and daily-life activity.

4 Classification Performance

4.1 Evaluation Approach and Indicators

Five-fold cross-validation is used to validate the classifier performance. This method ensures that all data will undergo training and testing, and it can effectively avoid the occurrence of over-fitting or under-fitting. It can be more persuasive and practical.

The confusion matrix is used to perform a contingency table analysis of the true class and the predicted class. For each instance, the number of true-negatives (TN), false-negatives (FN), false-positives (FP), and true-positives (TP) are calculated by comparing the truth labels and the predictions. The performance of the system can be evaluated by several parameters such as accuracy, precision, recall, and F1-score. The formulas for these parameters are calculated as follows:

$$accuracy = \frac{TP + TN}{P + N} \tag{4}$$

$$precision = \frac{TP}{TP + FP} \tag{5}$$

$$recall = \frac{TP}{TP + FN} \tag{6}$$

The F1-score is used to comprehensively consider the accuracy and recall rate, and it can be calculated by

$$F_1\text{-Score} = \frac{2}{\frac{1}{precision} + \frac{1}{recall}} \tag{7}$$

4.2 Classification Results

The activity recognition conversion matrix is shown in Table 2. The first column of the table represents the actual activity. The figures in the table represent the probability that the actual activity is recognized as the activity in the first row, and the diagonal indicates the accuracy of the activity recognition. Finally, statistics are divided into violent activities and nonviolence activities. The “push”, “push-down”, and “beat” are classified as violent activities, while other activities are classified as nonviolent activities. Violent activity recognition confusion matrix is shown in Table 3.

Table 2. Conversion ratio of physical activity recognition (unit: %)

	Beat	Jump	Play	Push	Run	Stand	Walk	Push-down	Fall-down
Beat	76.3	0.0	3.0	5.9	0.0	2.2	4.4	5.9	2.3
Jump	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Play	1.1	0.0	90.0	3.3	0.0	0.0	5.6	0.0	0.0
Push	15.7	0.0	2.9	71.4	0.0	1.4	8.6	0.0	0.0
Run	0.0	0.0	0.0	0.0	99.1	0.0	0.0	0.9	0.0
Stand	2.0	0.0	3.3	0.0	0.0	94.0	0.7	0.0	0.0
Walk	0.0	0.0	1.4	2.1	0.0	0.0	96.5	0.0	0.0
Push-down	15.6	0.0	4.4	8.9	0.0	4.4	0.0	55.6	11.1
Fall-down	12.0	0.0	2.0	6.0	0.0	6.00	4.0	26.0	44.0

The evaluation index can be calculated from the PA recognition confusion matrix. The recognition accuracy is about 90% and the recall reaches approximately 84%, in

Table 3. Confusion matrix of PA recognition (unit: %)

	Violence	Nonviolence
Violence (actual)	83.8	16.2
Nonviolence (actual)	7.2	92.8

addition, the precision is 82% and the F1-score is equal to 83%. The experimental results show that we can distinguish between violent activities and nonviolent activities accurately.

5 Conclusion

School bullying has a great influence on the physical and mental health of teenagers. This paper proposes an active violence detection algorithm which analyzes the accelerometer and gyro data collected by sensors. After features extraction and selection and classifier design, we can distinguish between violent activities and daily-life activities. Totally nine types of activities were tested, and the average recognition ratio was approximately 90%. This method can be applied to students' wearable devices with built-in sensors to better protect students and prevent violence in schools.

Acknowledgements. This work was supported by the National Natural Science Foundation of China (61602127), and partly supported by the Directorate General of Higher Education, Indonesia (2142/E4.4/K/2013), and the Finnish Cultural Foundation, North Ostrobothnia Regional Fund. The authors would like to thank those people who have helped with these experiments.

References

1. Gupta, P., Dallas, T.: Feature selection and activity recognition system using a single triaxial accelerometer. *IEEE Trans. Biomed. Eng.* **61**(6), 1780 (2014)
2. Politi, O., Mporas, I., Megalooikonomou, V.: Human motion detection in daily activity tasks using wearable sensors. In: *Signal Processing Conference*, pp. 2315–2319. IEEE (2014)
3. Altini, M., Vullers, R., Van Hoof, C., et al.: Self-calibration of walking speed estimations using smartphone sensors. In: *IEEE International Conference on Pervasive Computing and Communication Workshops*, pp. 10–18. IEEE Computer Society (2014)
4. San Buenaventura, C.V., Tiglao.: Basic human activity recognition based on sensor fusion in smartphones. In: *2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*
5. Wang, Z., Huo, Y.: A multi-attribute fusion acceleration feature selection algorithm for activity recognition on smart phones. In: *International Conference on Information Science, Electronics and Electrical Engineering*, pp. 145–148. IEEE (2014)
6. Coskun, D., Incel, O.D., Ozgovde, A.: Phone position/placement detection using accelerometer: impact on activity recognition. In: *IEEE Tenth International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, pp. 1–6. IEEE (2015)

7. Ayu, M.A., Ismail, S.A., Mantoro, T., et al.: Real-time activity recognition in mobile phones based on its accelerometer data. In: International Conference on Informatics and Computing, pp. 292–297. IEEE (2017)
8. Gao, L., Bourke, A.K., Nelson, J.: A system for activity recognition using multi-sensor fusion. In: International Conference of the IEEE Engineering in Medicine & Biology Society, Embc., p. 7869 (2011)
9. Kushwah, A., Kumar, S., Hegde, R.M.: Multi-sensor data fusion methods for indoor activity recognition using temporal evidence theory. *Pervasive & Mob. Comput.* **21**, 19–29 (2015)
10. Rahman, M.M., Charoenlarnnoppa, C., Suksompong, P.: Signal processing for multi-sensor E-nose system: acquisition and classification. In: International Conference on Information, Communications and Signal Processing, pp. 1–5 (2015)
11. Gao, L., Bourke, A.K., Nelson, J.: A comparison of classifiers for activity recognition using multiple accelerometer-based sensors (2012)
12. Zebin, T., Scully, P.J., Ozanyan, K.B.: Inertial sensor based modelling of human activity classes: feature extraction and multi-sensor data fusion using machine learning algorithms. In: *eHealth 360°*. Springer International Publishing (2017)
13. Chowdhury, A., Tjondronegoro, D., Chandran, V., et al.: Physical activity recognition using posterior-adapted class-based fusion of multi-accelerometers data. *IEEE J. Biomed. & Health Inform.* **PP(99)**, 1–1 (2017)
14. Gao, L., Bourke, A.K., Nelson, J.: Evaluation of accelerometer based multi-sensor versus single-sensor activity recognition systems. *Med. Eng. Phys.* **36(6)**, 779–785 (2014)
15. Pirttikangas, S., Fujinami, K., Nakajima, T.: Feature selection and activity recognition from wearable sensors. In: *Ubiquitous Computing Systems*, pp. 516–527. Springer (2006)
16. Wichit, N.: Multi-sensor data fusion model for activity detection. In: International Conference on ICT and Knowledge Engineering, pp. 54–59. IEEE (2015)



Multichannel High-Speed Data Caching System on FPGA for RAID Storage

Haixin Wang, Xue Bai, and Qiongzhi Wu (✉)

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

whxscpzh@163.com, wqz_bit@bit.edu.cn

Abstract. Channelization RAID storage system requests multichannel data transmission and high transmission bandwidth. We design a data caching system which is inserted between fore-end data source interface and the back-end RAID interface on a FPGA implementation. The caching system uses DDR3 as the external memory because of its large storage capacity and high storage rate. It uses a special channel management system and only needs three clock cycles to complete the read–write scheduling of different channels. The caching system provides the AXI4-Lite interface, so it can be dynamically configured by the AXI4-Lite bus. After testing, the caching system can satisfy the request of multichannel storage task.

Keywords: Multichannel · Caching system · DDR3 · FPGA

1 Introduction

In multichannel storage task, due to the wide dynamic range of the fore-end data interface and the uneven storage rate of the back-end RAID, the data integrity cannot be guaranteed without data caching. When data are transmitted, there will be multiple storage channels working at the same time. And the storage task usually has high transmission bandwidth. So, it is important to implement multichannel high-speed data caching system on FPGA.

The DDR3 SDRAM is a high-speed synchronous dynamic random-access memory with burst-8 transmission [1]. The DDR3 has high bandwidth, low power consumption, and large storage capacity, and it becomes the preferred scheme in high-speed and real-time data storage system design [2].

The used FPGA belongs to the Xilinx Kirtex-7 family. In this family, Memory Interface Generator (MIG [3]) core can be used to generate memory controllers and interfaces for FPGA. In [4], the paper analyzes the good adaptability and flexibility of MIG. But MIG Memory Controller can't adapt the complete transmission task. In [5], the paper uses MIG to implement a multichannel channelization data cache for Polyphaser DFT filter. It can cache high bandwidth parallel data. But it can't manage the multichannel reading and writing of different independent channels. In this paper, the caching system uses a special channel management system for scheduling different channels.

The caching system connects other module with two standard Xilinx interfaces which are AXI4-Stream and AXI4-Lite. AXI4-Lite is used for simple, low-throughput memory-mapped communication and AXI4-Stream is used for high-speed streaming data [6]. The standard interface can make this system adapt different application. In this article’s figs, AXI4-Stream is abbreviated to “axis”.

With the increment of single transmission length, the transmission efficiency of DDR3 is increasing and approaches an upper limit, which is indicated in [7]. When the single transmission length is short, the transmission efficiency is low. And when the single transmission length is long, the system doesn’t have enough scheduling clock cycles. In this paper, the system’s single transmission length of data package is a dynamic value which can be configured through AXI4-Lite interface. So, the data package length can be adjusted to a proper value in different transmission task.

2 Methods

2.1 Storage Features of RAID and DDR3

In this paper, the RAID consists of SSD [8]. Different SSDs have different storage performance. But they all have no response time, which causes steep decline of storage speed. In this paper, the no response time of the RAID is about 30 ms. If the transmission speed is high, the caching system accumulates large amounts of data which is usually up to dozens of MB in no response time. FPGA’s Rom source is not enough to cache the data. So DDR3 is chosen to be the external memory because of its high bandwidth and large storage capacity. The average write speed of RAID is 16 Gbps and read speed of RAID Controller is 19.2 Gbps. When RAID is in writing, the fore-end channels’ total transmission bandwidth can’t outnumber 16 Gps.

In this paper, DDR3 has 32 bit data width and 2G Byte capacity. Its physical clock is 600 MHz. The transmission bandwidth of the DDR3 is 38.4 Gbps. There are various delays and latent periods when DDR3 works. This physical characteristic of DDR3 will cause start-up time and stagnation time in actual work. So, the transmission speed of DDR3 is uneven. For example, in Fig. 1, after sending “app_cmd” and “app_addr”, “app_rd_data_valid” is pulled up after more than 20 clock cycles. And “app_rd_data_valid” can be suddenly pulled down during the transmission.

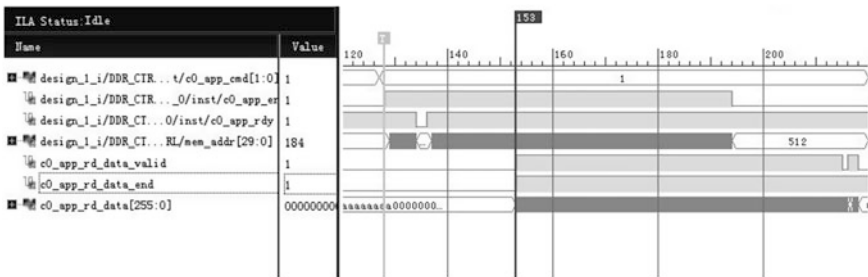


Fig. 1. Reading transmission of MIG memory controller

2.2 System Framework Design

The transmission speed of DDR3 and RAID are both uneven. So, the data caching of the systems should have two stages. Stage-1 caching is in FPGA which stores single and continuous data. Stage-2 caching is in DDR3 which stores long-length data packages. Caching in Stage-1 is called L1 caching and caching in Stage-2 is called L2 caching. Every channel has L1 caching and L2 caching.

The caching system consists of Data Stream Controller (DSC) and Channels Scheduling Module (CSM) which is shown in Fig. 2. DSC wraps MIG Memory Controller. It converts complete UI into simple AXI4-Stream interface. CSM manages data transmission of channels. It controls each channel to communicate with MIG Memory Controller, which is based on time division multiplexing.

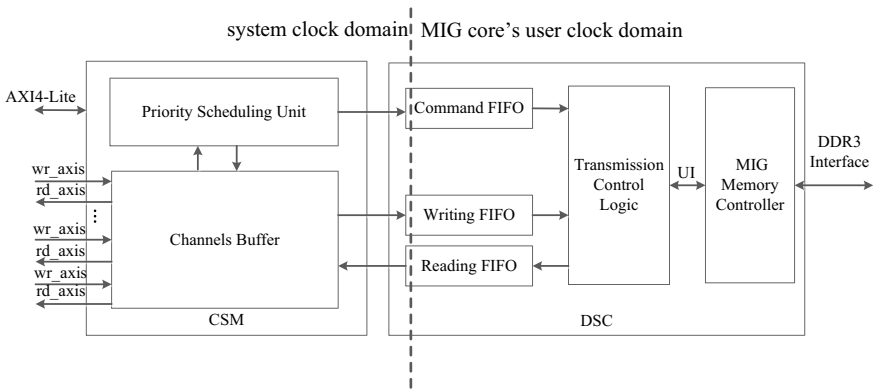


Fig. 2. FPGA logic of the multichannel data caching system

CSM and DSC work in different clock domains. CSM works in system clock domain. DSC works in MIG core's user clock domain, and the clock frequency is higher than system clock. The communication between the two modules is through asynchronous FIFO which isolates clock domains.

There are $2 * N$ AXI4-Stream interfaces connecting the N fore-end data channels and two AXI4-Stream interfaces connecting the back-end RAID interface for reading and writing data. In addition, there is an AXI4-Lite interface for accessing its function register. And the ports of DDR3 interface are constrained to FPGA's special IO ports.

2.3 UI of MIG

UI is MIG Memory Controller's user interface which is shown in Fig. 3. In the transmissions of UI, written data, read data, command, and address are separated into different FIFOs. The address indicates the location of the burst transmission. The command indicates the transmission is written data or read data. When the "app_en" and "app_rdy"

are asserted, “app_cmd” and “app_addr” can be written to the FIFOs. Data can be pushed into “app_wdf_data” when “app_wdf_wren” is asserted and “app_wdf_rdy” is high. Data can be popped from “app_rd_data” when “app_rd_data_valid” is high.

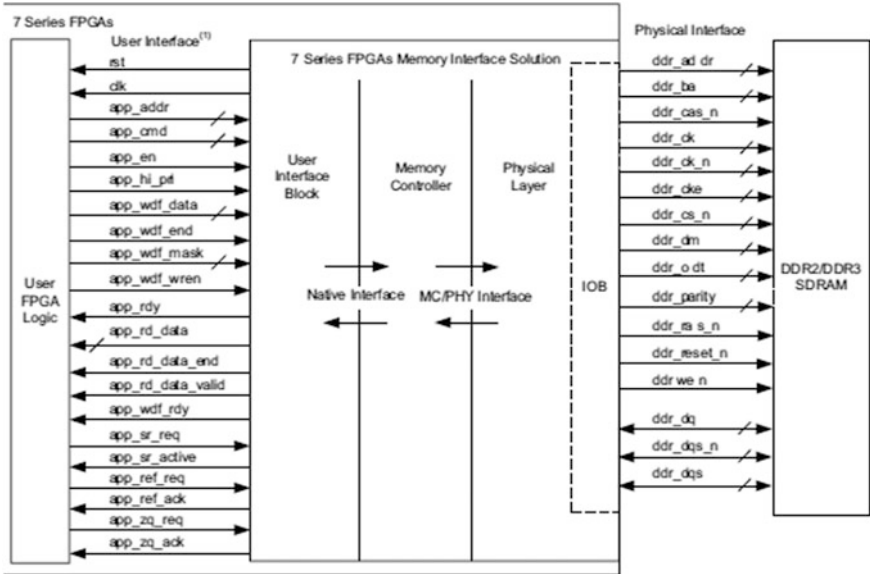


Fig. 3. Logic of MIG memory controller

2.4 Data Stream Controller

In this paper, Data Stream Controller (DSC) wraps MIG Memory Controller and uses asynchronous FIFO to isolate clock domains. It’s shown in Fig. 4. UI’s clock is asynchronous to system clock and the UI is a complete interface. So, it’s necessary to avoid MIG Memory Controller being connected with other modules directly. There are three asynchronous FIFOs including reading FIFO, writing FIFO and command FIFO whose interface is AXI4-Stream. In DSC, there is a finite-state machine in Transmission Control Logic to control data transmission between asynchronous FIFOs and UI. The state transition of the finite-state machine is shown in Fig. 6.

2.5 Channels Scheduling Management

Channels Scheduling Management (CSM) consists of Channels Buffer and Priority Scheduling Unit. The FPGA logic of CSM is shown in Fig. 5. CSM converts a multichannel transmission into a single channel transmission to connect DSC. L1 caching is in Channels Buffer. Priority Scheduling Unit determines the channel of the transmission according to L1 caching depth and L2 caching depth. The single transmission length of L1 caching and L2 caching can be configured by AXI4-Lite.

As shown in Fig. 7, Channels Buffer has $(N + 1)$ groups of FIFOs. Each group includes a read FIFO and a write FIFO. Every FIFO has an AXI4-Stream interface. N groups of FIFOs store the data of fore-end channels. The last one group caches the data of RAID. Every FIFO reports depth of cached data to Priority Scheduling Unit. A switch is inserted between DSC and Channels Buffer. The switch selects different channels based on the “ch_sel” signal of Priority Scheduling Unit.

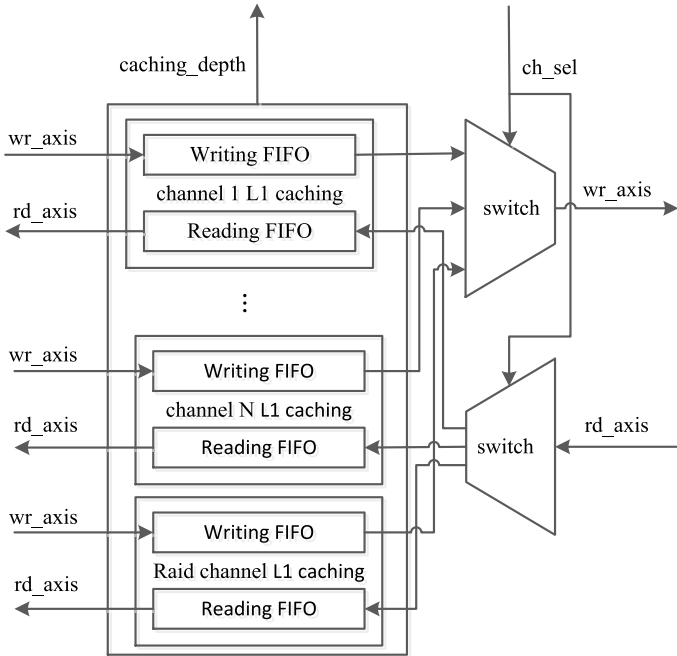


Fig. 7. Logic of channels buffer

The FPGA logic of Priority Scheduling Unit is shown in Fig. 8. The Arbitration Logic gets L1 caching depth and L2 caching depth and chooses a top priority channel. Then the Command Generation Logic generates a corresponding command which includes the L2 caching’s head address and transmission length. The first transmission address is provided by L2 Caching’s Address Management Logic and the transmission length can be configured by AXI4-Lite. The length is named package length. L2 Caching’s Address Management Logic can allocate storage space for every channel by AXI4-Lite and update the space when a channel has data package transmission.

There are four kinds of tasks in data package transmission of every channel, we call them M_1 , M_2 , M_3 , M_4 which is shown in Table 1. The priority of the four tasks is $M_1 > M_2 > M_3 > M_4$.

Each channel has a 4-bit vector in Arbitration Logic. Each bit of the vector presents a task is valid or not. N channels’ vectors constitute an $N * 4$ Mission Table. For example, in Table 2, there are three channels and the vectors include “1100”, “0100”, and “1010”. There is a three-step arbitration method to choose a top priority channel.

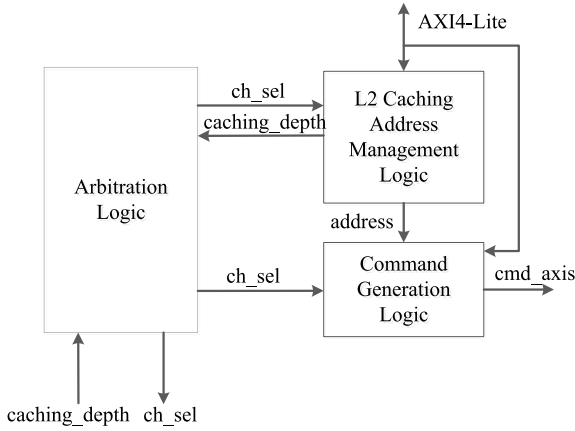


Fig. 8. Logic of priority scheduling unit

Table 1. Four kinds of package transmission

	M1	M2	M3	M4
Write/Read	Write package	Read package	Read package	Write package
Trigger Situation	L1 caching's writing FIFO is almost full, L2 caching's space lasts more than one package	L1 caching's reading FIFO is almost empty, L2 caching has more than one package data	L1 caching's reading FIFO lasts more than one package, L2 caching has more than one package data	L1 caching's writing FIFO has one package data, L2 caching space lasts more than one package

Table 2. A three-channel vectors' table

	M1	M2	M3	M4
Channel 1	1	0	0	1
Channel 2	0	0	0	1
Channel 3	0	1	1	1

- Step 1: Do an N-bit “OR” calculate for every column in Mission Table to get a 4-bit Priority Vector. For example, in Table 2, we get “1111”.
- Step 2: Input Priority Vector to a priority decoder to select a task M x. For example, in Table 2, we get M1.
- Step 3: Taking the N-bit vector of x-th column as the input, a channel number is selected through a priority decoder. For example, in Table 2, we get channel 1 as the top priority channel through the first column vector “100”.

Through the above three steps, a top priority channel can be selected by only three clock cycles. Time complexity will not increase proportionally as the number of channels increases.

2.6 The Verification System

The verification system is shown in Fig. 9. Data Generation is a three-channel data source which can be the fore-end data source. It can generate or receive incrementing number. RAID Controller can be controlled by Microblaze [9] to send or receive caching data from Multichannel Data Caching System.

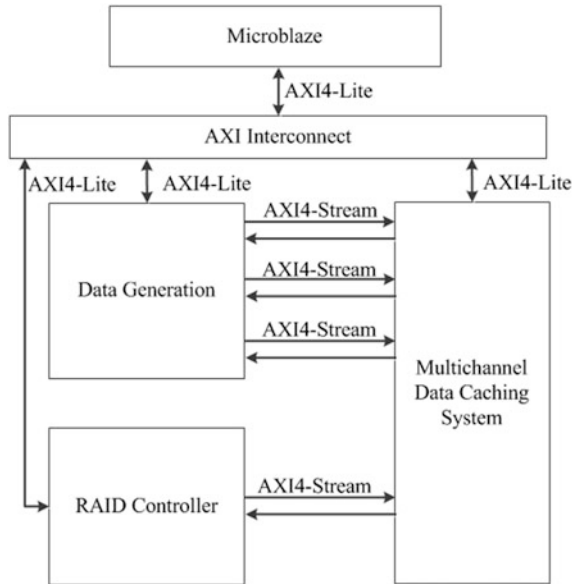


Fig. 9. Verification system

When RAID Controller imports data, DDR3 is switching in writing and reading. Generally, B_{ch} should be less than $B_{ddr3} * E_{ddr3} * 50\%$. B_{ddr3} represents the transmission bandwidth of the DDR3. B_{ch} represents the transmission bandwidth of the fore-end data channels. E_{ddr3} represents transmission efficiency of DDR3.

The transmission speed of every channel in Data Generation can be adjusted. The maximum amount of accumulated data in RAID's no response time is about 54 MB. So, choosing 64 MB as the single transmission length of L2 caching can satisfy caching performance. The [7] indicates that when single transmission length exceeds 1 KB, the transmission efficiency of DDR3 is approach up limit. So, 16 KB can be chosen to be single transmission length of L1 caching. The transmission bandwidth of RAID Controller is higher than Data Generation. The transmission speed should be limited when RAID Controller exports data.

3 Results

The verification flow is executed by the program in verification system's Microblaze. The verification method is that turn on incrementing number of three channels at the same time and receive the data with RAID Controller in Fig. 9. After receiving, the RAID Controller sends the stored incrementing number and Data Generation receives the data then checks data's correctness. The whole process can be controlled by Microblaze. After repeated testing, the result is shown in Table 3.

Table 3. Result of verification test

Channels' total bandwidth (Gbps)	16.0	16.8	17.6	18.4	19.2
Efficiency of DDR3 (%)	83.3	87.5	91.6	–	–
Verification result	No error	No error	No error	Error	Error

The result shows that the top caching bandwidth of this system is about 17.6 Gbps and the top efficiency of DDR3 is about 91.6%. The transmission rate of FPGA can be accelerated by increasing ports' I/O width or increasing clock frequency. The caching bandwidth's upper limit is the speed of DDR3.

4 Conclusion

This paper presents a multichannel data caching system which has high transmission bandwidth and multichannel scheduling management. The system has configurable data transmission package length and standard Xilinx interfaces. It resolves the requirements of different application of data caching for RAID storage. It provides a reference value for real-time storage on FPGA.

References

1. JEDEC Solid State Technology Association: DDR3 SDRAM Specification (JESD79-3A). In: JEDEC Standard. September, 2007
2. Wang, B., Du, J., Bi, X., Tian, X.: High bandwidth memory interface design based on DDR3 SDRAM and FPGA. In: Soc Design Conference, pp. 253–254. IEEE (2016)
3. Xilinx. UG586: Zynq-7000 AP SoC and 7 Series Devices Memory Interface Solutions v4.2 User Guide. October 04, 2017
4. Jiao, S., Cheng, R.: Design of a DDR3 controller based on FPGA. *Electron. Sci. Technol.* **28** (7), 41–43 (2015)
5. Ye, W., Li, H.: FPGA based DDR3 applications in a multichannel channelization data cache. In: International Symposium on Computational Intelligence and Design, pp. 54–57. IEEE (2017)
6. Xilinx. UG761: AXI Reference Guide v14.3. November 15, 2012
7. Ming, S., Zhao, Y., Lin, Q.: Design and optimization of DDR3 SDRAM controller based on FPGA. *Electron. Sci. Technol.* **29**(11), 47–50 (2016)

8. Liu, S., Aseffa, D., Wu, C.: Performance considerations for writing data to solid-state drives. *ICIC Express Lett.* **10**(3), 691–697 (2016)
9. Xilinx. WP469: Using the MicroBlaze Processor to Accelerate Cost-Sensitive Embedded System Development. June 06, 2016



Research on Multiple Switched Flat-Top Beam Smart Antenna

Wei Liang, Xiuzhen Luan^(✉), and Kejun Tan

School of Information Science & Technology, Dalian Maritime University,
Dalian 116026, China
xiuzhenluan@dlmu.edu.cn

Abstract. A novel method of multiple switched flat-top beam is proposed for solving the problem of unequal radiation in different directions of the same beam in the multiple switched beam smart antenna in this paper. Taking 4-element Quasi-Yagi antenna array for an example, the feed currents of 4-element flat-top beam antenna array are calculated by using Woodward-Rosen synthesis method. An improved 4×4 Butler Matrix feed network is designed by using the double-layer dielectric structure. The antenna array is simulated by HFSS simulation software. The simulated results show that, compared with the beams of uniform antenna array, the gain fluctuations of the flat-top beam antenna array are reduced by 6.4–25.9 dB in the main lobe, the sidelobe levels are reduced by 2.7–17 dB, and with the characteristics of small size and wide bandwidth.

Keywords: Multibeam · Flat-top beam · Smart antenna · Butler Matrix

1 Introduction

With the development of mobile communication, the number of mobile users increases rapidly, and the contradiction between the growth of mobile communication traffic and the limited spectrum resource becomes increasingly prominent. One effective way to solve this contradiction is to improve the effective utilization of spectrum. Smart antenna, as a technology to improve spectrum efficiency, has attracted wide attention in recent years. Smart antennas are divided into two types: fully adaptive and multiple switched beams. Multiple switched beam smart antenna system offers increased capacity and is less complex compared to the fully adaptive smart antenna, so it has strong practicability [1].

In multiple switched beam smart antenna, a set of narrow beams with different beam directions are set in advance at the receiving or transmitting end, the multiple beams are used to cover the entire user area, and each beam direction is fixed, as shown in Fig. 1a. In Fig. 1a, each beam is usually radiated by an uniform array, the intensity of the signal is different when the user is in the different directions of the same beam. When the user is in the central direction of the beam, the received signal is the strongest; when the user is at the edge of the beam, the received signal is the weakest. Therefore, the signal received by the user is unstable while the user is in moving. In order to solve this problem, a novel method of multiple switched flat-top beam smart antenna is presented in this paper, in which, every beam is designed as a flat-top beam.

The radiation pattern of the flat-top beam antenna array is sector shape in polar coordinates, as shown in Fig. 1b. The radiation is the same in all directions of the same sector region, so when the user moves in the area, the received signal is very stable. The design and implementation methods of multiple switched flat-top beam smart antenna are studied in this paper.

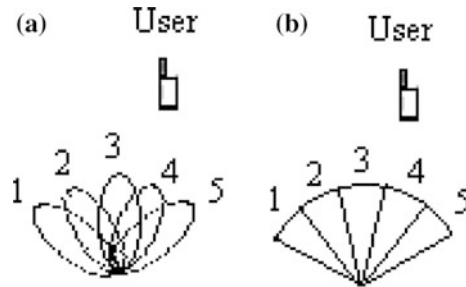


Fig. 1. The radiation patterns of multiple switched beam antenna

2 Design and Simulation

The design and implementation methods of multiple switched flat-top beam as shown in Fig. 1b are studied by taking a 4-element Quasi-Yagi antenna array for example.

2.1 Synthesis of Flat-Top Beam Antenna Array

Woodward-Rosson synthesis method is used in synthesis of flat-top beam antenna array in this paper. For a flat-top beam with a main lobe range of $(\theta_1-\theta_2)$, the desired pattern function can be expressed as follow [2]:

$$F_a(\theta) = \begin{cases} 1 & (\theta_1 < \theta < \theta_2) \\ 0 & (\text{others}) \end{cases} \tag{1}$$

The normalized feed current amplitudes of 4-element flat-top beam antenna array with 40° main lobe width are obtained and shown in Table 1.

Table 1. The normalized feed current amplitudes of 4-element flat-top beam antenna array

No. of antenna element	1	2	3	4
Normalized current amplitudes	0.41	1	1	0.41

2.2 Design of Feed Network

According to the principle of phased array, by controlling the phases of the feed currents, the antenna arrays with different beam pointing can be obtained. Butler Matrix

is often used as the feed network of multiple switched beam smart antenna. Traditional 4×4 Butler Matrix consists of four 3 dB couplers, two 45° phase shifters and two cross couplers, as shown in Fig. 2a [3]. When the signal is input from one of the input ports 1–4, the signals will output at four output ports 5–8 with equal amplitude and equal-difference phase shift. Table 1 shows that the normalized current excitation amplitudes of the antenna elements in a flat-top beam antenna array are different, therefore, in order to realize the multiple flat-top beam smart antenna, improvement to traditional Butler Matrix is needed. The improved Butler Matrix is shown in Fig. 2b, it consists of two 3 dB couplers, two 45° phase shifters, four unequal power dividers, four 90° phase shifters, four power combiners, and seven cross couplers.

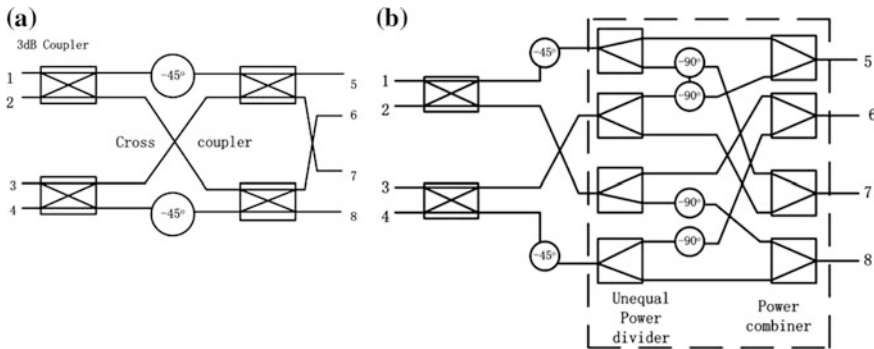


Fig. 2. 4×4 Butler Matrix topologies

Typical implementation method of Butler Matrix is: 3 dB coupler is implemented by microstrip dual branch directional coupler; phase shifter is implemented by extended transmission line section; cross coupler is implemented by microstrip three-branch coupler; power divider/combiner is implemented by Wilkinson divider [3]. The principle of this structure is simple, but it's size is large, and the frequency bandwidth is narrow. In Ref. [4], a class of compact and wideband couplers were presented, which exploited broadside coupling between top and bottom elliptical microstrip patches via an elliptical slot located in the midground plane layer. Broadside coupling phase shifter can be implemented by making two ports of the four-port device open circuit, and the other two ports are the input and output ports [5]. They are both with characteristics of small size and wide bandwidth, so are used for design in this paper. Figure 3 shows the structure diagram of the improved microstrip 4×4 Butler Matrix designed in this paper. This is a two-layer dielectric substrate configuration, the relative permittivity of the dielectric substrate is 4.4, the thickness is 0.8 mm, the two layers dielectric substrate are separated by a slotted ground plane layer. Four input ports 1–4 are located on two sides of the dielectric substrate, respectively. In order to realize the current amplitude ratio of 0.41:1:1:0.41, the improved unequal Wilkinson power divider structure is adopted. In which, one of branches is realized by parallel coupling line to solve the problem of the microstrip line width too narrow. Because the phase shift quantity of the broadside coupling phase shifter is limited, the 90° phase shifter is

realized by cascading two 45° phase shifters. The cross couplers are realized by the coupling between top and bottom elliptical microstrip patches via an elliptical slot located in the mid ground plane layer. Finally, four Wilkinson dividers are used for power combiners, signal output from four ports 5–8 on two sides of the dielectric substrate, respectively. The improved Butler Matrix works in the 2.45 GHz band.

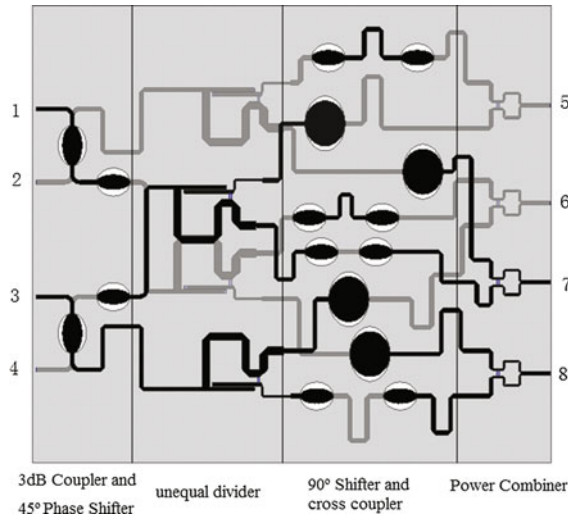


Fig. 3. Structure of the improved Butler Matrix

The simulation results of the improved Butler Matrix at the center frequency are shown in Table 2. It can be seen from Table 2 that when the signals are input from Port 1 to Port 4, respectively, the phase differences of output signals at output port 5–8 are close to the ideal results of 45°, -135°, 135°, and -45°, and the output signal amplitude ratio is close to the expected ratio of 0.41:1:1:0.41.

Table 2. Simulation results of the improved Butler Matrix

Input port number <i>i</i>	$ s_{ii} /\text{dB}$	$ s_{5i} /\text{dB}$	$\angle s_{5i}/\text{deg}$	$ s_{6i} /\text{dB}$	$\angle s_{6i}/\text{deg}$	$ s_{7i} /\text{dB}$	$\angle s_{7i}/\text{deg}$	$ s_{8i} /\text{dB}$	$\angle s_{8i}/\text{deg}$
1	-23	-17.2	41.2	-9.6	-5.3	-10.4	-54.1	-18.1	-95.9
2	-29	-17.5	-48.5	-10.5	83.4	-10.4	-139.2	-18.1	-6.5
3	-37	-17.7	37.6	-10.5	-95.4	-9.8	127.8	-18.0	-5.6
4	-23	-17.8	-52.4	-10.0	-9.8	-9.8	37.9	-18.0	82.9

2.3 Design and Simulation of Quasi-Yagi Antenna Array

In this paper, a 4-element microstrip Quasi-Yagi antenna is designed as an antenna array element, it is fabricated on the dielectric substrate with permittivity $\epsilon_r = 4.4$, the constituted 4-element antenna array is shown in Fig. 4. The radiation patterns of the 4-element antenna array using the feed network shown in Fig. 3 are shown in Fig. 5. Because the radiation patterns are symmetrical about the direction of 0° when the output phase differences are 45° , -45° and 135° , -135° , respectively, so only the radiation patterns when the phase differences equal to 45° and 135° are given in Fig. 5. Figure 5a shows the radiation patterns when the signal is input from the port 1 and the phase difference approximately equals to 45° at the output ports of the improved Butler Matrix. The radiation pattern of antenna array with uniform feed is also given in Fig. 5a for comparison. It can be seen from Fig. 5a that the maximum radiation directions of antenna array are both at 101° . In the main lobe range of 81° – 121° , the gains of the flat-top beam antenna array change from -3.7 to 8.3 dB, its sidelobe level is about -21 dB; but the gains of the uniform antenna array change from -9.3 to 9.1 dB, and its sidelobe level is about -4 dB.

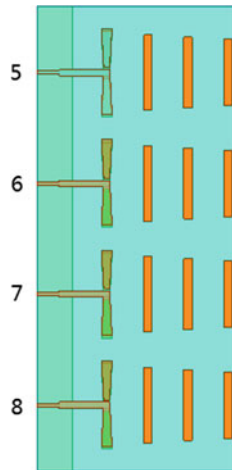
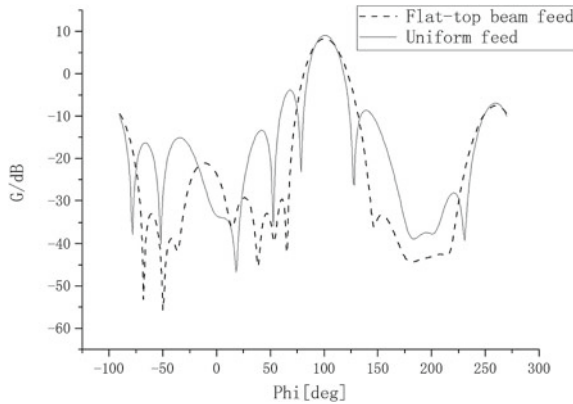
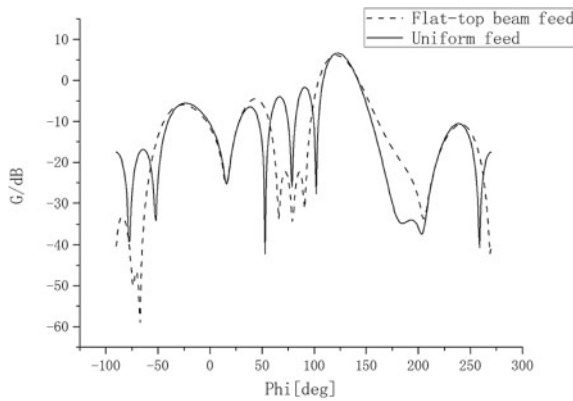


Fig. 4. Quasi-Yagi antenna array

Figure 5b shows the radiation pattern when the signal is input from the port 3 and the phase difference approximately equals to 135° at the output ports of the improved Butler Matrix. The radiation pattern of antenna array with uniform feed is also given in Fig. 5b for comparison. It can be seen from Fig. 5b that the maximum radiation directions of antenna array are both at 122° . In the main lobe range of 102° – 142° , the gains of the flat-top beam antenna array change from -2.3 to 6.1 dB, and its sidelobe level is about -4.5 dB; but the gains of the uniform antenna array change from -27.7 to 6.6 dB and its sidelobe level is about -1.8 dB.



(a) The radiation patterns when the phase difference equals to 45°



(b) The radiation patterns when the phase difference equals to 135°

Fig. 5. The radiation patterns of flat-top beam array and uniform array

Therefore, it can be seen that, under the same conditions, compared with uniform array, gain fluctuations of flat-top beam antenna array in main lobe are reduced by 6.4–25.9 dB, and the sidelobe level are reduced by 2.7–17 dB, the performances of the antenna array are improved greatly.

Figure 6 shows the radiation patterns of the multiple switched flat-top beam Quasi-Yagi antenna array when the four ports are input, separately. It can be seen from Fig. 6 that when input port switching between the port 1, port 2, port 3 and port 4, the antenna can cover the wide range of about 41° – 142° . In this range, the radiation intensity of the antenna array remains stable, thus the user can receive stable signals during movement.

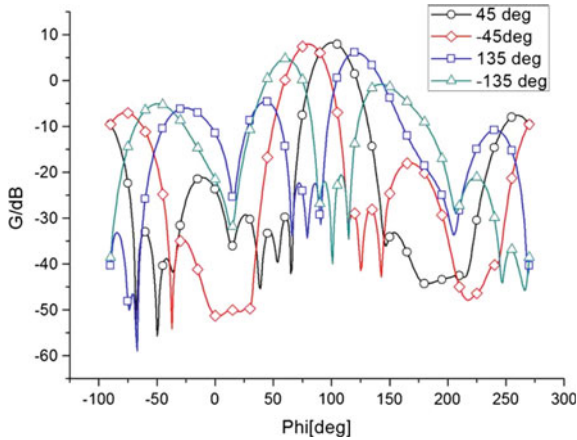


Fig. 6. The radiation patterns of the flat-top beam array when the four ports are input separately

3 Conclusion

The design and implementation methods of multiple switched flat-top beam smart antenna are presented by taking a 4-element array as an example in this paper. The results show that, compared with uniform array beam, flat-top beam has the advantages of flatter main lobe and lower sidelobe. If more elements are used for designing flat-top beam array, these advantages will be more obvious.

References

1. Bobor-Oyibo, F., Foti, S.J., Smith, D.: A multiple switched beam smart antenna with beam shaping for dynamic optimisation of capacity & coverage in mobile telecommunication networks. In: 2008 8th International Symposium on Antennas, Propagation and EM Theory, pp. 356–359. Kunming (2008)
2. Patidar, H., Mahanti, G.K.: QPSO for synthesis of linear array of isotropic antennas to generate flat-top beam including multiple null placement. In: 2015 International Conference on Signal Processing and Communication (ICSC), pp. 46–50. Noida (2015)
3. Fakoukakis, F.E., Kyriacou, G.A., Sahalos, J.N.: On the design of Butler-like type matrices for low SLL multibeam antennas. In: 2012 6th European Conference on Antennas and Propagation (EUCAP), pp. 2604–2608. Prague (2012)
4. Abbosh, A.M., Bialkowski, M.E.: Design of compact directional couplers for UWB applications. *IEEE Trans. Microw. Theory Tech.* **55**(2), 189–194 (2007)
5. Abbosh, A.M.: Ultra-wideband phase shifters. *IEEE Trans. Microw. Theory Tech.* **55**(9), 1935–1941 (2007)



Design of Portable Power Supply System

Xinqiang Zhang^(✉), Jiaqi Li, Ya Tu, Changyun Ge, and Xiujie Zhao

Department of Electronic Engineering, Dalian Neusoft University of Information,
Dalian 116023, China
xinqiangzhang@aol.com

Abstract. The design of a portable multifunctional charger is presented in this paper. The charger supports two charging modes of 220 V alternating current and 12 V DC. The 12 V lead–acid battery is used to support AC 220 V output, DC 12 and 5 V output. The maximum power point tracking and battery charging curves are controlled to extend the life of the product and improve the performance of the product. Bluetooth can be connected to the mobile phone or pad to control the battery output and monitor the battery state. The design can be used in the field of home tourism.

Keywords: Portable charger · Battery charging management · MCU

1 Introduction

With the continuous development of human society, now it has entered the information society. All kinds of portable electronic devices are becoming more and more common, such as mobile phones, electronic dictionaries, MP3, notebook computers, CD machines, and a variety of small electronic games. When people go out and work, they often encounter the situation that these electronic devices have no electricity suddenly and there is no power around them. It brings some annoyance to life. In order to solve this problem, this paper designs a convenient charger that can charge electronic equipment anytime and anywhere, in order to meet the needs of field work and tourism [1].

The charger designed in this paper supports two charging modes of 220 V alternating current and 12 V vehicle direct current. Users can choose the appropriate charging mode according to the specific circumstances. Support AC 220 V, DC 12 V, DC 5 V three output modes. The output interface supports many ways such as vehicle, six corners, USB and so on. It can charge a variety of digital products such as mobile phones, PAD, video cameras, etc. Support the use of 220 V AC equipment under rated 80 W.

2 System Architecture

The system hardware includes three parts: main control panel, operation board, and inverter board. The main control board is responsible for the control of the whole charging and discharging circuit; the operation board is used for indication of power

and output information, and the inverter board is used for the transformation of 12 V DC to 220 V communication. The block diagram of the main control board system is shown in Fig. 1. The CONN1B interface connects the operating board, the operating board uses the ink screen to display the battery power, at the same time, the touch switch is used to select the port of the voltage output; the COON2 connects the DC-AC inverter, the inverter circuit uses the unipolar SPWM modulation mode, and uses the voltage and current double closed loop feedback control system; the CONN3 connection battery is used in this design. Panasonic lead–acid battery has a rated capacity of 12 V14A.

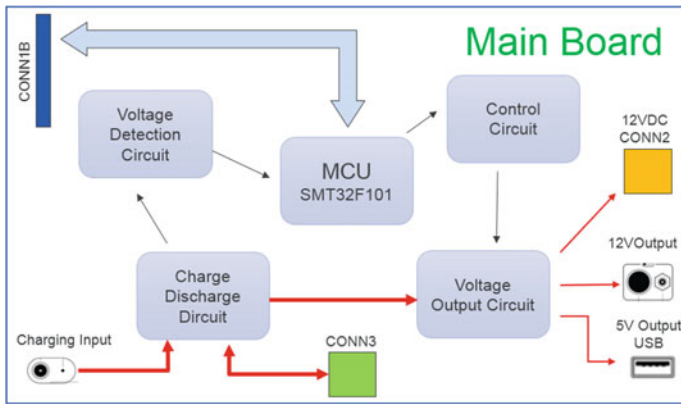


Fig. 1. Main board scheme

3 Design of Charge and Discharge Circuit

When charging, the battery can be charged by external 12V DC power supply. When the 220 V power is supplied, the system transfers the 220–12 V DC input through the external AC-DC module. The whole charging process is controlled by SMT32F101 MCU. The project adopts separate PWM controller and MOS-FET to control the whole charging process. The characteristics of the design are multiple components, various circuit, and better performance [2]. The system can fulfill the charging control while completing the functions of battery detection, error handling and timing control [3]. Charging process can be controlled by MCU. According to the charging curve of Ni MH battery, after detecting the voltage of the battery, choose the charging mode. According to the state of battery voltage, the whole charging process adopts four methods: fast charging, slow charging, trickle pre-charging, and automatic termination of charging. When the voltage is less than 11.5 V, the fast charging method is used, and the constant current charging method is mainly used. The slow charge method is used when the voltage is near 13.5 V, and the constant voltage charging mode is mainly used. When the voltage reaches 13.5 V, a trickle pre-charging mode is used. Battery charging and discharging characteristic curve as shown in Figs. 2 and 3.

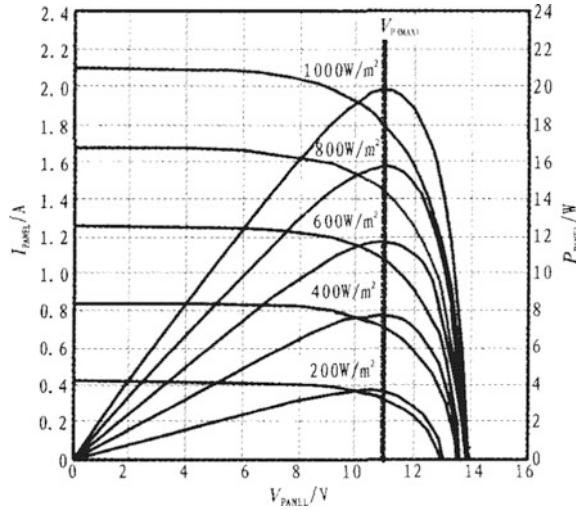


Fig. 2. Charging characteristic curve

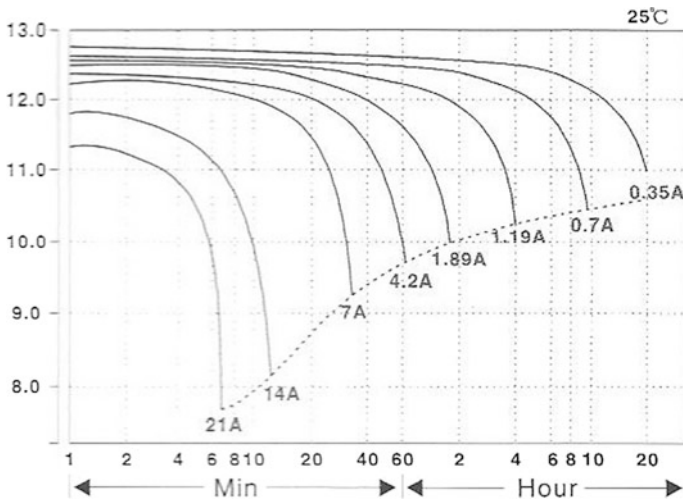


Fig. 3. Discharging characteristic curve

Products support 12/5 V DC and 220 V AC two output modes. The output voltage can be switched through the buttons on the front of panel. At the same time, there are green LED lights to show the state of the current output. The output of 12 V is controlled by a microcomputer to realize the breaking of the output circuit through the MOS tube. After sampling the current, the feedback is formed. When the output current is greater than the threshold, the output circuit is cut off by the microcomputer to realize the function of circuit protection. 5 V DC output uses adjustable step-down DC/DC switching regulator to ensure the stability of voltage output, including overheating

4 Battery Charge and Discharge Management

The PWM pulse width modulation method is used to control the charging current. STM32 microcomputer produces different modulation signals with different frequency and high level on IO3 and IO4 signals. By controlling the switch of Q1 and Q2 tubes, the purpose of controlling the charging current is achieved. Use the FTM module in STM32 to set up the PWM output mode. The outputs of the two PWM work in pairs in the same output mode. The actual charging current is connected to the ADC port of the microcontroller after amplification by detecting the current sampling resistance. After reading the charging current, the microcontroller compares with the previously set charging current. If the actual current is too small, the duty cycle of PWM will be adjusted to increase the charging current. If the actual current is larger than the set value, the charging current will be reduced.

When the battery is discharged, different design circuits are used according to different output. The DC 5 V output uses the AE2596 voltage regulator module. The DC 12 V of the battery is connected to the input port of 2596 after passing through the rectifier diode. The output voltage is passed through the filter circuit for peripheral use. The MCU controls the output switch through I/O port. The DC 12 V output circuit is similar to the charge circuit control. The control circuit is connected to the negative pole of the loop. The output of the dual MOS transistor can be enhanced by the transistor, and the output capability of the current can be improved. The realization of AC 220 V output circuit uses a full bridge inverter circuit. The whole inverter circuit also includes high-frequency step-up transformer, periodic wave transform circuit and output LC filter circuit. SPWM inverter modulation is used to control the inverter circuit, and waveform software is generated by synchronous sampling of monopole triangle wave symmetry rule. With this sampling method, the center line of pulse width can be equally spaced, and the consistency of each pulse width in each half cycle can be ensured. It is convenient for microprocessor implementation and controllable pulse width coefficient.

5 The Realization of Mobile Phone Control

This design can communicate with the mobile phone by Bluetooth, and the mobile phone can control the switch of the output circuit through Bluetooth. At the same time, it can display the current battery status on the mobile phone in real time. The microcomputer connects the Bluetooth module through the UART serial port. Bluetooth module uses CC2540 chip to support Bluetooth 4.0 standard. SCM and Bluetooth module communicate with each other through AT command.

In the process of using Bluetooth to communicate, the mobile phone is the main device and the charger is the slave device. The communication is transmitted to each other in the form of ASCII. When communicating, the data transmitter sends a handshake signal and transfers data after receiving the feedback signal. The receiver sends a feedback signal after the data is received. Data transmission adopts full duplex mode, and data structures are sent as follows: “?;Key1:Value1; Key2:Value2...” The number of data that represents the transmission. Data is given in the form of data pairs.

Key represents the content of the transmitted data, and Value represents the specific value of the data.

At present, the mobile terminal supports Android platform, the development environment uses Android Studio 2, and SDK uses API 21. The mobile terminal can display the current battery, charging state, battery temperature, environmental humidity and voltage output status. At the same time, the output of the battery can be controlled.

6 Conclusion

A portable multifunctional charger is realized in this paper. After the design and debugging of the hardware and software, the charger can realize charging and discharging functions. It can display battery information and control charger output by mobile phone. At the same time, the common abnormal conditions will be reported to failure. After testing, the input and output indicators meet the design requirements and achieve the design goal. Due to the large number of discrete devices used in the design process, the price of the overall design scheme is relatively low, and at the same time, it is conducive to the design of the later expansion function. Because the charging voltage of this design is low (12 V), isolation measures are not considered in the design process, which is also the shortcoming of this design.

References

1. Guo, H., Hu, S.: Design of a portable field charger. *Mod. Electron. Technol.* (10), 9–11 (2012)
2. Liu, G., Mao, J.: Design of charge controller for Ni MH battery based on MCU. *J. Univ. Inf. Eng.* **9**, 379–384 (2008)
3. Guo, Q., Liu, H.: Design of high efficiency battery charging system based on current source PWM rectifier. *Power Autom. Equip.* **35**(6) (2015)
4. Zheng, Y., Li, W.: Design of portable solar power generation system. *J. Zhejiang Univ. Sci. Technol.* **22**(3) (2010)



Design of Direct Current Motor Servo Control System Based on SOPC

Tu Ya^{1,2(✉)}, Jialu Du¹, and Peng Cui³

¹ School of Marine Electrical Engineering, Dalian Maritime University, Dalian 116026, Liaoning, China

tuya@neusoft.edu.cn, dujl66@163.com

² Dalian Neusoft University of Information, Dalian 116023, Liaoning, China

³ Liaoning Vocational College Light Industry, Dalian 116100, Liaoning, China
18532450@qq.com

Abstract. Nios II is applied in direct current motor servo system as core microprocessor of FPGA system in this paper. A SOPC system is created by Qsys and Quartus II. Feedforward-feedback control algorithm is adopted and Fuzzy-PID is taken as the control algorithm of feedback. The position loop, velocity loop, and current loop control in the system improve both dynamic performance and static characteristic. Fuzzy adaptive PID control algorithm are achieved by Nios II, and simulated in MATLAB/Simulink. The whole system's analysis, synthesis, simulation and configuration are completed by Quartus II. Results show that this method improves both system responding speed, and tracking precision in DC motor servo system.

Keywords: Nios II · SOPC · Fuzzy-PID control · Feedforward control

1 Introduction

The servo control system is an important field of automatic control system, which plays an important role in modern industrial and military fields. Direct current motor servo system occupies a leading position in the servo system with its wide speed regulation, excellent starting, braking characteristics, and high positioning accuracy. With the development of semiconductor and EDA technology, the emergence of high-performance microprocessors and FPGA/CPLD, makes digital direct current motor servo system develop rapidly. However, designs of most hardware in direct current motor servo system are complicated by MCU [1]. FPGA/CPLD has the characteristics of high integration, high speed and easily modifying. It can convert board-level design to chip-level design, which greatly reduces the difficulty of circuit design, and has high application value in the control system [2]. However, using HDL language compile complex control algorithm has certain difficulties. As the development of modern SOPC (System on Programming Chip) technology, the application and design based on Nios II have new design methods and design process [3].

In this paper, the complex control algorithm is realized by Nios II processor and HDL language. It not only reduces the development cycle but also embodies the design ideas for users and applications. Finally, this SOPC system is implemented by

Quartus II and Qsys. This control system is tested on the direct current motor servo system. Results show this system hardware and software design is reasonable and control system has good control performance.

2 Control Principle and Simulation

2.1 The Principle of Motor Speed Regulation

Electric, the mechanical and the electric-mechanical speed regulation are common motor speed regulation methods. The characteristics of separately excited direct current motor are shown:

$$n = \frac{U_N}{C_e \Phi_N} - \frac{R}{C_e C_T \Phi_N^2} T_{em} \quad (1)$$

where n is rotation speed, U_N is motor armature voltage, R is the external string resistance of the armature loop, T_{em} is electromagnetic torque, C_e is electromotive force constant, C_T is torque constant, Φ_N is excitation flux. According to Formula (1) to be seen, varying the motor speed can be realized by changing the parameters such as U_N , R , and Φ_N . Speed regulation by changing R is not good enough for its economy and smoothness. The range of speed regulation is too small, by changing Φ_N . When adjusting speed by voltage, the motor mechanical characteristic is constantly, and its speed range is largely. It is easy to realize stepless speed regulation, and better smoothness. Therefore, the voltage regulation is usually adopted in the field of motor speed control system. With the emergence of new electronic power devices, the PWM control mode with full-controlled switching power components becomes the mainstream.

2.2 The Principle of Control System

The system includes feedback, velocity feedback and current feedback position loops as shown in Fig. 1. In this system, current loop can restrain the interference of internal armature current and limit maximum current for protecting the motor. Velocity loop mainly restrain velocity fluctuation and improve the ability against load disturbing. Position loop is the primary control loop to realize position tracking. Three loops working together ensure great dynamic and static characteristic of system [4].

2.3 The Model of Direct Current Motor

The parameters of direct current motor are set as rotate speed without load is 4100 rad/m, decelerated ratio is 1/160 and rating voltage is 56 V. Omitting inductance of armature and viscosity damp factor we obtain the concise transfer function of direct current servo motor of which voltage $u_a(t)$ of armature is input variable and rotate speed $\omega(t)$ of motor is output variable:

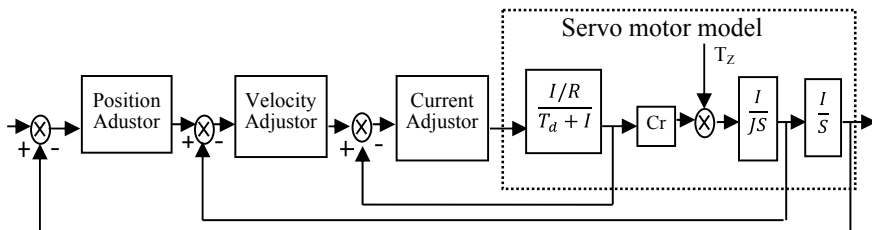


Fig. 1. System control principle diagram

$$H(s) = \frac{1/Ke}{T_m s + 1} \tag{2}$$

It is feasible to regard it as an angle sensor in view of those parameters: back electromotive force factor of motor $K_e = \frac{160 \times 60 \times 56}{4100 \times 360}$, time constant $T_m = 10$ ms, feedback factor $K_F = 15$ V/131.4°. In consideration of the transfer function that represents the relation between voltage and angle, an integral element is added to realize the transfer relation between voltage and angle.

2.4 Algorithm Design and Simulation

Fuzzy adaptive PID control algorithm is applied in this system. PID parameters are adjusted through Fuzzy reasoning in this system. The system can achieve optimal, fast response speed, high control accuracy, by adjusting the coefficients [5].

1. The Principle of Fuzzy adaptive PID Control

The Fuzzy adaptive PID algorithm has dual input and single output. We take displacement error (E) and error variety rate (Ec) as fuzzy inputs and PWM control signal (U) as output. The Fuzzy adaptive PID control algorithm is shown in Fig. 2. The discrete equation of PID is

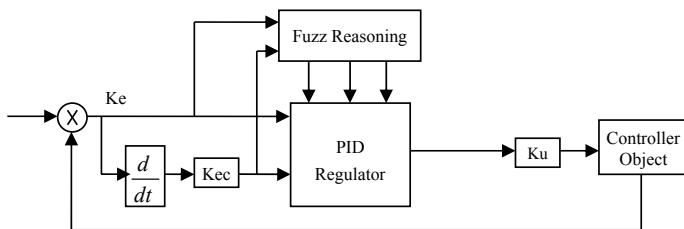


Fig. 2. The block diagram of fuzzy control

$$P(k) = K_P\{e(k) + \frac{T}{T_I} \sum_{j=0}^k e(j)T + \frac{T_D}{T}[e(k) - e(k-1)]\} \quad (3)$$

where k is sampling sequence, T is sampling period, $e(k)$ is the deviation at the time of the k -th sample, $e(k-1)$ is the deviation at the time of the $(k-1)$ -th sample, and $P(k)$ is the output of controller at the time of the k -th [6].

2. System Simulation

According to above discussions we get system model with MATLAB/Simulink. The Fuzzy adaptive PID algorithm improves the response speed and increases the system bandwidth. In this algorithm 10 V, 5 Hz sine wave tracking waveform is shown in Fig. 3.

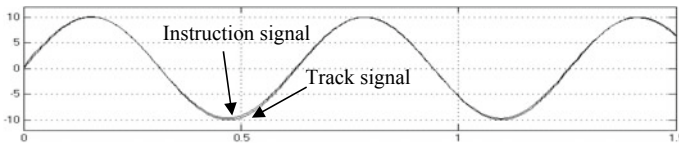


Fig. 3. Sine wave tracking with Fuzzy adaptive PID control

3 System Implementation

3.1 Hardware Design

EP4CE115F29C7 is the core controller in this system, and structure diagram is shown in Fig. 4. The position loop, current loop, and velocity loop are composed of FPGA and various external circuits. The position detection sensor outputs signal and computer demands signal are delivered to the data acquisition circuit. Position deviation and its rate of change can be obtained. Position detection sensor uses precise potentiometer with accuracy of 0.1%. The analog–digital conversion uses AD9235. The position detection sensor composes the position loop and velocity loop of the system. The current detection sensor uses Hall sensor. Its maladjustment current is less than 30 mA and respond speed is less than 10 s. The analog–digital conversion is implemented with AD1674 with conversion accuracy of 1/4096. The current detection sensor loop composes the current loop of the system [7].

In addition, the direct motor has high voltage and current, so it is important to apply isolated electric circuit 6N137 in the controlling part and motor driving part to enhance the stability of the system. The driving circuit is monopole, limited and reversible driving circuit. The PWM driving circuit is composed of NMOS IRF9150 and PMOS IRF150 group to control prograde and reversal of the motor. The motor drive circuit is shown in Fig. 5.

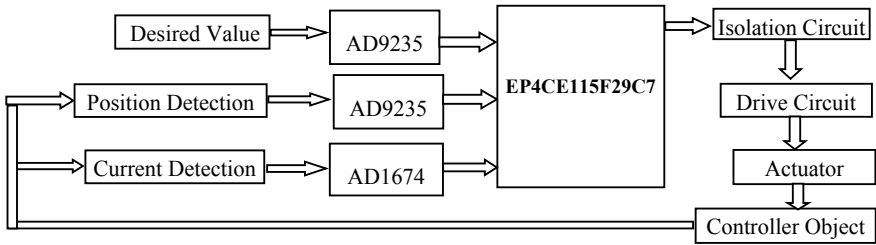


Fig. 4. The block diagram of system structure

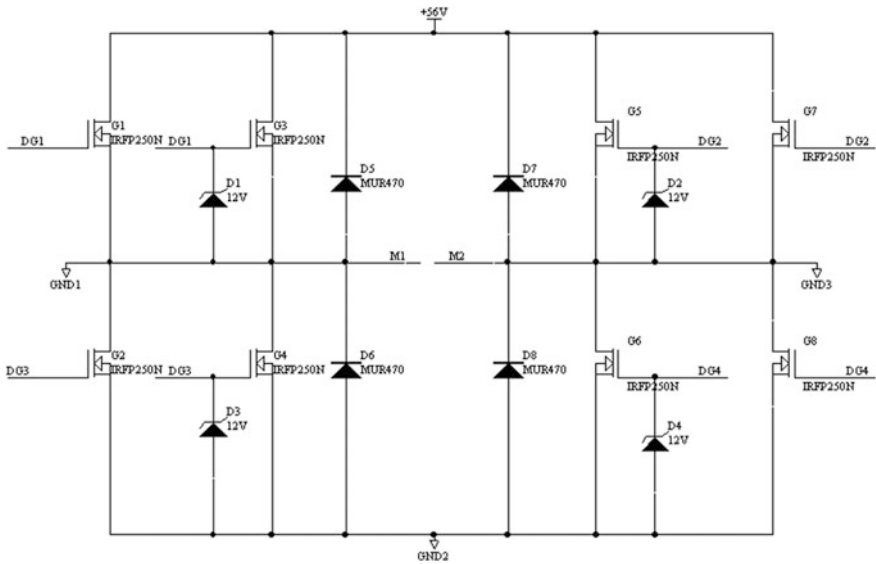


Fig. 5. The circuit of motor drive

The positive and negative rotation of the motor are controlled based on the duty ratio of two groups PWM. The two groups of PWM waves set the proper dead zone so as to avoid the current overcurrent caused by the MOSFET at the same time. The output signal of PWM wave and its enable signal are generated by FPGA. As is shown in Fig. 6, a group MOSFET are controlled by PWM wave that generated by A and C. Another group MOSFET are controlled by B and D. The dead zone time control of PWM wave in two groups is also realized by this circuit. The dead zone time can be achieved by adjusting the values of R17, C21, R19 and C22 in Fig. 6. The dead zone time is: $\tau = RC = 1\text{ K} \times 0.01\ \mu = 10\ \mu\text{s}$.

3.2 Software Design

The design of software is carried on TOP-DOWN method based on Quartus II. The whole design's analysis, synthesis, simulation, and configuration are completed by

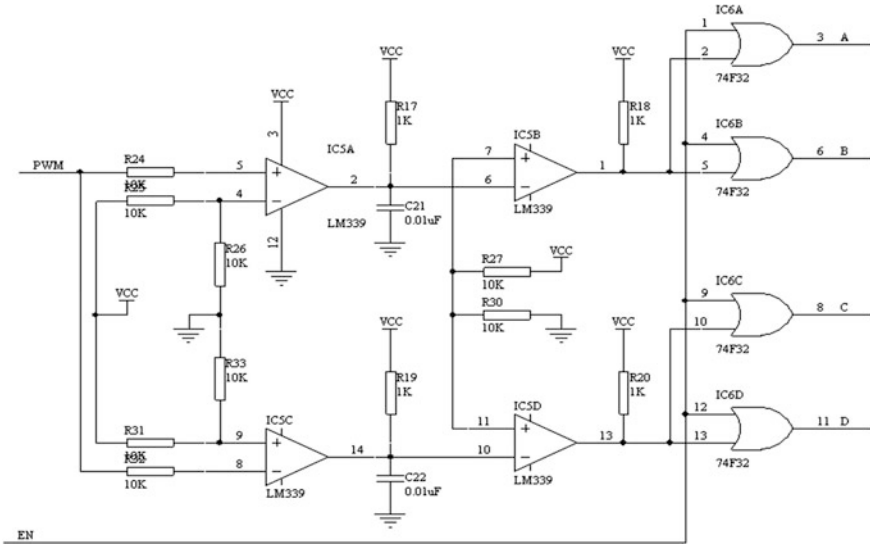


Fig. 6. The circuit of PWM generation

Quartus II. Qsys is an object-oriented development tool for SOPC, which was introduced by Altera [8]. The Nios II processor core and peripheral interfaces for this project can be built through the Qsys tool. The design of SOPC platform is shown in Fig. 7. Where AD_9235CTRL and AD_1674_CTRL are A/D converter driver modules. Control algorithm and motor control are completed in Nios II processor [9].

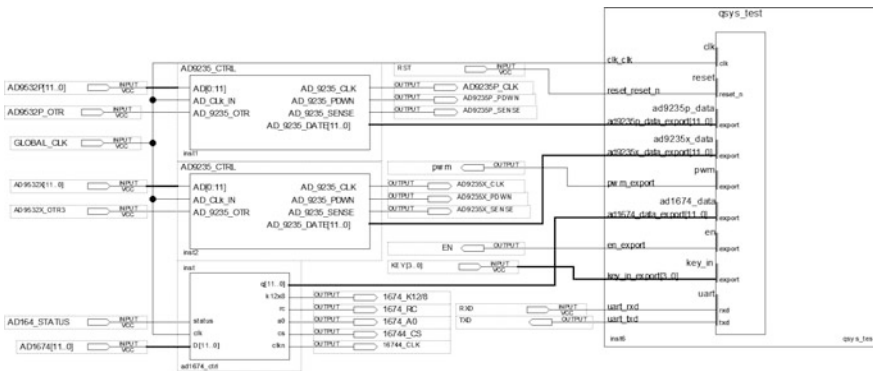


Fig. 7. The SOPC platform

3.3 System Debug

By feedforward and fuzzy-PID control algorithms, the system bandwidth and system response speed are improved, and the system response is improved obviously. The following three figures as shown in Figs. 8, 9 and 10 are some measuring results that we obtained from experiments. It can be seen that in these waveforms, the sine wave and square wave have better tracking characteristic, such as low overshoot and short adjustment time. Among these signals, the square wave has the maximum overshoot and the longest adjustment time. The overshoot of 0.5 Hz square wave with 10 V is

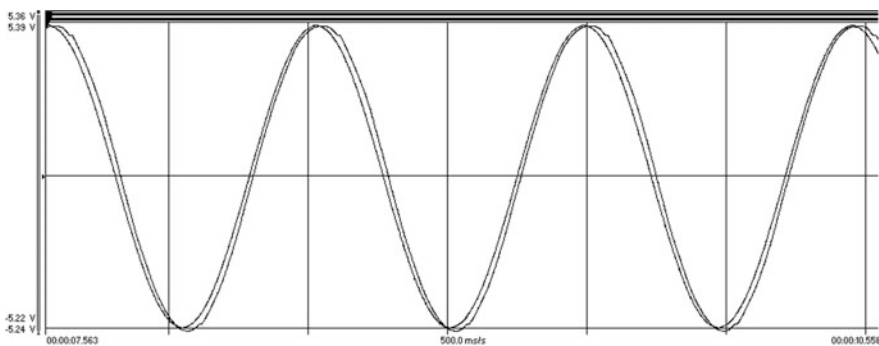


Fig. 8. Tracking of 1 Hz sine wave with 10 V peak to peak

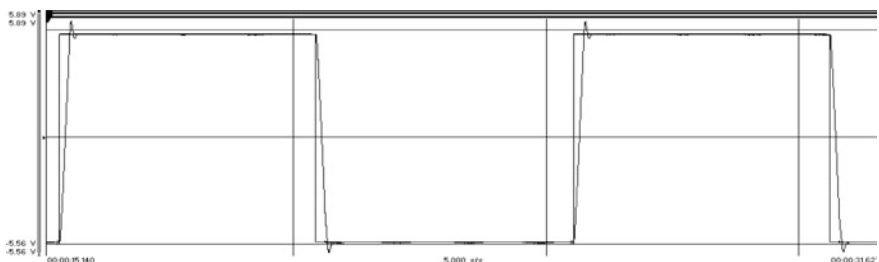


Fig. 9. Tracking of 0.5 Hz square wave with 10 V peak to peak

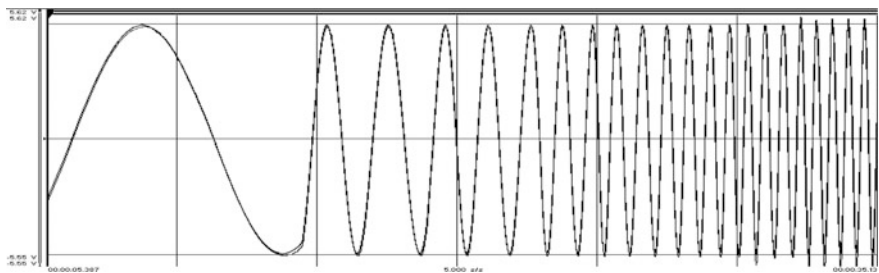


Fig. 10. Tracking of random frequency sine wave with 10 V peak to peak

0.2 V and its adjustment time is 36 ms. And the overshoot of 1 Hz square wave with 10 V is 0.7 V and its adjustment time is 160 ms. We can see with frequency analyzer, that the 3 dB bandwidth of system is 3 Hz when instruction signal is 10 V wave.

4 Conclusion

The design using FPGA with the advantages of high speed, work stability and easily modifying have realized three loops control in direct current motor servo system, and corresponding with the external circuit, the whole system performance is greatly improved. The design of the control algorithms with Quartus II and the Nios II embedded processor is customized with Qsys, which have the advantages of a great deal of flexibility, augmentability, and short design development cycle.

References

1. Li, G.: Servo drive technology and its development. *Electrotech.* (12) (2000)
2. Lati, A., Irki, Z., Sakhi, S., Nemra, A., Hamerlain, M.: Implementation of features detection and matching algorithms on FPGA using Nios II. In: *International Conference on Modelling*, pp. 55–60 (2017)
3. Solet, D., Béchenec, J.L., Briday, M., Faucou, S., Pillement, S.: Hardware runtime verification of embedded software in SoPC. In: *Industrial Embedded Systems*, pp. 1–6 (2016)
4. Zhou, R., Du, Y., Yuan, W.: Application of fuzzy-PI controller with feed forward control in direct current motor servo system. In: *2005 International Conference on Neural Networks & Brain (ICNN&B '05)*, vol. 2 of 3, pp. 1262–1267 (2005)
5. Subiramonian, S., Jawhar, S.J.: FPGA implementation of ABC & Cuckoo search optimization based PI controller for speed control of DC servo motor drives. *Int. J. Appl. Eng. R....* **9**(24), 24349–24360 (2014)
6. Sun, H., Wang, X., Lin, Q., Wang, X., Su, S.: The design of the DC servo motor controller based on fuzzy immune PID algorithm. In: *Chinese Control Conference*, pp. 4724–4729 (2017)
7. Sathyan, A., Milivojevic, N., Lee, Y.J., Krishnamurthy, M., Emadi, A.: An FPGA-based novel digital PWM control scheme for BLDC motor drives. *IEEE Trans. Industr. Electron.* **56** (8), 3040–3049 (2009)
8. Yang, Y., Liao, Y., Li, F.: Design of BLDCM actuator position servo controller based on FPGA micromotor. **43**(3), 48–51 (2010)
9. Li, F., Wu, J.F., Wu, Y.H., Li, H.W., Zhang, P., Xuan, M.: Research and realization of BLDC acceleration control based on SOPC. *Electron. Technol. Appl.* (04) (2010)



Design and Implementation of a Numerical Control High-Precision Low-Temperature Drift Constant Current Source

Xiwei Guo, Chaochao Yang^(✉), and Deliang Liu

Missile Engineering Department, Shijiazhuang Campus, Army Engineering University of PLA, Shijiazhuang 050003, China
1508022921@qq.com

Abstract. According to the testing demand of a certain type of equipment, the current design scheme of constant current source is compared and analyzed, and the overall scheme is determined by considering the factors such as cost and index. Next, it is designed and hardware selected of power supply circuit, constant current circuit, and control circuit, and the hardware circuit is built. The system software design flow is analyzed. Finally, the system is tested and the results show that the design requirements are met.

Keywords: Numerical control · High-precision · Constant current source

Constant current source is a kind of unit circuit or test platform widely used in electronic system, which plays an extremely important role in the field of modern measurement and measurement. It is widely used in practical projects and is commonly power supply of storage battery charging, LEDs Drive, sensor design, electroplating, laser drive, and other fields.

The current output of the more mature constant current source products in the market may be up to milliamperes to hundred amperes level, However, the output current size, stability, and accuracy of the indicators cannot be very good to meet the requirements of a certain type of equipment system testing Applications. Therefore, this paper designs and manufactures a numerical control high-precision low-temperature drift current source.

1 Requirements Analysis and Overall Solution Design

For a certain type of equipment system testing requirements, the constant current source system using numerical control, DC output current stability, can be in 10–500 mA the range of arbitrary settings, not with the load and ambient temperature changes, output current error range is \pm mA.

The key to realizing the numerical control constant current source is the high-precision constant current, and there are three common methods to get that: constant current diode or transistor, four-terminal adjustable constant current device, voltage-controlled constant current source. There are many ways to achieve numerical control, such as digital logic device, programmable device CPLD/FPGA, single-chip micro-computer, etc. [1].

1.1 Constant Current Source Scheme Analysis and Design

Scheme 1: the constant output current is obtained by using two identical transistors with a relatively stable BE voltage, and the structure is shown in Fig. 1a. Output current size $I = V/B$. The constant current source is simple and easy, but the same type of triode also has certain individual differences, so it is not suitable for the constant current demand with high precision.

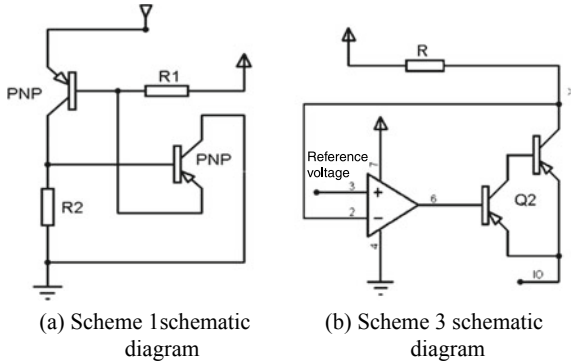


Fig. 1. Constant current source schematic diagram

Scheme 2: Using four-terminal adjustable constant current source to achieve cross-flow, this device output current can reach 2200 mA, by changing the peripheral component (resistance) parameters, so that the current can be adjusted to achieve the purpose. There are usually two ways of doing this: first, through manual adjustment to change output current, this method can not meet the design task of the NC adjustment requirements, the second is to change the required resistance parameters through digital potentiometer, although it can achieve the purpose of numerical control, the digital potentiometer for each step of the resistance is relatively large, It is difficult to realize the output of the step current in the design task.

Scheme 3: The constant current source consisting of a reference voltage source, a comparison amplifier, an adjustment tube, a sampling resistor, and its structure as shown in Fig. 1b. The output current is converted to a voltage by the sampling resistor, the output voltage is fed back to the input end of the comparator amplifier, and then compared with the reference voltage, the amplifier adjusts the error voltage to control the regulating conduction state, adjusts the output current, and maintains the output current constant.

Based on the above analysis, option three is chosen as the design scheme of the constant current source.

1.2 Numerical Control Scheme Analysis and Design

The Scheme 1 is shown in Fig. 2. This scenario uses a set of decimal counters, on the one hand, the counter completes the voltage decoding display, on the other hand the output of the counter is entered as the address of the EPROM, and the output of the EPROM is transformed by the D/A transform to achieve the output stepping by controlling the reference voltage of the magnified error. However, because this scheme uses Open-loop control strategy, the circuit is simple and the cost is low, but the final output cannot be adjusted and modified, so that the output current accuracy is not high, and control data burning in the EPROM, reducing the system design flexibility and the adaptive ability is poor [2].

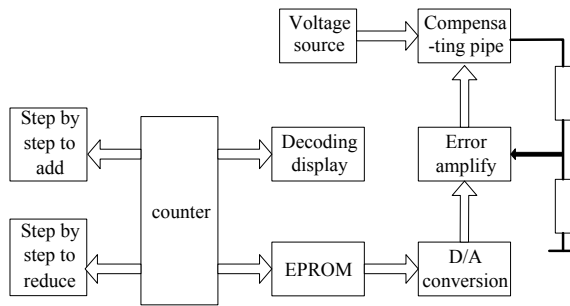


Fig. 2. Schematic diagram of scheme 1

The Scheme 2 is shown in Fig. 3, it is mainly based on SCM as the core to build the controller, through the keyboard to preset the current value, SCM output the corresponding digital signal, after D/A conversion, signal amplification, level conversion, voltage control constant current source, output current signal. The actual output current uses the precision resistance sampling to become the voltage signal, after the high input impedance differential amplifier, the A/d conversion, the signal is fed to the microcontroller to compare the output feedback signal with the preset value, send out the adjustment signal, then output the new current, In this way, the closed-loop regulation is formed, the output current is locked, and the precision and stability of the output current are improved. This scheme adopts single-chip microcomputer to control, display and preset number, which makes the system flexible and convenient, and the current output precision and stability are higher. However, the stability of this scheme is limited by the ability of the SCM to process data.

The whole schematic diagram of scheme 3 is approximately the same as scheme 2. It uses the FPGA chip to make the system, and uses the embedded soft-core processor for the overall control, algorithm operation, display and number of functions, with VHDL language Design Digital Hardware control module control, with fast running, stable and reliable characteristics.

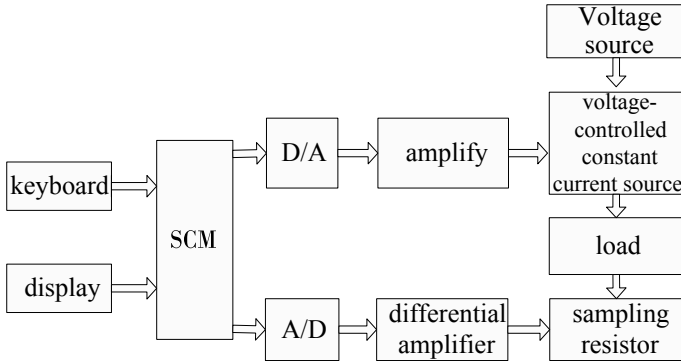


Fig. 3. Schematic diagram of scheme 2

The above three NC implementation schemes have their own characteristics. After comparison can be seen, scheme three is the optimal solution, but given the heavy workload adopts digital control hardware module design and debugging complex, so the final option two, that is, single-chip computer as the core to achieve numerical control functions.

Based on the above analysis, the project is improved on the basis of program two of the whole scheme, combining it with the three-phase scheme of the constant current source scheme, which is composed of MCU as the core of building controller, through the keyboard to preset the current value or the key to gradually fine-tune, single-chip output corresponding digital signal, after the D/A conversion, signal amplification, voltage control constant current source, output current signal. The system composition is shown in Diagram 4.

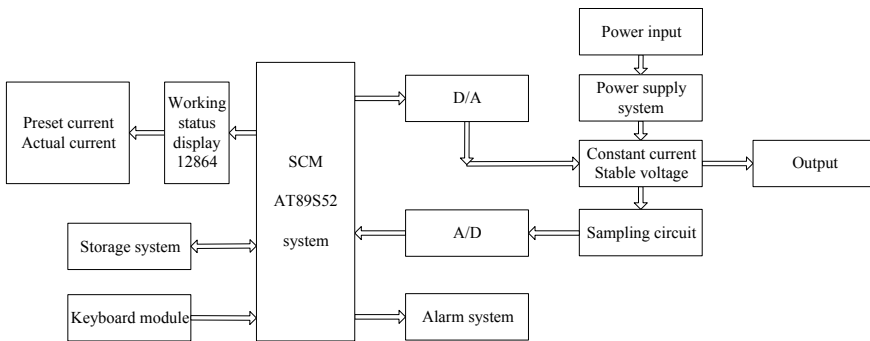


Fig. 4. System composition diagram

2 Main Hardware Circuit Design

2.1 Power Supply Circuit Design

For the sake of design aesthetics, the power supply circuit of the system is integrated into one board when designing the hardware circuit. The system uses mains power supply, and the Rectifier Bridge (KBP307) is rectified after the transformer step-down. Because the voltage after rectification is not very smooth, so the system design coupled with a filter circuit to smooth the voltage, the supply voltage is only a small ripple, which also to improve the accuracy of constant current source laid the foundation. Finally, the project uses a three-terminal voltage regulator (7815, 1915, 7805 and 7905) level two step-down for constant current source and SCM system power supply [3]. The specific power supply circuit design is shown in Fig. 5.

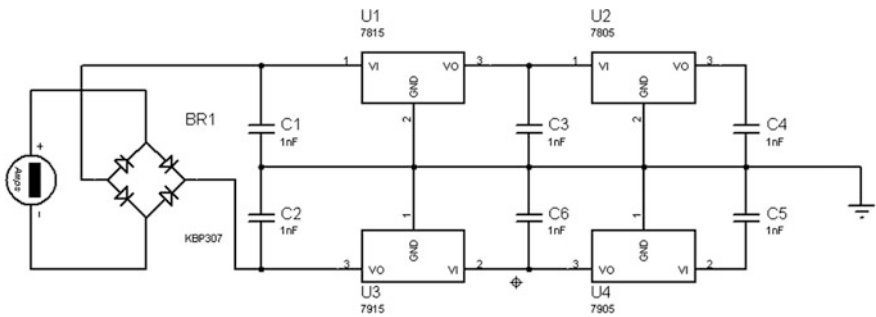


Fig. 5. Power supply circuit design

2.2 Design of Constant Current Circuit

Constant current circuit principles as shown Fig. 6 shows.

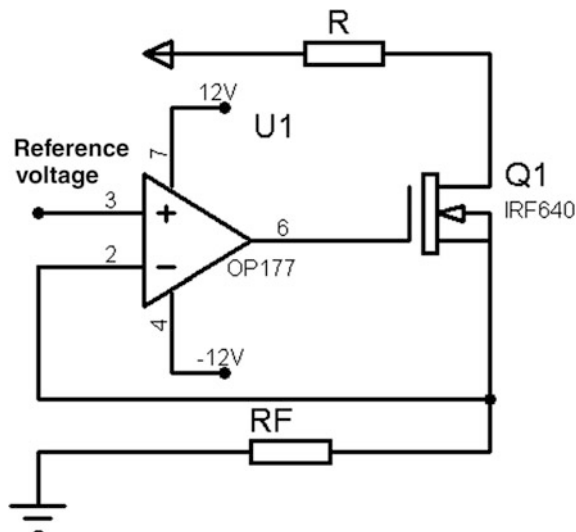


Fig. 6. Constant current circuit

Constant current circuit is the main circuit in the whole system, its working principle: by the D/A converter produces a controllable constant voltage to the same phase input of the comparator. It is known by the virtual short principle of the amplifier that the voltage on the reverse input end is the same as the input end of the phase, so the sampling resistor to the ground voltage is D/A the voltage value output by the converter, It can be seen from $I = U/R$ that the current flowing through the sampling resistance is a constant value, and this voltage value is controlled by the single-chip computer. Because the comparator has high impedance, the current can only be supplied by the field-effect tube, because the voltage at both ends of the sampling resistor is constant, the amplifier provides the source of the FET, making it in a certain conduction state, and when the load changes, the feedback changes, which leads to the change of the conduction state of the FET and the voltage at both ends of the sampling resistor will not change, that is, the system can get a constant current value.

In order to realize high-precision and low-temperature drift, the sample resistance is used as the sampling resistance of the manganese copper precision resistance with no temperature drift. In addition, in the sample resistance connection, the method of combination of sampling series and parallel is used to offset the influence of temperature on the system precision by using different characteristics of different resistance temperature drift coefficient. OP177 high-precision operational amplifier is selected as the most comparator, with small temperature drift and large open loop gain, which further improves the system precision.

2.3 Control Circuit Design

The control circuit of SCM is shown in Fig. 7.

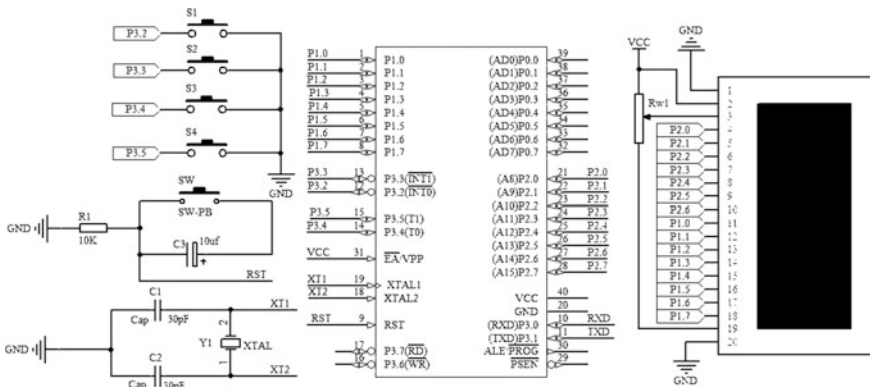


Fig. 7. SCM system circuit diagram

The control module uses the key to realize human-computer interaction for the basic test mode and drawing mode switching. The design selects LCD12864 to display working status, making the entire control system and display interface more user-friendly.

3 System Software Design

The function of software can be divided into two categories: the first is the monitoring software (main program), which is the core of the whole control system and is used to coordinate the relationship between the execution module and the operator. The second category is the implementation software (subroutine), which is used to perform various substantive functions such as input, control, sampling, alarm, display, etc. Each execution software is also a small function execution module. Each execution module is listed here, and the function definition and interface definition are defined for each execution module. After each execution module is planned, the monitoring program can be planned [4].

The overall design block diagram of the system software programming is shown in Fig. 8.

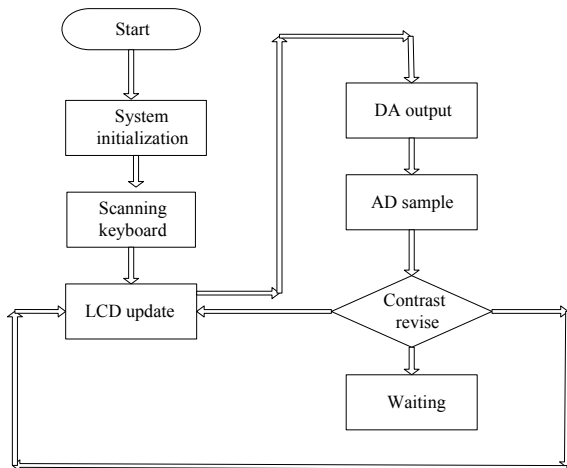


Fig. 8. Main program flowchart

4 System Test

After the design and production of software and hardware, the physical circuit board is shown in Fig. 9.

The system test is the main means to reflect the performance of the equipment, so the step current test and the load resistance change test are respectively carried out.

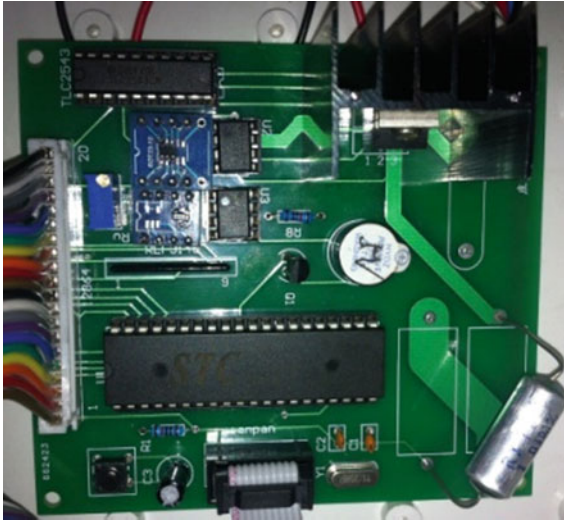


Fig. 9. Board physical drawing

4.1 Step Current Test

This system adopts the 12 serial D/A as the part of mc system, through the button control output analog and analog to control the voltage sampling resistance, voltage constant current source are formed by sampling resistance is converted to current. During the test, we used the keyboard control simulation to achieve the step and used a multimeter to test the actual output current. Measurement data as shown in Table 1 (current unit: mA).

Table 1. Step current test

Input current	51	52	53	54	55	56	57	58	59	60
Measured current	51.0	51.9	52.9	53.9	54.9	55.9	56.9	57.9	58.9	59.9
Difference	0.00	-0.10	-0.10	-0.10	-0.10	-0.10	-0.10	-0.10	-0.10	-0.10
Error (%)	0.00	-0.19	-0.19	-0.19	-0.18	-0.18	-0.18	-0.17	-0.17	-0.17
Input current	101	102	103	104	105	106	107	300	400	500
Measured current	101	102	103	104	105	106	107	300	400	500
Difference	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Error (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

4.2 Load Resistance Change Test

By changing the load of the system, test the system with load capacity, the test method is the same as above.

The test results are shown below (Table 2).

Table 2. 4 Load resistance change test

<i>When unloaded</i>						
Input current value	10	50	100	150	200	500
Measured current value	9.90	49.9	100	150.1	200	500
Difference	-0.10	-0.10	0.00	0.10	0.00	0.00
Error (%)	-1.00	-0.20	0.00	0.07	0.00	0.00
<i>Load is 1 Euclidean</i>						
Input current value	10	50	100	150	200	500
Measured current value	9.90	49.9	100	150.1	200	500
Difference	-0.10	-0.10	0.00	0.10	0.00	0.00
Error (%)	-1.00	-0.20	0.00	0.07	0.00	0.00
<i>Load is 3 Euclidean</i>						
Input current value	10	50	100	150	200	500
Measured current value	9.90	49.9	100	150.1	200	500
Difference	-0.10	-0.10	0.00	0.10	0.00	0.00
Error (%)	-1.00	-0.20	0.00	0.07	0.00	0.00
<i>Load is 5 Euclidean</i>						
Input current value	10	50	100	150	200	500
Measured current value	9.90	49.9	100	150.1	200	500
Difference	-0.10	-0.10	0.00	0.10	0.00	0.00
Error (%)	-1.00	-0.20	0.00	0.07	0.00	0.00

5 Conclusion

According to the test result, the technical performance index of the system is analyzed:

- (1) The current output range can be from the minimum 10 mA to start stepping to 500 mA.
- (2) Changes the load size, has no effect on the output current, and realizes the purpose of constant current.
- (3) System temperature drift small, high-precision, to meet the design requirements.

The measured data shows that using this scheme the measured current value and the actual current value of error is very small, only found in measurement data may fluctuate in 1 mA, sampling resistor of resistance because of the change in temperature may cause the change of resistance. Therefore, there are slight changes in the output current value. But it has some advantages over conventional constant current source.

References

1. Deng, M.-s., Yan, J., Liao, W.-x.: Design of a combined-type and precise constant-current power source. *Power Supply Technol. Appl.* **13**(10), 42–45 (2010)
2. Xue, X.-l.: Design and realization of a new high-precision digital controlled constant current source. *J. Minjiang Univ.* **31**(5), 29–32 (2010)

3. Lin, Z.-h., Mao, Y.-l., Zhan, P.: Design and realization of current source system in energy-saving lamp test equipment. *Mach. Build. & Autom.* **38**(4), 132–134 (2009)
4. Xue, Yi: Design of a high precision program-controlled constant-current source. *Process. Autom. Instrum.* **30**, 63–65 (2009)



Compact Hybrid-Integrated Circular Polarized Double-Ring Antenna for Satellite Application

Sihao Chen¹(✉), Dongliang Fei², and Lianxing He^{1,2}

¹ ShanghaiTech University, Shanghai, China

chensh@shanghaitech.edu.cn, 18800295770@163.com

² Innovation Academy for Microsatellites, Chinese Academy of Sciences, Shanghai, China

feidongliang@sina.com, 13816474811@139.com

Abstract. In this paper, a compact hybrid-integrated circular polarized double-ring antenna is proposed based on the Yagi-Uda concept. With two cross-slotted patch hybrid coupler integrated inside the square ring, the proposed antenna can effectively radiate a circular polarized wave with lightweight and more compact size. With six other rectangular rings aligned at the right positions to act as wave directors, a gain up to 10 dBi can be achieved. The proposed antenna is thus especially suitable for satellite communications where weight and size are major constraints.

Keywords: Compact microstrip Yagi-Uda antenna · Satellite-borne antenna · Double-ring antenna

1 Introduction

With the rapid growth of ultra high frequency (UHF) antennas on satellites, airplanes, and boats, the design of higher performance antennas becomes much more urgent. Due to strict design constraints, satellite UHF antennas have high demands in term of gain, preferably with smaller size and lighter weight at the same time. Existing satellite UHF antennas, such as metal Yagi-Uda antennas, are too big and heavy to meet the demand of the satellite applications, especially for microsatellites, where the weight and size are extremely restricted. Thus, a highly integrated compact antenna with miniaturized size and high gain is needed.

Among the different antennas with profile miniaturization, the square-ring antenna seems to be a good candidate for a compact antenna because its side length of ring is about a quarter-wavelength when it works at its fundamental mode, which is only a half of the conventional patch radiator [1, 2]. There are two traditional ways to generate circular polarization. The first one is by introducing perturbation segments along the ring resonator, which generates circular

polarization with narrow AR bandwidth [2–4], and the other way is to use an external hybrid coupler to produce two waves with the same amplitude and 90° phase difference [5–7]. Even though the AR bandwidth of the second method is wider than first one, the total size of the antenna system is much larger due to the externally hybrid coupler.

There are several ways to make such antennas more compact by reducing the size of hybrid coupler. The most common way is using a meandered cross-slot line patch to make hybrid coupler more compact [8]. Since the size of hybrid coupler is relatively small, it can be easily embedded into the ring antenna. Due to the finite ground plane, the backside radiation will increase which can cause the decrease of the antenna gain. Another meandered ring resonator below the main radiator with a certain distance can be used as an inductive reflector to reduce the backside radiation and obtain a directional radiation pattern [9].

In this paper, a ring antenna with meandered cross-slot line patch hybrid coupler inside is proposed. A parallel-coupled feeding is necessary as a matching network in order to generate two orthogonally fundamental modes. Another meandered ring resonator is placed below the primary radiator. The proposed antenna is designed in the general idea mentioned above and fabricated on FR4 substrates.

2 Configurations of the Proposed Antenna

The top view of the proposed compact hybrid-integrated circular polarized double-ring antenna is shown in Fig. 1a. Microstrip rectangular ring antenna is on the top of FR4 substrate with a cross-slotted patch hybrid coupler embedded inside. The length of the ring antenna is about half of wavelength. Due to the large input impedance, a parallel-coupled feeding is designed for feeding the ring resonator. The width, the length, and the gap between the microstrip and the resonator are properly tuned to meet the operating frequency requirements. Space-based satellite application normally requires antennas that radiate a uni-directional pattern with a significant front-to-back ratio to ensure high security and efficiency in the propagation. A reflector, shown in Fig. 1b, is thus designed for strengthening the directional radiation pattern. The total length of the meandered ring is chosen to be larger than half of wavelength, to acts as an inductive reflector. By tuning the length of the meandered ring and the distance between reflector and radiator, the induced current in the reflector becomes inductive which enhance the directional radiation pattern at the far field. Furthermore, six other rectangular rings are aligned sequentially with certain distance, which makes further contributions on the directional radiation pattern as well. The complete model of proposed antenna is shown in the Fig. 1c (Fig. 2).

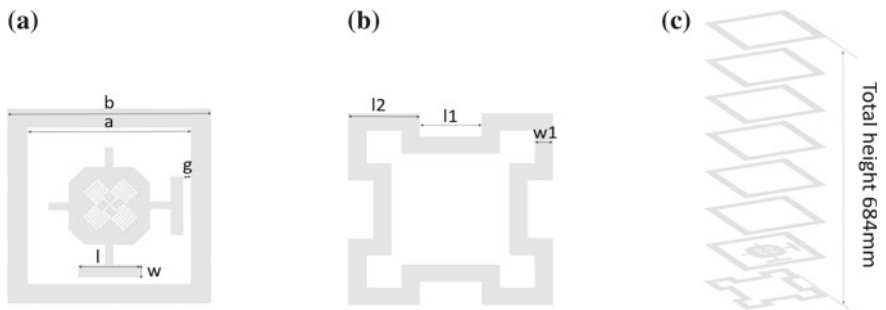


Fig. 1. Configuration of the as-proposed antenna: **a** top view of the radiation patch; **b** the reflector; **c** layout of the proposed antenna

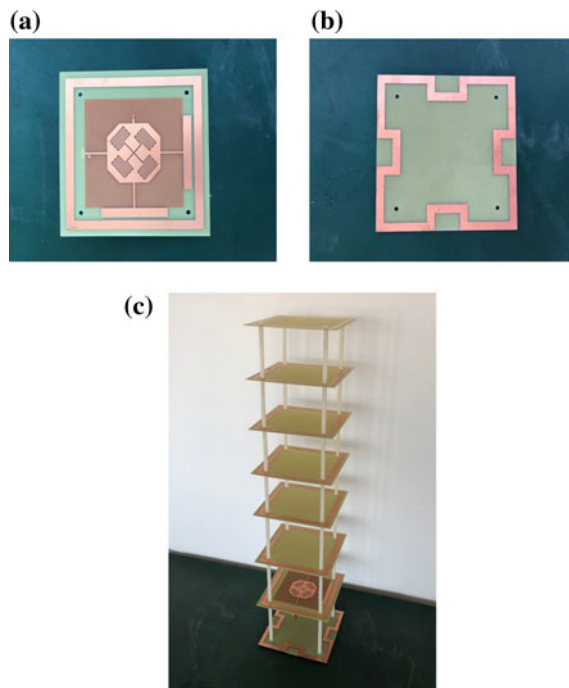


Fig. 2. Pictures of the antenna: **a** top view of the radiation patch; **b** the reflector; **c** layout of the proposed antenna

3 Simulation Results and Experiment Verification

The simulated and measured S11 of proposed antenna is shown in Fig. 3, the measured result is slightly off the predicted result; however, the bandwidth is large enough to cover the requested central frequency. The slight deviation might be caused by the fabrication error of the patch.

The simulated and measured axial ratio of proposed antenna is shown in Fig. 4. Although there are a small discrepancies in the low- and high-frequency range, axial ratio in the required frequency range (around 435 MHz) are both smaller than 3 dB, the antenna produces good circular polarization performance.

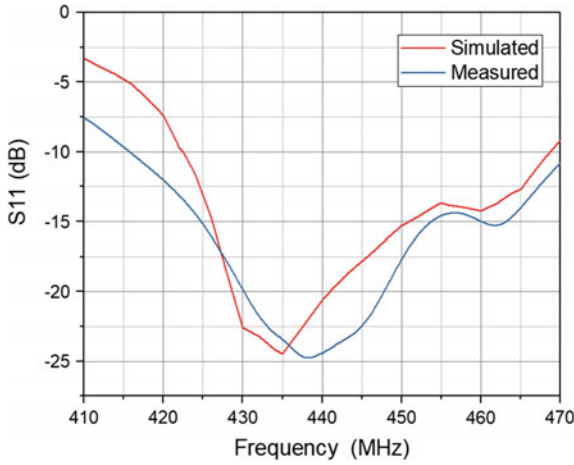


Fig. 3. S11 of the simulated and measured antenna

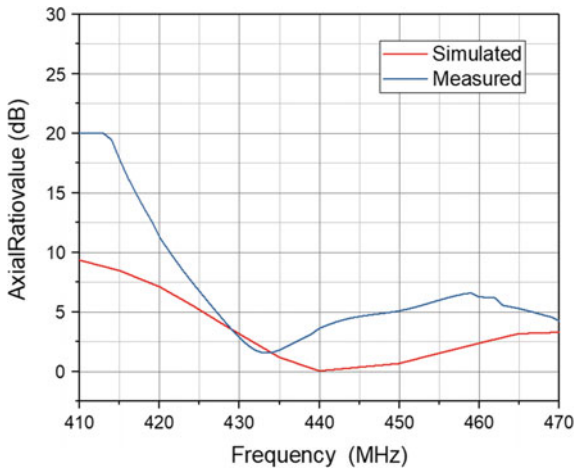


Fig. 4. Axial ratio of simulated and fabricated antennas

The radiation patterns of the as-developed prototype are shown in Fig. 5. The simulated gain around the operating frequency is about 11.5 dB while the measured gain is about 9.8 dB. There are slight discrepancies in gain and radiation pattern, and the measured directivity is not as good as simulated. That might be caused by the unstable performance of the FR4 substrate with high dielectric loss and inaccurate thickness of the air layers, as the proposed antenna uses six rectangular rings as a wave director which increase the uncertainty thickness of the air layers. Also, the feeding probes and the “frame” sticks connecting different layers are not considered in the simulation, and may be interfering with performance.

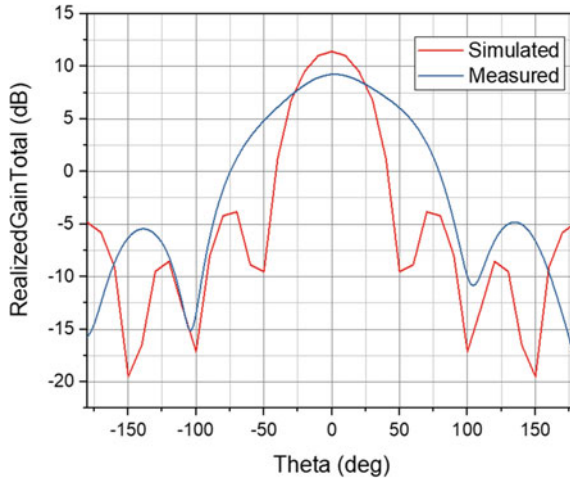


Fig. 5. Radiation patterns of simulated and measured antennas

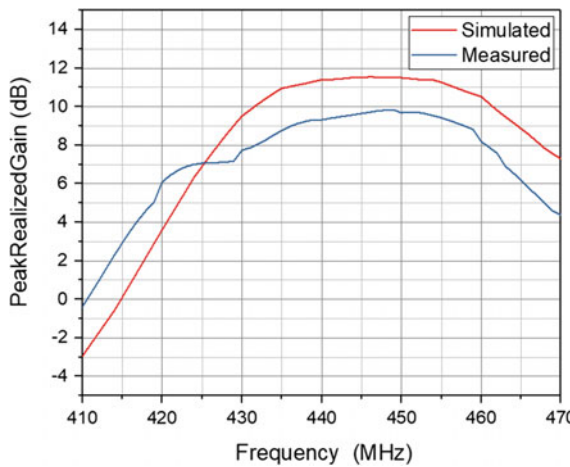


Fig. 6. Gains relationship with frequency of simulated and measured antennas

The gains as a relationship with the frequencies are shown in Fig. 6. The measured gain at 435 MHz is about 9 dB and the simulated result is nearly 2 dB higher than that. The reason causing the discrepancy might be the position of the feed point which is not exactly the same as in the simulation model. And that means, the impedance matching might not be as good as the simulation model.

4 Conclusion

In this paper, a compact hybrid-integrated circular polarized double-ring antenna with six layers wave directors is proposed and verified by a prototype. The antenna performed well at 440 MHz by tuning the parameters mentioned above. The cross-slot line patch plays an essential role in making the antenna compact. The additional six rectangular rings proved to work well as wave directors, producing high directivity and gain. A gain up to 10 dBi can be achieved. This antenna can be further stowed into a very compact volume for accommodation during launch and once orbit it is unfurled to a deployed length as showed in Fig. 2. As weight and power are very limited and expensive resources on board a satellite. The proposed antenna is especially suitable for satellite communication where weight and power are normally restricted.

Acknowledgements. The projects are supported by the National Natural Science Foundation of China (Grant No.61671304) and Shanghai Sailing Program (16YF1411000).

References

1. Zhu, L., Li, K.: CPW-fed rectangular microstrip ring antenna for suppression of parasitic backside radiation. *Microw. Opt. Technol. Lett.* **36**(1), 65–67 (2003)
2. Row, J.S.: Design of square-ring microstrip antenna for circular polarization. *Electron. Lett.* **40**(2), 93–95 (2004)
3. Chen, H.M., Wong, K.L.: On the circular polarization operation of annular-ring microstrip antennas. *IEEE Trans. Antennas Propag.* **47**, 1289–1292 (1999)
4. Liu, J.C., Zeng, B.-H., Badjie, L., Drammeh, S., Bor, S.-S., Hung, T.-F., Chang, D.-C.: Single-feed circularly polarized aperture-coupled stack antenna with dual-mode square loop radiator. *IEEE Antennas Wirel. Propag. Lett.* **9**, 887–890 (2010)
5. Row, J.S.: Design of aperture-coupled annular-ring microstrip antennas for circular polarization. *IEEE Trans. Antennas Propag.* **53**(5), 1779–1784 (2005)
6. Qing, X.M., Chia, Y.W.M.: Circularly polarised circular ring slot antenna fed by stripline hybrid coupler. *Electron. Lett.* **35**(25), 2154–2155 (1999)
7. Lin, Y.F., Chen, H.M., Chu, F.H., Pan, S.C.: Bidirectional radiated circularly polarised square-ring antenna for portable RFID reader. *Electron. Lett.* **44**(24), 1383–1384 (2008)
8. Sun, S., Zhu, L.: Miniaturized patch hybrid couplers using asymmetrically loaded cross slots. *IET Microw. Antennas Propag.* **4**(9), 1427–1433 (2010)
9. Li, Y., Sun, S., Yang, F.: A miniaturized Yagi-Uda-oriented double-ring antenna with circular polarization and directional pattern. *IEEE Antennas Wirel. Propag.* **12**, 945–948 (2013)



Influence of Center of Mass Movement on Steering Characteristics of Front-Wheel Steering Vehicles and Four-Wheel Steering Vehicles

Shuaijun Yang^(✉), Yinshan Wang, and Shaoyun Lu

School of Automobile and Transportation, Tianjin University of Technology and Education, Dagou South Road 1310, Tianjin 300222, China
shuaijunyang209@163.com

Abstract. In order to analyze the effect of car center of mass on the steering stability of front-wheel steering vehicles and four-wheel steering vehicles, and based on the two-degree-of-freedom vehicle dynamics model, this article uses Matlab/Simulink to carry out the joint simulation of its steering characteristics. Based on the two conditions of front-wheel steering and four-wheel steering respectively, the stability of vehicle before and after the change of center of mass at different vehicle speeds is discussed and analyzed. The result shows that vehicle stability is greatly reduced after the center of mass moves. But when at low-speed, the steering stability after the center of mass movement of the four-wheel steering vehicle is still better than that of the front-wheel steering vehicle of the normal parameters.

Keywords: Centroid transfer · Front-wheel steering · Four-wheel steering · Matlab/Simulink

1 FWS Mathematical Model of Steering Movement

Four-wheel steering can provide good help for the stability of the vehicle when steering. The parameters such as the turning radius and the lateral slip angle of the center of mass of 4WS are better than those of front-wheel Steering. However, based on the considerations of maneuverability and economy in modern vehicles, there are still a number of vehicles using front-wheel steering [1–3].

The existing researches on vehicle handling stability are mostly based on the constant center of mass. However, due to the change of the vehicle load during actual driving, the center of mass of the vehicle will be shifted and the steering characteristics will be changed [4]. In this paper, the Centroid angle and the yaw rate are used as research parameters to study the effects of vehicle speed, centroid position, and other factors on the steering characteristics of front-wheel steering and four-wheel steering.

According to the two-degree-of-freedom FWS car model, establish differential equations:

$$\begin{cases} m(\dot{v} + u\omega_r) = (k_1 + k_2)\beta + \frac{1}{u}(ak_1 - bk_2)\omega_r - k_1\delta \\ I_z\dot{\omega}_r = (ak_1 - bk_2)\beta + \frac{1}{u}(a^2k_1 - b^2k_2)\omega_r - ak_1\delta \end{cases} \quad (1.1)$$

where m —total vehicle mass; u —forward speed of the vehicle; v —vehicle lateral speed; ω_r —vehicle yaw rate; β —vehicle centroid angle; I_z —vehicle yaw inertia; a , b —centroid pair Front and rear axle distance; k_1 , k_2 —front and rear wheel cornering stiffness; δ_f —front-wheel corner.

Simplify Formula (1.1), and we get:

$$\begin{cases} \dot{v} = \frac{(k_1 + k_2)}{mu}v + \left[\frac{ak_1 - bk_2}{mu} - u \right] \omega_r - \frac{k_1\delta}{m} \\ \dot{\omega}_r = \frac{(ak_1 - bk_2)}{I_z u}v + \frac{(a^2k_1 - b^2k_2)}{I_z u} \omega_r - \frac{ak_1\delta}{I_z} \end{cases} \quad (1.2)$$

Make the state variable $X = \begin{bmatrix} v \\ \omega_r \end{bmatrix}$. Input variable $U = \delta_f$. The above formula can be abbreviated as:

$$\dot{X} = AX + BU$$

$$Y = CX + DU$$

$$A = \begin{bmatrix} \frac{k_1 + k_2}{mu} & \frac{ak_1 - bk_2}{mu} - u \\ \frac{ak_1 - bk_2}{I_z u} & \frac{a^2k_1 + b^2k_2}{I_z u} \end{bmatrix}; \quad B = \begin{bmatrix} -\frac{k_1}{m} \\ -\frac{ak_1}{I_z} \end{bmatrix};$$

C-output matrix, in order to directly output the state variable X , C is set to the unit matrix; D -direct transfer matrix, because the output matrix is set to the identity matrix, so D is set to zero matrix.

2 Simulation Analysis of FWS Vehicle Movement

The MATLAB/Simulink module is now used to simulate the above state-space equations by a special State-Space module [5]. The front-wheel is given an angle step input of 0.135 rad. The simulation uses a car parameter, as follows.

total vehicle mass $m = 2045$ kg, vehicle yaw inertia $I_z = 5428$ kg m², centroid pair front axle distance $a = 1.488$; centroid pair rear axle distance $b = 1.712$ m, front-wheel cornering stiffness $k_1 = -77.85$ kN/rad, rear wheel cornering stiffness $k_2 = -76.51$ kN/rad.

2.1 Influence of Different Vehicle Speeds on Transient Response Characteristics

In order to study the transient response characteristics of the front-wheel-steered vehicle at different vehicle speeds, the simulation result tests of vehicles with vehicle speeds of 30, 80 and 120 km/h were conducted. The simulation results are shown in the figure below.

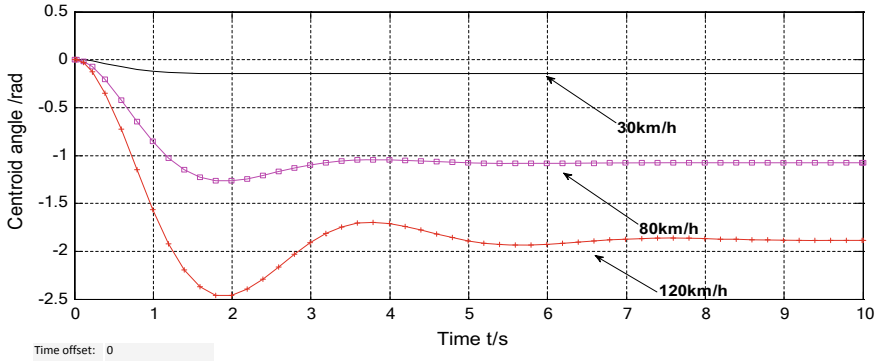


Fig. 1. Comparison chart of the frontal steering centroid angle

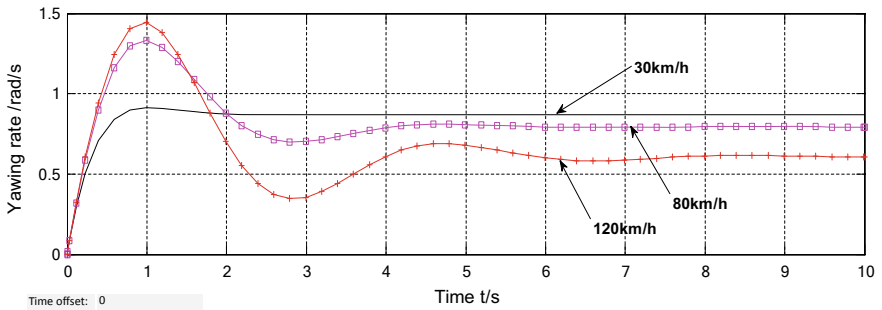


Fig. 2. Comparison chart of front-wheel steering yaw rate

As can be seen from Fig. 1, as the speed increases, the vibration amplitude of the centroid side-angle vibration increases, and the settling time also increases. At a vehicle speed of 30 km/h, the lateral slip angle of the center of mass tends to be stable at about 2 s, but when the speed reaches 120 km/h, the lateral slip angle of the centroid gradually becomes stable after 5 s. It is proved that the stability of the vehicle at low speed is higher than that at high speed.

At the same time, it can be seen from Fig. 2 by the comparative analysis of each curve, as the vehicle speed increases, the steady-state value of the yaw rate decreases, the overshoot increases, the stabilization time increases, and the reaction time also increases.

2.2 Influence of Center of Mass Movement on Front Steering

In the following, we will move the center of mass of the vehicle 200 mm backward to analyze the influence of the mobile centroid on vehicle stability. And as a comparison, the vehicles with the speeds of 30, 80 and 120 km/h were again selected to conduct the simulation result tests. The simulation results are as follows (Figs. 3 and 4).

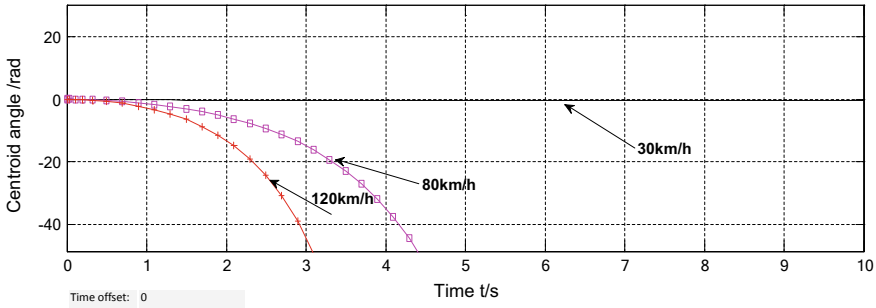


Fig. 3. Comparison chart of the frontal steering centroid angle after the centroid change

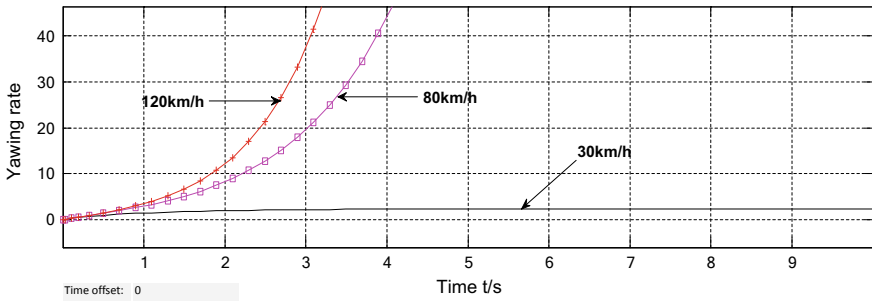


Fig. 4. Comparison chart of front-wheel steering yaw rate after center of mass change

It can be clearly seen from the above two figures that the change of the center of mass has a great influence on the stability and stability of the vehicle. At low speeds, the vehicle can still maintain stability. When the vehicle speed is too high and reaches 80 km/h or more, the centroid slip angle and yaw rate do not converge and the vibration amplitude is large. It proves that the influence of center of mass movement on the handling stability of the vehicle gradually increases with the increase of vehicle speed. At the same time, we can see from the comparison of the yaw rate that the convergence can still be maintained at a low speed of 30 km/h, while after one second of driving at a high speed, the yaw rate of the vehicle begins to increase exponentially and the vehicle stability decreases linearly. With the increase in speed, the situation will further deteriorate.

3 Simulation Analysis of 4WS Vehicle Movement

3.1 Mathematical Model of 4WS Steering Movement

Also based on (1.1):

$$\begin{cases} m(\dot{v} + u\omega_r) = (k_1 + k_2)\beta + \frac{1}{u}(ak_1 - bk_2)\omega_r - k_1\delta \\ I_z\dot{\omega}_r = (ak_1 - bk_2)\beta + \frac{1}{u}(a^2k_1 - b^2k_2)\omega_r - ak_1\delta \end{cases} \quad (1.1)$$

There are many types of four-wheel steering. Here we use a form of four-wheel steering with the ratio (ξ) of the front and rear wheel angles satisfying Eq. (3.1) [6, 7].

$$\xi = \frac{\delta_r}{\delta_f} = \frac{-b + \left[\frac{ma}{k_2(a+b)} \right] u^2}{a + \left[\frac{mb}{k_1(a+b)} \right] u^2} \quad (3.1)$$

Based on the derivation method in the Sect. 1, after adding the corners of the rear wheels, recalculate the differential equations of motion, and abbreviate the form of the space equation. Since the rear wheel angle is taken into account, B changes.

$$B = \begin{bmatrix} -\frac{k_1 + \xi k_2}{m} \\ -\frac{ak_1 - \xi bk_2}{I_z} \end{bmatrix}.$$

3.2 Effects of Different Vehicle Speeds on Transient Response Characteristics

For comparison, we still select the above parameters as the vehicle parameters. Input these parameters into (3.1), $\xi = 0.6702$. Again, the three conditions of 30, 80, and 120 km/h are selected for simulation, and the results are as follows (Figs. 5 and 6).

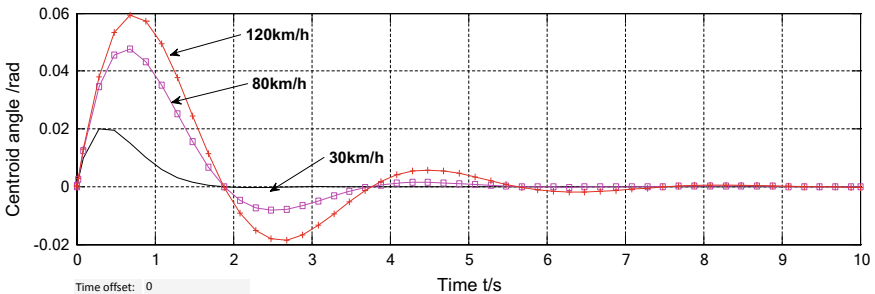


Fig. 5. Comparison chart of four-wheel steering centroid angle

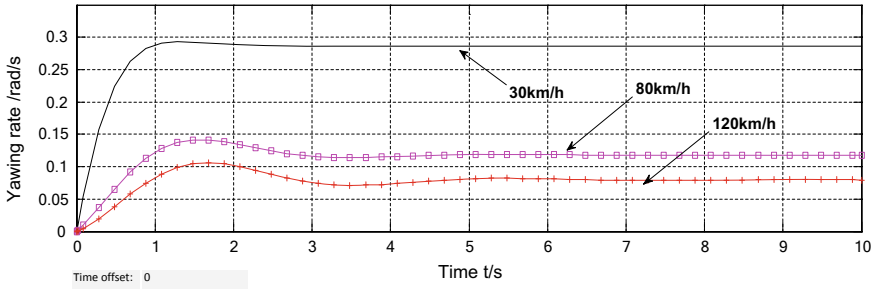


Fig. 6. Comparison chart of four-wheel steering yaw rate

It can be seen that when the vehicle uses four-wheel steering, the centroid angle of the center of mass during steering is much better than that of the front-wheel steering vehicle. However, at the same time, the yaw rate is significantly reduced, the overshoot increases, and the reaction time extends, resulting in a partial decrease in steering wheel sensitivity. However, it is not obvious and it will not cause driving trouble [8, 9].

3.3 Influence of Center of Mass Movement on Four-Wheel Steering

To facilitate comparison, we move the centroid back 200 mm, the same as in Sect. 2.2 above. Simulate and analyze the transient response characteristics of the vehicle steering, and the results are as follows (Figs. 7 and 8).

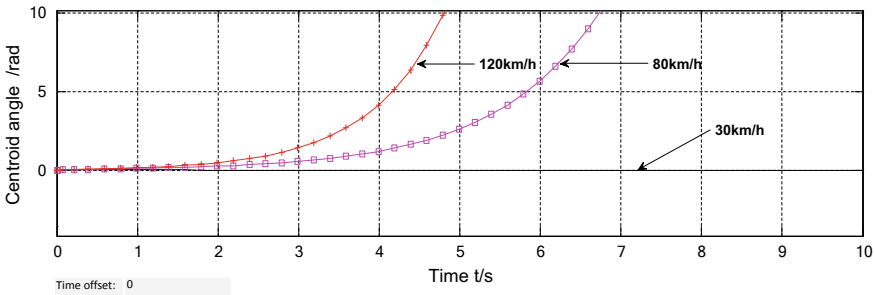


Fig. 7. Comparison chart of four-wheel steered mass centroid angle after centroid change

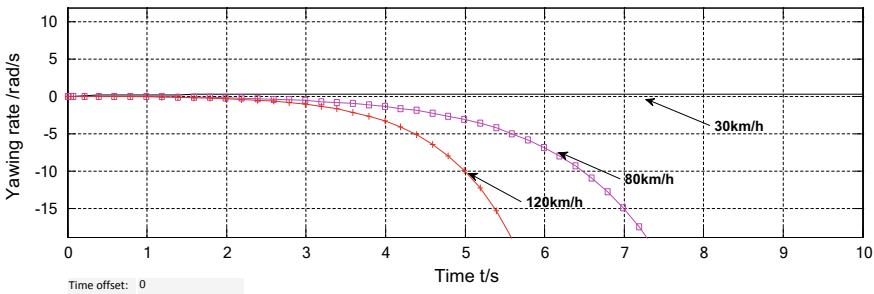


Fig. 8. Comparison chart of four-wheel steering yaw rate after centroid change

From the above figure, we can see that like front-wheel steering vehicles, four-wheel steering vehicles are still stable in low-speed steering, but the curves tend to diverge at high speeds, causing unstable conditions. However, we can see from the lateral angle of centroid that the stability of the four-wheel steering vehicle is obviously better than that of the front-wheel steering vehicle, the time for maintaining stability is also longer than that for the front-wheel steering vehicle, and the safety is relatively improved [10].

It can also be clearly seen from the yaw rate that the four-wheel steering vehicle maintains a more stable duration and amplitude than front-wheel steering vehicle. This shows that in the case of a change in centroid, the ability to maintain vehicle stability with four-wheel steering is still stronger than that with front-wheel steering.

Here we compare the condition of the front-wheel steering vehicle of normal parameters with the condition of the four-wheel steering vehicle after changing the center of mass. In order to simulate the normal driving speed, we select the common speed in the city, 40 km/h, as the comparison speed. The comparison chart is as follows (Figs. 9 and 10).

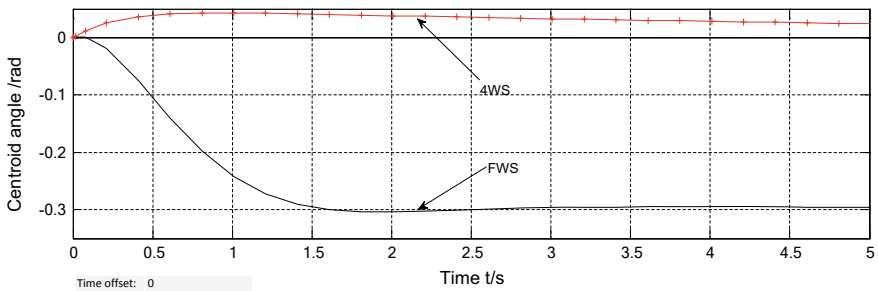


Fig. 9. Contrast of lateral centroid angle

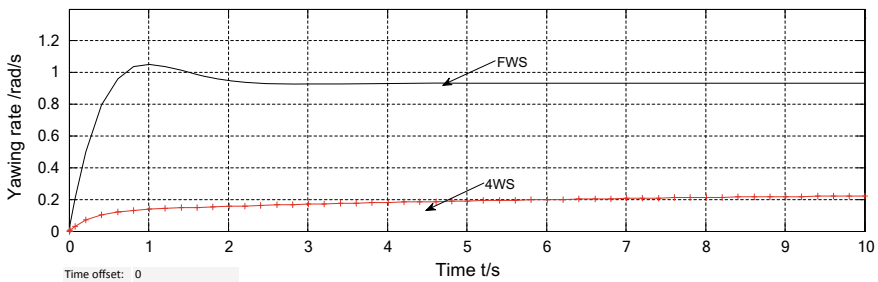


Fig. 10. Comparison of yaw rate

From the above comparison we can clearly see that at the speed of 40 km/h, even after changing the centroid, the four-wheel steering vehicle is superior to the front-wheel steering vehicle of normal parameters in terms of the amplitude and reaction time of the lateral slip angle of the centroid. Once again it proves the four-wheel steering's

stability. From the aspect of the yaw rate, the overshoot of four-wheel steering is reduced, the stabilization time is shortened, and the reaction time is prolonged.

However, due to little interference with the driver, the overall safety of the four-wheel steering vehicle is still higher than the front-wheel steering vehicle. This proves that if we use vehicles mainly under conditions of low speed and heavy load, we should give priority to four-wheel steering vehicles as heavy-duty driving tools, which can greatly increase safety. Four-wheel steering vehicles have a better settling time than front-wheel steering vehicles even in the event of an unexpectedly high road speed.

4 Conclusions

By modeling the 4WS vehicle in which the front-wheel angle is in proportion to and rear wheel angle and the FWS vehicle at different speeds, a simulation analysis of the transient response before and after the change in the center of mass position is made. The results are as following:

1. With the same parameters of vehicle, the steering stability of the 4WS vehicle is better than that of the FWS vehicle. Although the reaction time of the 4WS vehicle has increased, the overall safety is still higher than that of the FWS vehicle.
2. The offset of the center of mass has a great influence on the stability of the vehicle. No matter 4WS vehicles or FWS vehicles, they can maintain stable at low speeds, but may sideslip or roll over at high speeds. At the same time, too much rearward movement will cause the reaction time of the car to prolong and greatly decrease the steering stability of the vehicle.
3. In low-speed driving conditions, the steering stability of 4WS vehicles after changing the center of mass is still better than that of FWS vehicles without changing the center of mass. It shows that if for some reason it is necessary to drive with a heavy load which will change the center of mass, the 4WS vehicle should be given priority to choose, and the vehicle safety under this condition is even higher than that of the FWS vehicle whose center of mass has not changed.

References

1. Zhang, B.: Study on lateral dynamics characters and control of 4WS vehicle, Tianjin University (2006)
2. Itoh, H., Oida, A., Yamazaki, M.: Numerical simulation of a 4WD–4WS tractor turning in a rice field. *J. Terramech.* **36**(2) (1999)
3. Liyun, L.I.: Simulation analysis of coupled automobile chassis suspension and steering system. *Adv. Mater. Res.* **3559**(1046) (2014)
4. Wu, L.Y., Xu, T.Q., Luo, T.: Research and realization of simulation for automobile steering resisting torque. *Appl. Mech. Mater.* **3830**(733) (2015)
5. Zhao, L.: Vehicle braking stability analysis in turn condition. *Appl. Mech. Mater.* **3343**(607) (2014)
6. McCluggage, D.: Four-wheel steering redux. *Autoweek* **64**(3) (2014)

7. Li, M., Jia, Y.: Precompensation decoupling control with H_{∞} performance for 4WS velocity-varying vehicles. *Int. J. Syst. Sci.* **47**(16) (2016)
8. Li, M., Jia, Y., Matsuno, F.: Attenuating diagonal decoupling with robustness for velocity-varying 4WS vehicles. *Control Eng. Pract.* **56** (2016)
9. Liu, Q.J., Chen, J., Zhang, W.: The simulation of four wheel steered car run in close-loop driving conditions. *Appl. Mech. Mater.* **3082**(543) (2014)
10. Oksanen, T., Linkolehto, R.: Control of four wheel steering using independent actuators. *IFAC Proc.* **46**(18) (2013)



Reflectors for Multiple Applications Based on Flux Compensation Method

Wang Guangzhen^{1(✉)}, Hou Yu², and Li Jia¹

¹ Foundation Department, Tangshan University, Tangshan, Hebei 063000, China

zhen_g_wang@163.com

² Department of Physics, Tangshan Normal University, Tangshan 063000, China

Abstract. The reflector of one surface for Light-emitting diode (LED) is designed based on flux compensation method. It is easier to be fabricated compared with freeform lens. Moreover, the reflector can effectively reduce material absorption loss. This design method is simple and other researchers can understand and repeat more simply. The designed reflector also has a small exit angle for uniform illumination. It has high illuminance uniformity and flux efficiency, which are more than 90% and 95%, respectively. Reflectors designed here can be used in optical signal coupling, display cabinets, and other optical transmission applications.

1 Introduction

At present, LED lighting components [1–4] have been used in various applications, including Backlight [5], headlamp and projector [6]. But the research of LED reflector is more difficult because LED has special emitting. Usually, lenses have been used in LED system widely for uniform illumination. But the lens has several refractive surfaces, which increase absorption loss and scattering loss greatly. In order to solve the above questions, a new type of reflectors are designed here, which have higher efficiency and uniformity. With the improvement of LED source, research and application of LED lighting [7, 8] develop very fast in recent years. The main design methods of the reflectors are as follows: SMS design method, various optimization design methods, energy mapping method [9, 10] and so on. SMS design method uses multiple surfaces at the same time so it is more difficult to master. Optimization design method needs to adjust parameters repeatedly, which is time-consuming. In addition, most of these methods only shows the design process but the details of these methods haven't been described. In this article, the surface's coordinates are calculated by solving equations simply. First of all, theoretical illuminance of the illuminated surface was obtained by the Square Inverse Law and the Cosine law. And then, the coordinates of the reflector

Wang Guangzhen—Masterminded EasyChair and created the first stable version of this document.
Li Jia—Created the first draft of this document.

were solved by the energy compensation method. Finally, the effects of these reflectors with different exit angles were studied.

2 Design and Simulation

2.1 Theoretical Modeling

For a single-surface reflector, the total flux from the LED source is usually distributed in the following four styles shown in Fig. 1. Parameters r_1 and r_n are the lengths from the two edge points of the reflector to LED, respectively. Parameter θ_m is half of the exit angle and y_m is the radius of the illuminated plane. Reflectors designed by Fig. 1b–d have larger size compared to reflectors designed by Fig. 1a. So Fig. 1a is used here to introduce the calculation process. In this style, the light rays are reflected in a crossway.

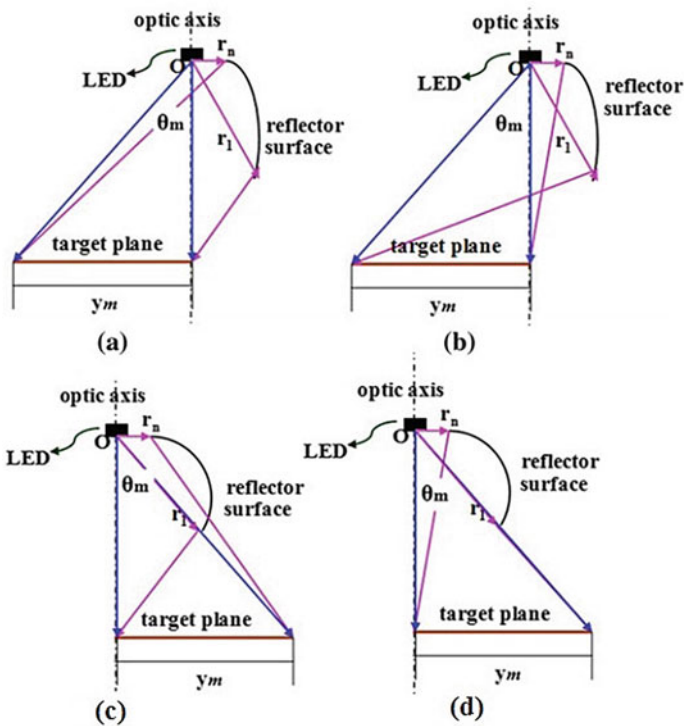


Fig. 1. Four types of flux distributions

Figure 2 is a diagram of LED reflection system. Currently, Lambert LED source is the most popular light source in the market. So it is used here for designing reflectors. The light intensity at the direction of θ for Lambert LED source is:

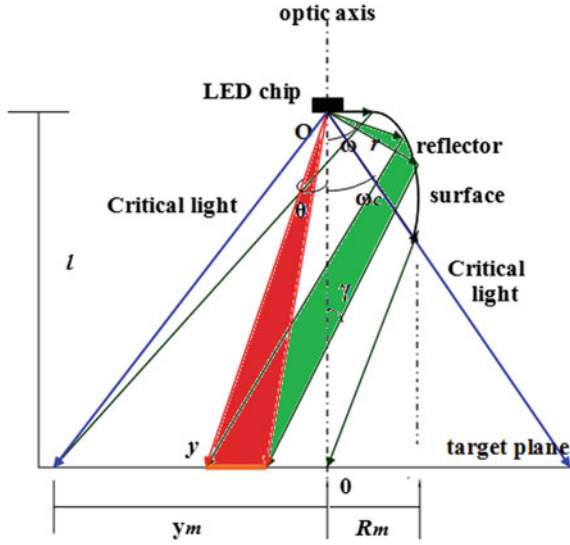


Fig. 2. Sectional drawing for the lighting system

$$I_\theta = I_0 \cos \theta \tag{1}$$

where I_0 is the light intensity in the direction of optic axis ($\theta = 0$). It is located at the center of the reflector's bottom for a rotational and symmetry system. The energy of LED is divided into two parts: the part (emitting angle is θ) which incidents directly to the illuminated plane and the part (emitting angle is ω) which is reflected to the illuminated plane. The angle of the critical light ray is ω_c . Simultaneously, it is the ray which goes to the edge of the illuminated plane. The vertical distance from LED to the illuminated plane is l . The reflector has the maximum radius of R_m . The polar radius of arbitrary point on the reflector's surface is r .

According to the Square Inverse Law, the illuminance in the plane, which is perpendicular to the irradiation direction, is proportional to I_θ . Simultaneously, this illuminance is inversely proportional to the distance's square, which is also called the normal illuminance. It is shown in Eq. (2).

$$E(l, \theta) = I_\theta / (l / \cos \theta)^2 = I_0 \cos^3 \theta / l^2 \tag{2}$$

According to the Cosine Law, the illuminance in the horizontal surface is proportional to I_θ and $\cos \theta$. Simultaneously, this illuminance is inversely proportional to the vertical distance from LED to the illuminated point. This illuminance is also called the horizontal illuminance which is shown by Eq. (3).

$$E(y) = I_\theta \cos \theta / (l / \cos \theta)^2 = I_0 \cos^4 \theta / l^2 \tag{3}$$

A very important issue must be noted: $E(y)$ has the biggest value $E(0)$ at $\theta = 0$.

Here, uniform horizontal illuminance is needed. If there is no energy loss, the average illuminance is written as:

$$E_{ave} = \phi/S = \pi I_0/\pi y_m^2 = I_0/(l \tan \omega_c)^2 \quad (4)$$

where ϕ is the total flux from LED. Parameter S is the area of the illuminated plane. Therefore, the illuminance that the reflector should provide at the central position $(y, 0)$ is:

$$E(y)' = E_{ave} - E(y) = I_0/y_m^2 - I_0 l^2/(l^2 + y^2)^2 \quad (5)$$

It is clearly impossible to use Eq. (5) if $E(0) > E_{ave}$. So a single-surface reflector has a critical angle: $\omega_c = \theta_m = 45^\circ$ with $E(0) = E_{ave}$.

For the light rays reflected by the reflector, $\omega_c \leq \omega \leq \pi/2$, the coordinate y of the illuminated plane is obtained using Eq. (6) and flux conservation (7).

$$\begin{aligned} I_0 \cdot 2\pi \cdot \cos \omega \sin \omega d\omega &= E'(y) \cdot 2\pi y dy \\ &= [I_0/y_m^2 - I_0 l^2/(l^2 + y^2)^2] y dy \end{aligned} \quad (6)$$

After integration, Eq. (7) is got

$$\sin \omega^2 \Big|_{\omega_c}^{\omega} = y^2/y_m^2 \Big|_0^y + l^2/(l^2 + y^2)^2 \Big|_0^y \quad (7)$$

It can be written as

$$(1/y_m^2)y^4 + (l^2/y_m^2 - 1 - \sin^2 \omega + \sin^2 \omega_c)y^2 + l^2 \sin^2 \omega_c - l^2 \sin^2 \omega = 0 \quad (8)$$

If $A = 1/y_m^2$, $B = l^2/y_m^2 - 1 - \sin^2 \omega + \sin^2 \omega_c$ and $C = l^2 \sin^2 \omega_c - l^2 \sin^2 \omega$, y can be solved by Eq. (9).

$$y = [(-B + (B^2 - 4AC)^{1/2})/(2A)]^{1/2} \quad (9)$$

Combining Fig. 2, it can be obtained the relationships between r, ω and y in Eqs. (10) and (11).

$$F = dr/r d\omega = \tan[(\pi - \omega - \gamma)/2] \quad (10)$$

$$\tan \gamma = (r \sin \omega + y)/(l - r \cos \omega) \quad (11)$$

They can be solved by Euler's formula. Where F is the function used in solving Euler's formula.

2.2 Simulation and Analysis

All coordinates (y, ω) of the reflector's are solved by Eqs. (10) and (11). Then these values are imported into the software to form an entity. The obtained reflector's shape is shown in Fig. 3. The θ_m and l of the reflector are 26.6° and 2 m, respectively.

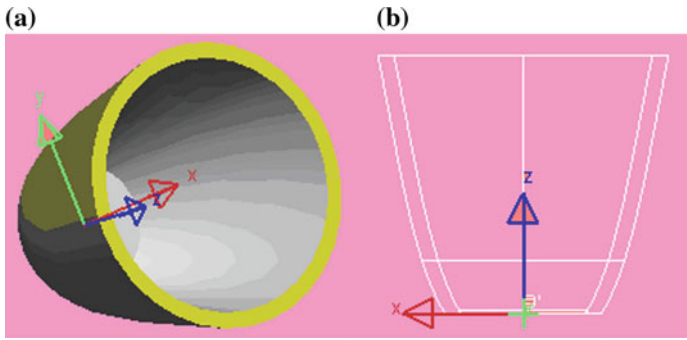


Fig. 3. Entity and sectional view of reflector

An optical system including light source, reflector and the illuminated plane is constructed for simulation and analysis. A $1\text{ mm} \times 1\text{ mm}$ LED chip is used in the simulation. It's flux is 100 lm and the incident wavelength is 0.55 micron. One hundred thousand rays are traced on Monte Carlo method. The average illuminance in the illuminated plane is 32.2 lx. The illuminance distribution is shown in Fig. 4.

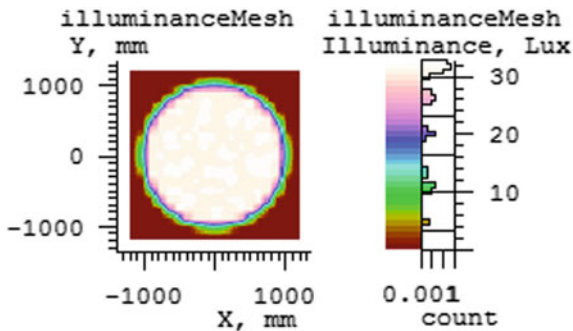


Fig. 4. Schematic diagram of illuminance

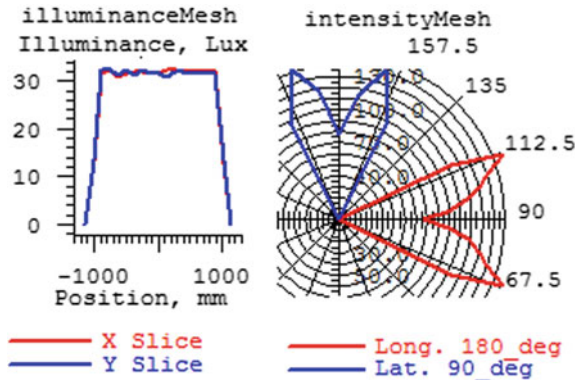


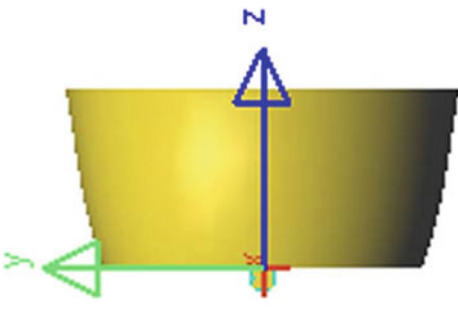
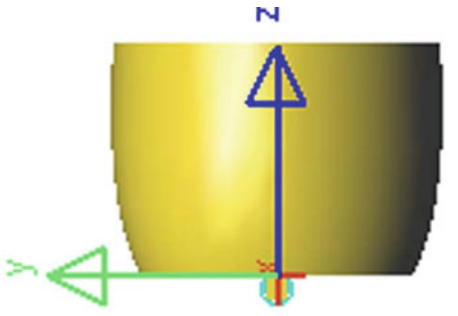
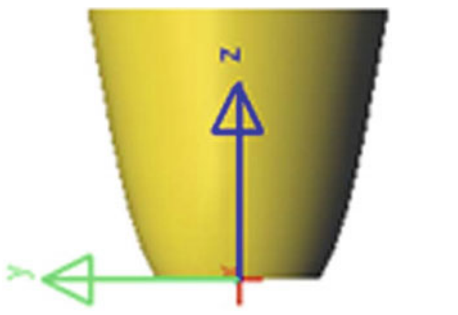
Fig. 5. Line charts of illuminance and light intensity

Figure 5a and Fig. 5b are line charts of illuminance and intensity, respectively. The illuminance uniformity, which is the ratio of minimum illuminance to average illuminance, is more than 95%. Ignoring all the absorption and scattering loss, flux efficiency, which is defined as the ratio of flux in the illuminated plane to that from LED source, is about 99%. In Fig. 5b, the light intensity distribution is bat-wing type, which is the feature of uniform lighting.

In addition, the reflector designed on this method can achieve miniaturization and has good lighting effect. For example, a reflector with $\omega_c = 22.6^\circ$ is designed. The top and bottom diameters are $\phi = 10$ mm and $\phi = 7.6$ mm, respectively. The average illuminance in the illuminated plane is 322200 lx. It's small size makes it be used as a package reflector. Many instrument lighting and signal coupling lighting may use these reflectors. Table 1 shows detailed information of reflectors with different exit angles (a) $\omega_c = 45^\circ$, (b) $\omega_c = 22.6^\circ$ (c) $\omega_c = 22.6^\circ$ and (d) $\omega_c = 11^\circ$. "Ave E" represents the average illuminance in the illuminated plane. Table 1 reveals that the sizes of the reflectors which can be adjusted according to special need. They all have high illuminance uniformity and efficiency.

Taking Table 1a, b and d as an example, the relationship between LED size and illuminance uniformity is studied. The simulations use square LED chip. It is found that the LED size has no significant influence on the lighting efficiency. The uniformity has little difference which is shown in Fig. 6. It also can be seen that, if 2.5 mm × 2.5 mm LED chip is used, the uniformity decreases slightly. But the reflectors still meet requirements of many application cases.

Table 1. Comparison of reflectors with four exit angles

	(a) $\omega_h = 45^\circ$	(b) and (c) $\omega_h = 22.6^\circ$	(d) $\omega_h = 11^\circ$
Side view			
Application	Reading	Coupling	Stage
l (mm)	200	20	5000
y_m (mm)	200	10	1000
R_m (mm)	20	10	60
Ave E (lux)	794	322200	88.4
Uniformity (%)	98.3	94.4	96.1
Efficiency (%)	99.76	99.9	99.9
		Courtyard	
		2000	
		1000	
		80	

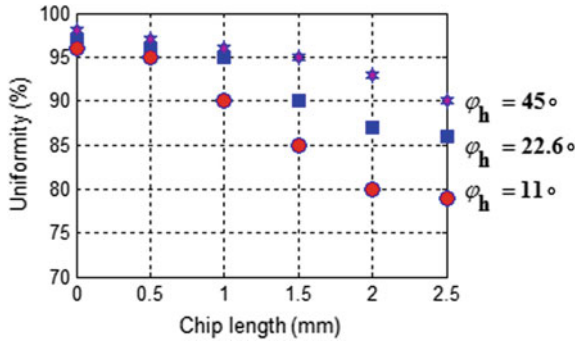


Fig. 6. LED size’s influence on uniformity for reflectors

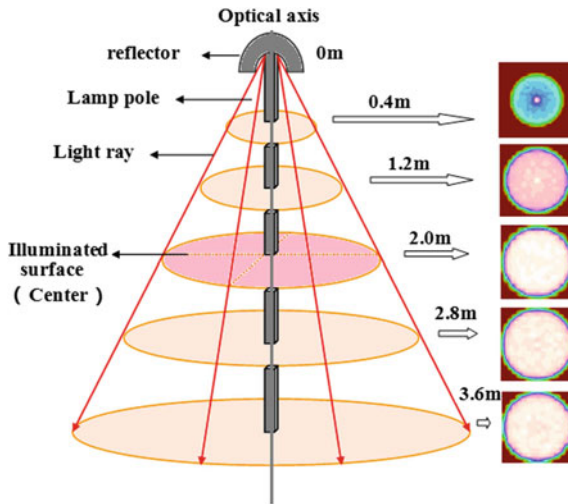


Fig. 7. Illumination depth of the reflector

The illumination depth is a specific parameter must be considered in many applications. It is defined as the length between two special points along the optical axis. The illuminance at the two special points reaches 20% of that in the center illuminance. Figure 7 shows the illuminance distributions for Table 1c at several distances along

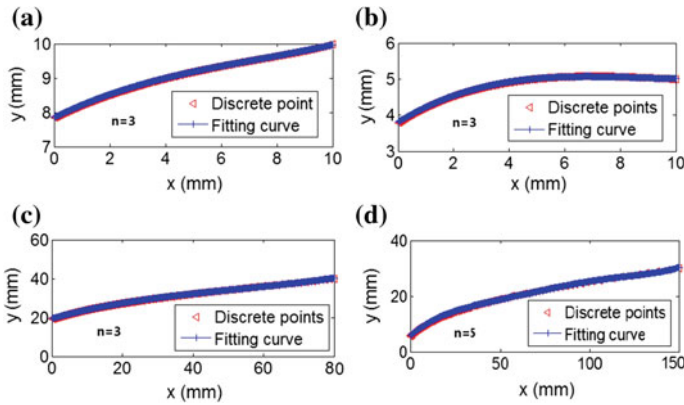


Fig. 8. The fitted lines of reflectors in Table a–d

optical axis. According to flux conservation $E\pi y_m^2 = 20\% E(\pi y'_m)^2$, when $y_m = 1$ m we get $y'_m = (5y_m)^{1/2}$ and $l = 2y'_m = 4.47$ m. In Fig. 7, when $l = 1.2$ m, illuminance uniformity is also high in the illuminated plane. So the illumination depth of this reflector can reach $4.47 - 1.2 = 3.27$ m. It is noteworthy that uniformity declines correspondingly if the distance reduces.

Another feature of the reflectors designed using above method is that their surfaces can be fitted on polynomial fitting method. The polynomial function can be expressed as follows:

$$y = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \tag{12}$$

Figure 8 shows four curves in Table 1. The fitting curves and the discrete points coincide perfectly, and n is the maximum value in the fitting equations. It can be seen that the curve, which has the smallest exit angle and the maximum n .

Table 2 lists the polynomial function’s coefficients of the four fitting curves in Fig. 8.

3 Conclusion

A kind of single-surface reflectors is designed on the simple method based on flux conservation method. Several reflectors’ models are obtained. The sizes, performances and other parameters of these reflectors are compared and analyzed. It is proved that

Table 2. The coefficients of the polynomial equations

	a_5	a_4	a_3	a_2	a_1	a_0
Figure 8a	0	0	0.00157691779588832	0.0346163407075388	0.400831969389786	7.8491551728743
Figure 8b	0	0	0.00213774790130996	0.0562779201005193	0.469187152686452	3.81030331647692
Figure 8c	0	0	3.93505797570177e-05	0.00618564153959688	0.503736532382858	19.4851400692772
Figure 8d	2.63543633986824e-09	1.07002001104112e-06	0.000162897243676495	0.0119387830932206	0.566809293028201	5.78481160283112

this kind of reflectors can realize a wide range of illumination with good illuminance uniformity and efficiency. In addition, the reflector has a great illumination depth. They also can be directly used for lighting and optical signal transmission.

Acknowledgements. This work was supported by the Science and Technology Projects of Hebei Province (Grant No. 16274522) and the Science and Technology Projects of Hebei Province (Grant No. 17210403).

References

1. Jae Suk Yang, A., Park, J.H., Beom-Hoan, O., Park, S.G., Lee, S.G.: Design method for a total internal reflection LED lens with double freeform surfaces for narrow and uniform illumination. *J. Opt. Soc. Korea* **20**, 614–622 (2016)
2. Wang, Xiaoli: LED ring array light source design and uniform illumination properties analysis. *Optik* **140**, 273–281 (2017)
3. Golmohammadi, S., Rabbani-Shabestari, S.J.: Design of white LED using GaN/InxGa (1 - x)N multiquantum well. *Optik* **126**, 5820–5824 (2017)
4. Wang, G., Wang, L., Li, F., Zhang, G.: Collimating lens for light-emitting-diode light source based on non-imaging optics. *Appl. Opt.* **51**, 1654–1659 (2012)
5. Lee, T.X., Chen, B.S.: High uniformity and tolerance design for direct-lit LED backlight illumination using lagrange interpolation. *J. Disp. Technol.* **12**, 1403–1410 (2016)
6. Sun, W.S., Tien, C.L., Lo, W.C., Chuo, P.Y.: Optical design of an LED motorcycle headlamp with compound reflectors and a toric lens. *Appl. Opt.* **54**, E102–E108 (2015)
7. Moiseev, M.A., Dосkolovich, L.L.: Design of TIR optics generating the prescribed irradiance distribution in the circle region. *J. Opt. Soc. Am. A* **29**, 1758–1763 (2012)
8. Tsai, C.Y.: Free-form surface design method for a collimator TIR lens. *J. Opt. Soc. Am. A* **33**, 785–792 (2016)
9. Wanga, G., Hou, Y., Hu, L., Tang, W., Gao, J., Wang, L.: Reflector designed for light-emitting-diode lighting source in three-dimensional space. *Optik* **126**, 4534–4538 (2015)
10. Prins, C.R., Boonkamp, J.H.M.T.T., Roosmalen, J.V., Ijzerman, W.L., Tukker, T.W.: A Monge–Ampère-Solver for free-form reflector design. *SIAM J. Sci. Comput.* **36**, B640–B660 (2014)



The Design of Low-Power and High-Precision Electronic Scale Based on Single Chip

Dun Liu and Fang Qu^(✉)

School of Electronic Engineering, Tianjin University of Technology and Education (TUTE), Tianjin 300222, China
dtqufang@126.com

Abstract. This paper introduces the design and fabrication of a low-power and high-precision electronic scale with resistance strain gauge as the weight sensor. According to the demand of commercial measurement, the paper compares and proves the various design schemes, chooses the cantilever-type resistance strain sensor with the STC89C52 single chip, and uses the 24-bit high-precision A/D converter chip HX711 to convert the signal to A/D. The system has the advantages of high integration, fast response speed, high precision, low power consumption, and strong anti-interference, which lowers the cost and improves the performance and reliability of the whole machine.

Keywords: Single-chip · Strain gauge · Electronic scale · HX711

1 The Composition and Scheme Demonstration of the System

1.1 The Composition and Working Principle of the System

The electronic scale system designed in this paper can be divided into six parts, such as SCM control circuit [1–3], sensor circuit, A/D conversion circuit [4], LCD display, matrix keyboard input, and buzzer alarm module, and its system composition diagram is shown in Fig. 1.

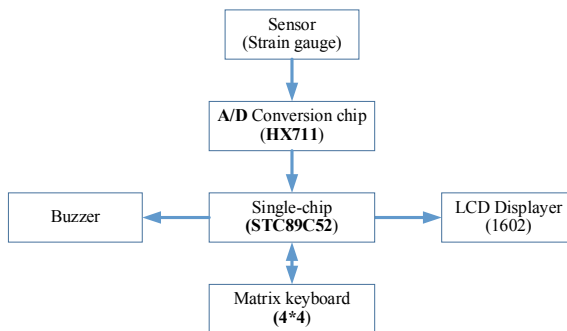


Fig. 1. System chart

The system converts the pressure to the electrical signal through the sensor, that is, after the resistance strain gauge in the sensor is induced to the pressure, the resistance will change slightly. Through the full bridge measurement circuit, the tiny change of the resistance is converted to a small change of the voltage. HX711 [5] is a high precision 24 bit A/D converter chip, and the signal is converted to A/D, and the electrical signal is amplified 128 times, and then the collected 24 bits of data are sent to the single-chip processor through its D_{OUT} end for processing. The MCU transfers the processed data to the LCD display circuit and outputs the measurement results by the display.

1.2 The Scheme of the Core Part of the System

(1) Control part of the scheme

The system adopts 89C52 as the control core. 51 MCU has 12 MHz main frequency, 32 I/O ports, 8 K flash program storage space, 256 byte RAM storage, three timings/counters and five interrupt sources, and its price is low, the C language program is easy to write, the control is convenient, the program can be transplantable, and the precision requirement can be reached.

(2) The scheme of the A/D conversion part

The system uses HX711 as the chip of A/D conversion, which integrates low-noise amplifier internally that can be used to amplify and collect pressure signals. HX711 is a 24 bit A/D converter chip designed for high-precision weighing sensors. The chip occupies less I/O port, has simple circuit and high precision. It is very suitable for data acquisition and A/D conversion of resistance strain gauge pressure sensor.

2 Theoretical Analysis and Calculation

2.1 Theoretical Analysis and Calculation of Piezoelectric Sensors

The working principle of the resistance strain gauge is based on the resistance strain effect, that is, when the conductor produces mechanical deformation, its resistance value changes accordingly [6]. The cantilever-beam force sensor is adopted in this design. The sensor has simple structure and high sensitivity and is suitable for small pressure measurement.

With a resistor wire under the action of external force, it will cause resistance change ΔR , and there is

$$\frac{\Delta R}{R} = \frac{\Delta l}{l} - \frac{\Delta S}{S} + \frac{\Delta \rho}{\rho} \quad (1)$$

In Formula (1), ρ is resistivity of resistance wire; l is length of electrical resistance wire, and S is area of resistance wire.

Supposing the axial effect of the resistance wire is $\varepsilon = \Delta l/l$, it is known from the mechanics of materials:

$$\Delta r/r = -\mu(\Delta l/l) = -\mu\varepsilon \quad (2)$$

Poisson's coefficient of the resistance wire material in the Formula (2) is μ , after finishing we can get:

$$k_0 = \frac{\Delta R/R}{\varepsilon} = (1 + 2\mu) - \frac{\Delta\rho/\rho}{\varepsilon} \quad (3)$$

From the Formula (3), it is known that the sensitive coefficient k_0 is composed of two parts, $(1 + 2\mu)$ indicates the change of the geometrical dimension of the material after the force, and $\frac{\Delta\rho/\rho}{\varepsilon}$ indicates the change of the material resistance. For metal materials, the $\frac{\Delta\rho/\rho}{\varepsilon}$ term is negligible, so $k_0 = 1 + 2\mu$. The experiment shows that the relative change of resistance is proportional to the ratio of resistance to wire in the tension limit of resistance wire, that is, $k_0 = 1.7-3.6$. Therefore, the Formula (3) can be further expressed as

$$\frac{\Delta R}{R} \approx k_0\varepsilon \quad (4)$$

2.2 Theoretical Analysis of Reverse Conversion of Gravity Value by AD Value

According to the design requirements: the measurement of the full range of output voltage = excitation voltage * sensitivity (1.0 mV/V). If the supply voltage is 5 V, there is 5 V * sensitivity (2.0 mV/V) = full range 10 mV, which is equivalent to the voltage of 10 mV when 10 kg gravity is generated.

Suppose the actual weight is A Kg, the value measured by the sensor is Y; if the pressure sensor outputs is 8.6 mV at 10 kg, the voltage sent to the A/D module is: A Kg * 8.6 mV/10 kg = 0.86 * A mV; after 128 times amplification, it is 128 * 0.86 * A mV \approx 110 * A mV, then convert it to 24 bit digital signal, it is Y = 110 * A mV \times $2^{24}/4.3$ V \approx 429,185 * A (full-scale output is 4.3 V voltage). So, we can get:

$$\begin{aligned} A &= Y/429185 \text{ Kg}/100 \\ &\approx Y/4.30\text{g} \end{aligned} \quad (5)$$

3 System Hardware Design

3.1 System Structure

The hardware structure of this system is shown in Fig. 1. The main modules are the MCU STC89C52RC [7] minimum system control circuit module, the resistance strain pressure sensor module HL-8, the high precision and gain HX711 A/D conversion chip module for the electronic scale, the 1602LCD LCD module and the buzzer alarm module.

3.2 Introduction of Main Control Chip STC89C52 Single-Chip Microcomputer

STC89C52 is a low-power, high-performance CMOS 8-bit microcontroller, also suitable for conventional programming. With a smart 8-bit CPU and a programmable Flash memory unit in the system, STC89C52 provides a flexible and effective solution for many embedded control applications.

3.3 Design of HX711 Interface Circuit for A/D Converter Chip

According to the design requirements, the system requires the output current signal to be 20–1000 mA, the step is 1 mA, and the numeric value display is required. Therefore, at least 12 bits of conversion precision is required for the quantitative actuating element A/D converter. Considering the design requirements of the system and considering the constraints of the I/O interface resources of the single-chip microcomputer, the final determination is to select the HX711 whose quantization precision can reach $1/4096 < 1/1000$, and the precision requirements of the design can be reached completely. The HX711 interface circuit is shown in Fig. 2.

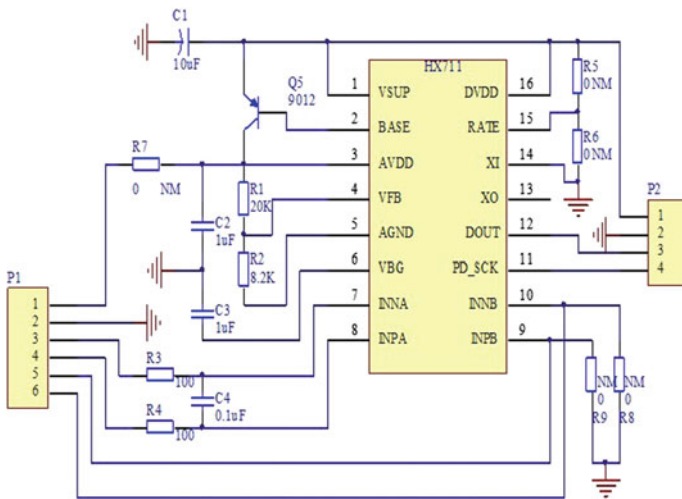


Fig. 2. Interface drawing of the HX711

The input selector switch can optionally select either channel A or channel B and be connected to its internal low-noise programmable amplifier. The programmable gain of channel A is 128 or 64, the corresponding full amount difference input signal amplitude is ± 20 mV or ± 40 mV, respectively. Channel B is a fixed 64 times gain for system parameter detection.

3.4 Design of Piezoelectric Sensor

The common measuring circuits for strain sensors are single arm bridge, differential half bridge, and differential full bridge, in which the differential full bridge can improve the sensitivity of the bridge, eliminate the nonlinear error of the bridge, and eliminate the common mode interference such as temperature error. In general, 4 strain gauges are used in the measurement to form a differential full bridge. The sensor used in this design is the full bridge measuring circuit. Its circuit diagram is shown in Fig. 3. The bridge-type measurement circuit has four resistors, of which anyone can be resistance strain gauge, one diagonal of the bridge is connected to the working voltage, and the other is diagonal output voltage. Its characteristics are: when the resistance of the four bridge arms reaches the corresponding relation, the output of the bridge is

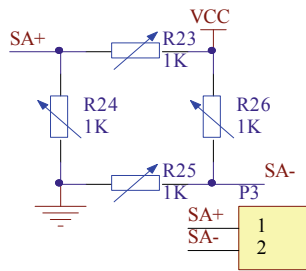


Fig. 3. Full bridge measuring circuit

zero, or the output of the voltage, which can be measured by using the sensitive galvanometer, so the bridge can accurately measure the small resistance changes. The strain resistance is used as the bridge arm resistance in the bridge circuit. When there is no pressure, the bridge is balanced and the output voltage is zero. When there is pressure, the resistance value of the bridge arm changes and the bridge loses balance. In the full bridge measuring circuit, two strain gauges with the same force properties are connected to the opposite side of the electric bridge. The output sensitivity is twice as high as that of the half bridge, and the nonlinear error and temperature error have been improved.

4 System Software Design

The software design of the system is mainly divided into system initialization, AD acquisition and operation, key and display processing modules.

4.1 System Main Program Flowchart

After the system is powered on, the main program completes the system initialization first, which includes the system variable definition and variable assignment initial value, and then call A/D collection function, to turn the A/D acquisition module output 24-bit binary serial data into the decimal, and then to adjust zero and calibration, and finally to separate four-digit decimal data hundred, 10-bit, Single-digit and decimal places, then call the liquid crystal display function, and send the corresponding values to the corresponding LCD display. The system main program flowchart is shown in Fig. 4.

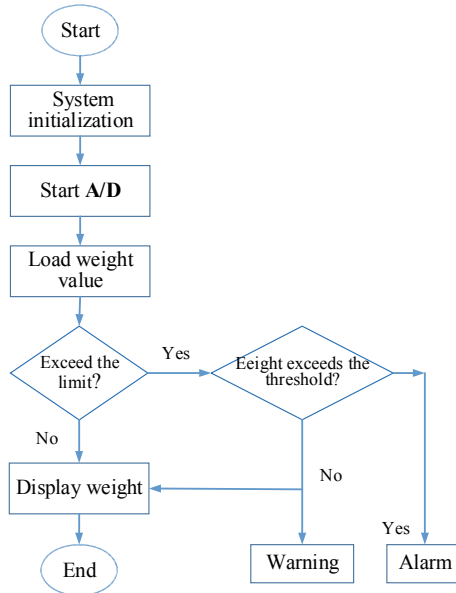


Fig. 4. Main program flowchart

4.2 Design of A/D Data Collecting Subprogram

As a chip for A/D conversion, HX711 is used to amplify and collect pressure signals which is a 24 bit A/D converter. The A/D data acquisition subprogram is mainly to collect the output signal of the piezoelectric sensor. The first 24 ADSK pulses collect

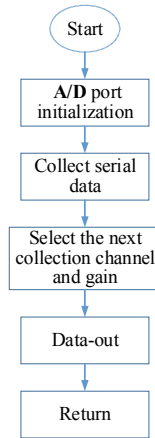


Fig. 5. A/D data acquisition flow chart

24-bit serial binary data, and the next 1–3 ADSK pulses choose the channel and gain of the next A/D acquisition, and use 1 ADSK pulses to select the channel A with the gain of 128. Its program flowchart is shown in Fig. 5.

5 Test Results

5.1 Test Scheme

First, check whether the welding of the circuit is correct and use multimeter to detect whether there is short circuit or open circuit in the circuit board. After confirming both are correct, then turn on the power, and the output voltage value of the power source should be measured by multimeter. After the test and it is confirmed everything is correct then it can be connected to all parts of the circuit. First press the reset button of single-chip microcomputer to reset the system and call the liquid crystal display function to display several numbers. After the display is normal, then connect the sensor and the conditioning circuit and the A/D conversion circuit made up of 24 bit serial A/D conversion chip and burn the whole program into the chip and observe whether the LCD display is 000.0. If it is not displayed, then need to adjust zero of the software, after the zero adjustment is completed, slowly give pressure to the sensor to see if the value of the liquid crystal display increases synchronously, and whether the value will return to 000.0 after releasing the pressure.

5.2 Test Results

Because the sensor is very sensitive, if the sensor is not fixed well in the horizontal direction, there will be some error. In addition, the sensor's lead is also very sensitive. It will cause errors even if it is touched slightly. The electronic scales is repeatedly tested

with the weight, and the error of the electronic scales is reduced by micro-adjusting the related parameters in the program according to the characteristic curve of the resistance strain gauge sensor given by the manufacturer. The test data is shown in Table 1.

Table 1. The test data

Weight (g)	The first measurement (g)	The second measurement (g)	The third measurement (g)	The average value (g)	The average error (g)
5	5	5.6	4.8	5.1	0.1
10	9.7	10.2	10.1	10	0
20	19.8	20.3	19.6	19.9	0.1
50	49.7	49.8	50.2	49.9	0.1
100	100.3	100.1	100.2	100.2	0.2
200	199.9	199.7	200.1	199.9	0.1
500	499.7	500.4	500.3	500.2	0.2

5.3 Error Analysis

We can see from the test data that the maximum error is 0.2 g, which accuracy is relative high, but it is found that factor such as the scales shaking, the position of the object in the scales, the vibration of the frame when it is placed on the heavy object, and the influence of the external magnetic field on the iron bracket, etc. can generate systematic error. A series of measures have been taken to control the system error, so as to design a reliable and accurate electronic scales system.

6 Conclusion

This design is based on the STC89C52 single-chip microcomputer to achieve a simple electronic scales, which can be used to measure the weight between 5 and 500 g. The simple electronic scales with the resistance strain gauge as the pressure sensor is designed according to the demand of commercial measurement, after the comparison of each scheme, and the sensor chooses the cantilever-beam resistance strain sensor. With 24-bit high-precision A/D converter chip HX711 for A/D signal conversion, the scales have advantages of high integration, fast response, anti-interference, and so on, which reduces the cost of electronic scales, improves the performance and reliability of the whole machine. After analyzing the test results, the error is within the allowable range and the expected requirements are met.

Acknowledgements. This work was supported by the Tianjin Vocational Training Package Project—Professional training package for maintenance of household electrical appliances under Grant No. TJPXB1-1 (JB4071002).

References

1. Zhang, Z., Xiong, G.: Design of the multi-function electronic scale based on single-chip microcomputer. *Mech. Electron.* **34**(11), 58–61 (2016)
2. Chen, J.: A high precision electronic scale based on STM32. In: *Proceedings of 2017 2nd International Conference on Automation, Mechanical Control and Computational Engineering (AMCCE 2017)* (2017)
3. Gao, M., et al.: Development of a new intelligent electronic scale system. *Mod. Electron. Technol.* **40**(14), 40–43 (2017)
4. Yu, F., Li, Q., et al.: Design of an electronic scale based on HX711. *Sens. World* **22**(12), 33–36 (2016)
5. Yang, R.: Strain-gauge-based electronic scale experimental system design. *China Instrum.* **5**, 65–68 (2017)
6. Wang, R., Liu, X., et al.: Design of the electronic scale based on STC89C52RC MCU. *Foreign Electron. Meas. Technol.* **36**(05), 94–97 (2017)
7. Zhang, L., Chen, Y., et al.: Design of a simple electronic scale based on resistance strain gauge sensor. *Electron. Sci. Technol.* **04**(04), 96–100 (2017)



The Main Damage Characteristics of Semiconductor Devices Under High-Power Microwave

Kaibai Chen^(✉)

Shijiazhuang Campus of Army Engineering University, Shijiazhuang, China
3029417012@qq.com

Abstract. This paper takes the PIN diode as the main object of analysis, theoretically analyzes its operating principle and the type of damage occurring under the impact of the high-power microwave, and proposes that the damage characteristics of the device should be based on the time and energy.

Keywords: Damage characteristics · High-power microwave · PIN diode · Secondary breakdown effects

1 Introduction

The studies have shown it theoretically and experimentally that the more sophisticated, precisely, and integrated the electronic information system is, the stronger its electromagnetic susceptibility it is to high-power microwave destruction. Under the effect of HPM, the reliability of devices and circuits gradually highlights its importance. With the continuous development of related theories and technologies, the commonly used researches on the damage mechanism and the evaluation of HPM parameters on the interference of electronic systems have been studied. There are two main methods: experimental research or the theoretical analysis. It is intuitive to study the HPM effect of devices and circuits relying on experiments. However, the cost of the experimental analysis is obviously high, and it is difficult to sum up universality conclusions. Therefore, theoretical research on the damage characteristics of HPM remains necessary.

2 High-Power Microwave Technology and Its Application

What is HPM? HPM are strong electromagnetic pulses with peak power greater than 100 MW or average power greater than 1 MW and frequencies between 300 MHz and 300 GHz. HPM technology refers to interfere or damage electronic information systems through the use of high-frequency electromagnetic radiation, this technology is mainly focused on highly integrated electronic systems, and degrades or invalidates their function. In a word, it is a new concept of technology with a great potential application prospect to the future.

HPM has the following characteristics: high power, wide target universality, and the ability to perform different degrees of damage to the target from far to near within the beam coverage depth range; high frequency, easy to transmit, launch or couple and function with the target; wide spectrum, which can play multiple types of targets at the same time; high-speed, it emits microwave energy at the speed of light which makes the goal no time to escape; easy to use, HPM is less affected by weather conditions than high-powered lasers, and has relatively low requirements for targeting accuracy; lower cost, HPM relies on electric source to provide energy with a high cost-effectiveness ratio; and finally it destroys the electronic components inside the target, which cause no external damage.

With the development of transmission and emission technologies, HPM generation technologies and pulse power technologies, The HPM production efficiency, energy level, and effective radiation power are continuously increasing. At present, the power level of a single narrow-spectrum HPM source has reached tens of GW levels, and the power level of a single ultra-wide spectrum HPM has reached 100 GW. In the near and foreseeable future, the power level of a single HPM source will be higher, and its damage effect cannot be underestimated.

3 Damage Characteristics of Semiconductor Devices in HPM Environment

The following five mechanisms of HPM damage to electronic systems and semiconductor components [1]:

- (1) High voltage breakdown;
- (2) Device burned. It refers to the permanent damage of semiconductor devices, most semiconductor devices have a minimum power of 10 W, and some sensitive devices have only 1 W;
- (3) Microwave heating;
- (4) Transient pulse voltage impact. In addition to radiating HPM inside and outside the electronic system directly, the large current generated by the pulses on the device metal shield will surge on the shell. Any gaps such as voids or exposed leads will introduce surge current into the shell. It is enough to cause damage to sensitive devices under most circumstances;
- (5) Instant interference; Sometimes, although the inductive power entering the electronic system is too less to cause system damage, however, the transient current generated by the induction may cause instantaneous failure of the components and make the electronic system out of normal operation.

For example, the PIN diode's failure mechanism is mainly based on the secondary breakdown effect. There are two key types of damage about the secondary breakdown: the thermal secondary breakdown and the current secondary breakdown. The common principle to the two types of failure mechanism is that the device temperature has risen to a critical value after the heat accumulates and then occurs the secondary breakdown. In the end, the diode was completely destroyed.

3.1 PIN Diode Structure and Characteristics

The PIN diode belongs to a special microwave-sensitive semiconductor device which owns fast switching speed, high reverse breakdown voltage, large controllable power, small loss, and good short-circuit and open-circuit characteristics under the positive and negative biases. It is widely used in radio frequency or microwave circuit design [2].

Ordinary semiconductor diodes are composed of PN structure. The P region functions as an acceptor doping and the N region works as the donor doping. The PIN diode means to add an intrinsic layer I between the P region and the N region. Once the I region is implanted, the PIN diode's on-resistance will be greatly reduced compared with single-pole devices, which known as conductance modulation. When it comes to reverse bias, the reverse breakthrough will require a higher voltage due to the existence of the I layer, meanwhile the higher reverse breakdown voltage. So, the PIN diode won't generate a nonlinear rectifying effect on the microwave signal, which differs fundamentally from the conventional diode. Figure 1 is the specific structure of PIN diode.

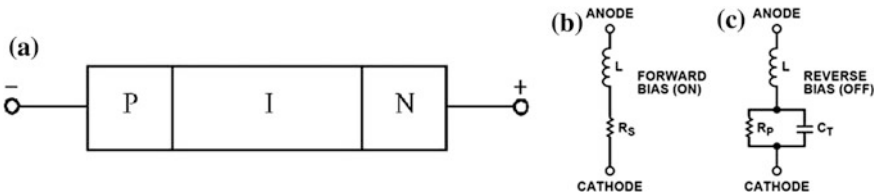


Fig. 1. PIN Diode and Equivalent Circuit

It can be seen from the figure that the PIN diode has the same forward and reverse characteristics as the PN junction, however, the PIN diode can withstand higher voltage than the PN junction when reverse bias is applied due to the presence of the I region.

Generally speaking, the reverse bias voltage will not grow unlimited, however, in some special circumstance the large reverse bias voltage will result in rapidly increasing reverse bias current. Under the action of electric field, the electrons and holes will obtain a large accelerated speed when they are crossing the space charge region, and meanwhile the rapid increase in speed leads to a certain amount of kinetic energy. When their kinetic energy reaches a certain level, they will hit the electrons inside the lattice atoms and produce new electron-hole pairs, the new electron-hole will hit the electrons in other atoms, it results in a series of chain reactions to produce a large number of electron-hole pairs. Since electron-hole pairs are produced at fast speed and in large quantities, it is just like an avalanche that nothing can stop it, so the breakdown is also called as avalanche breakdown.

3.1.1 PIN Diode Operation Characteristics

Excess carriers in the P and N regions will be injected into the I region when the PIN diode is applied with a forward voltage, along with the injection carrier and the composite carrier will tend to be equal, then the current reaches an equilibrium state.

Meanwhile, due to the accumulation of a large number of carriers, the resistance of PIN diode will become lower, the more the carriers inside the I region are, the lower resistance PIN diode will get.

Since the actual I layer contains a small amount of P-type impurities, when there is no voltage applied across the PIN diode, the holes in the I region will diffuse into the N region, and the electrons in the N region will diffuse into the I region, too, the situation happened the IN interface will then form the space charge region. The impurity concentration in the I region is low compared to the N region, so the depletion region is almost entirely in the I region. At the PI interface, due to the presence of a concentration difference (the concentration of the holes in the P region is much greater than that in the I region), the effect of the diffusion motion will remain, but it is smaller compared to the IN interface and most of the time can be ignored. Therefore, when the zero bias occurs, the PIN diode is in a high-impedance state because of the depletion region in the I area.

The reverse bias situation is similar to that of zero bias. The difference is that the built-in electric field will be strengthened. The effect is to widen the space charge region of the IN junction and mainly to expand it into the I region. The PIN diode at this time can be equivalent to the resistance plus capacitance. When the reverse bias voltage is too large, the depletion region is filled with the entire I region. I region punch-through occurs at this time, and the PIN tube does not work normally.

3.2 Two Types of Secondary Breakdown Effects

3.2.1 Thermal Secondary Breakdown

Figure 2 shows the voltage and current curves of a PN junction. The first breakdown produced by applying a negative bias is called Zener breakdown or avalanche breakdown. If the breakdown current caused by a negative bias increases continuously without control, the high temperature will probably exceed Si melting point (1688 K). In addition, the ohmic heating caused by the strong electric field can also lead to the metallized strip's melting which then will change into destructive secondary breakdown effect. Thermal secondary breakdown is mainly caused by high temperature in the semiconductor junction region. A large number of valence electrons in the semiconductor are going to be excited into carriers under the influence of high temperature, the increase of carrier current will enlarge the saturation current, and the increase of saturation current will further increase the temperature of the junction. Eventually, the semiconductor is ablated resulting in irreparable damage.

3.2.2 Current Secondary Breakdown

The current secondary breakdown, also known as avalanche secondary breakdown, occurs when the electric current caused by high voltage is much higher than the current made by thermal mode. Normally, the electric field of PN junction space charge region is controlled by the concentration of donor and acceptor inside the device, when the applied voltage is high, the avalanche effect of the depletion region will collide and ionize a large number of carriers, those electron-hole pairs move to both sides under the effect of the electric field in the depletion region, the concentration of carriers above the doping concentration determines the field strength in the depletion region, then the

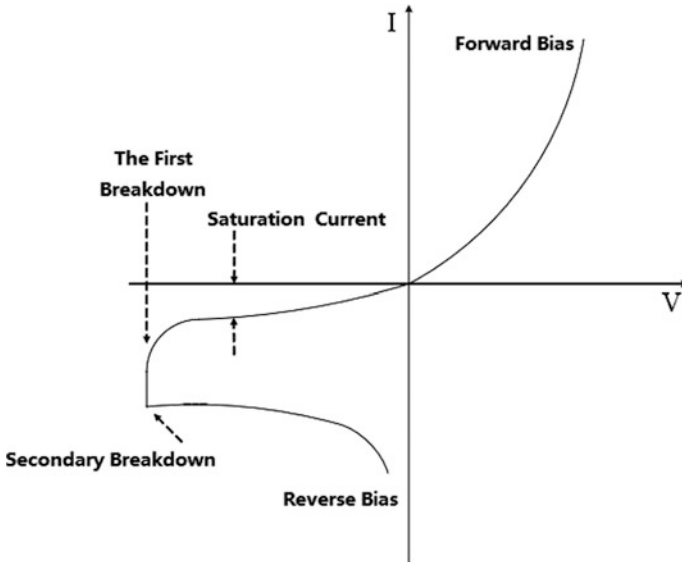


Fig. 2. Voltage and Current Effects of PN Junction

electron-hole pairs move to the edge of the depletion region and cause the peak of the electric field at the edge. Due to the high electric field strength, avalanche breakdown will occur at the edge of the depletion layer too, it generates dual injection of carriers. And if the velocity of carrier injection into the depletion region is bigger than that of the depletion region, the peak of the electric field in the depletion region will increase, eventually the thermal-current will be out of control and increase the current density and power consumption rapidly, generate a large amount of heat, and in the end increasing device temperature quickly. The strong current can also damage the device by burning metal strips elsewhere as with thermal breakdown. Along with the influence of negative resistance fluctuation, the current concentrates on a small area and form a current filament as long as the current density is too large, as a result, the rapidly increasing temperature will soon generate some hot spots which could cause device performance degradation as soon as the temperature exceeds 1000 K, once it goes to the silicon melting point, the short-circuit effect occurs in the PN junction area and the device will be permanently destroyed.

4 Conclusion

So it can be concluded that all types of breakdown effects are caused by semiconductors absorbing electromagnetic field energy. From the perspective of energy conservation, an electromagnetic field aroused by HPM can conduct its energy into semiconductor devices with thermoelectric effects by the form of coupling, radiation, etc. Once an irreversible physical and chemical change occurs, the device breakdown caused by the breakdown effect is also irreversible. Its voltage mainly falls on the

junction and the power is mainly released from the junction when a high-power microwave pulse impacts on a semiconductor device. One direction has been clearly pointed out that the main research goal should concentrate on analyzing both time and energy when studying the damage mechanism, determining the temperature variation of the electron-hole junction and the energy injection effect, and eventually making the analysis of the HPM radiation damage parameters, providing a clear theoretical basis for experimental research.

References

1. Schamiloglu, E.: High power microwave sources: where do we go from here? In: Conference Record of the Twenty-Fifth International IEEE Power Modulator Symposium, 2002 and 2002 High-Voltage Workshop. 30 June–3 July (pp. 190–191) (2002)
2. Doherty, W.E., Joos, R.D.: The PIN diode circuit designer' handbook. Microsemi Corp



A Configuration-Based Automatic Test System for Armored Vehicle Information Acquisition Devices

Linghui Zhang^{1(✉)}, Ruina Zhao², Lei Guo¹, Shao Li¹,
and Weizheng An¹

¹ China North Vehicle Research Institute, Beijing, China
{zhlh, guolei, lishao, anwzh}@noveri.com.cn

² Beijing North Vehicle Group Corporation, Beijing, China
zhaoruina2011@163.com.cn

Abstract. Against the situation that traditional armored vehicle information acquisition device can have multiple models, variable structures, long manual test cycle, complex test steps, low reliability, and other drawbacks, a configuration-based automatic test system for armored vehicle information acquisition devices is designed to realize the various performance testing. The system uses a combination of hardware and software virtual instrument technology which provides the ability to flexibly configure the test information. In addition, the test cases can be automatically configured and executed; the test results are automatically generated.

Keywords: Acquisition device · Automatic testing · Configuration-based

1 Introduction

Due to the increasing requirements in the performance of tanks and armored vehicles, the use of advanced electromechanical equipment is in great demand, and armored vehicles are developing in the direction of informatization. While the information system is an important part of tanks and armored vehicles, information acquisition device is the core electrical equipment of information system. The information acquisition device needs to collect the working status of various equipment and various types of sensor signals to ensure the normal operation of the entire armored vehicle system and that the system state is under full surveillance. At the same time, it provides the vehicle's fault diagnosis information for the crews to reduce the driving burden. For different types of tanks or armored vehicles, the classes and quantities of signals collected by the information acquisition device are disparate, and the correctness of the information collected by the acquisition device will directly affect the performance of the tank control system. Therefore, the test of the information acquisition device for tanks and armored vehicles is of great significance for the study of its performance and function, a stable and reliable information acquisition device and measurement/control system are necessary to guarantee the functions and action decisions of various parts of the vehicle. Due to the variety of information acquisition device types, the use of

different test benches to build the test circuit has high labor intensity and low working efficiency; the manual control test method is easy to produce misjudgment; the degree of automation is low and high skilled maintenance personnel are needed. With the development of computer, detection and electronic technologies, the development of universal, standardized and intelligent automatic test systems is the trend of automatic test systems at home and abroad. In order to examine the information acquisition devices more accurately, conveniently and quickly, the proposed system uses a configuration-based automatic test system to be able to configure different types of acquisition devices and achieve automatic testing, failing automatic alarms, indicating fault locations. It will greatly reduce the labor intensity of detection, and improve the accuracy of detection and working efficiency [1].

2 Overall System Design

According to the detection requirements of the information acquisition device, the system shall meet the following function requirements:

1. Complete the detection of various information acquisition devices on a single detection device;
2. When testing, the test signal can be introduced through the external socket of the information acquisition device;
3. When changed to another type of information acquisition device, only some simple settings in software are needed, no additional circuit is required to set up;
4. Have the ability to report test results automatically for operator's observation [2].

The major functions of the acquisition device include analog signal acquisition, switching signal acquisition, provision of stabilized voltage supply for sensors and external bus communication. The analog signal acquisition mainly includes voltage type and resistance type signals. Switching signal acquisition includes high and low-level signal, high-level with high-resistance signal, low-level with high-resistance signal, etc. The typical sensors' power supply voltages are 5, 12 and 24 V. The external communication of the acquisition device is mainly through CAN bus and FlexRay bus. All the collected and the external bus communication signals are attached to the device through the connectors on the case. Due to the variety of information acquisition device types, the analog signals, switch signals and external communication methods are different, moreover, the type of external socket and the number of connector pins are also not the same. In order to overcome the difficulty caused by type variety and meet the system requirements, a dedicated configuration-based automatic testing system is designed as illustrated in Fig. 1.

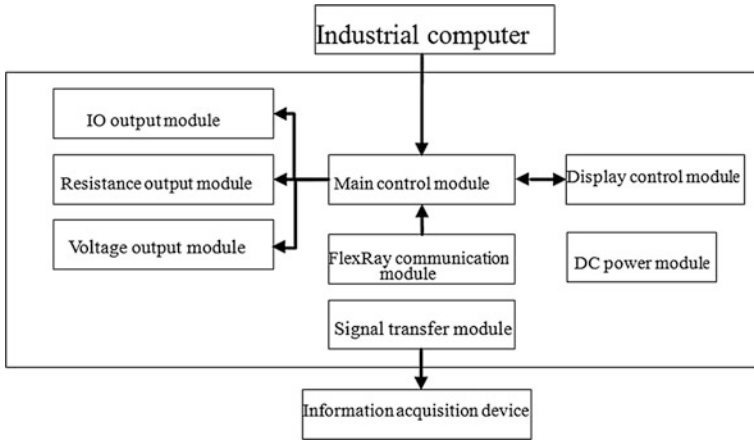


Fig. 1. Overall system composition

3 Main Modules of the System

3.1 Design of the Test Master Module

The test master control module, shown in Fig. 2, uses ARM as the core control chip, which has the advantages of high running speed and powerful functions. The test master module can receive and save the test cases set up by the industrial computer that will be executed by default before the next overwriting. The test master module will control the switch output module, the resistance signal output module as well as the voltage signal output module to generate the simulated test signals according to the test cases configured by different channel tests [3]. The test master module can also control the communication between the communication module and the acquisition device and judge the communication state of the acquisition device, meanwhile, the simulated test signals are compared with the collected signals uploaded by the acquisition device to verify the correctness of the information from the collection channel so that to form a closed-loop test. At last, the test master control module drives the display control module to show the test results.

3.2 Design of the High-Level, High-Resistance Signal Test Circuit

The high-level output adopts the BTS721L1 chip, the scheme is illustrated in Fig. 3, this chip's high-level part can have very large current output while the low-level part does not support input current function, and the chip's package structure is shown in Fig. 4. The high-level output circuit can be used in the test of the high or low-level signals and high-level with high-resistance signals.

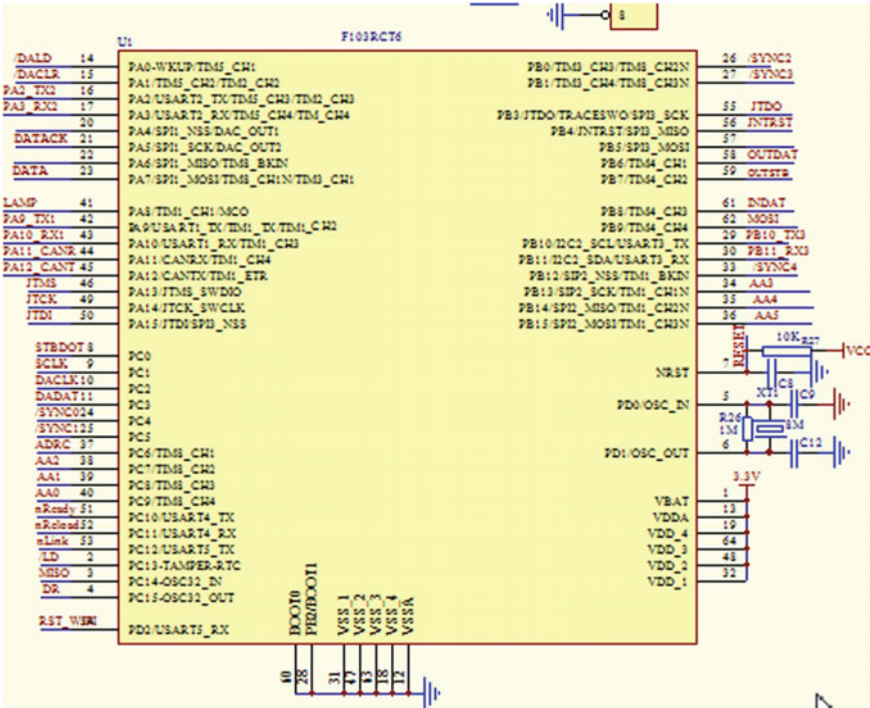


Fig. 2. The scheme of the control chip

3.3 Design of the High-Level, Low-Level Signal Test Circuit

As shown in Fig. 5 the electric relay RY4 outputs the OUT signal under the control of the control system CTRL. When the relay is not energized, OUT is connected to ground, and when the relay is energized, OUT outputs 24 V through a current limiting resistor.

It can be seen from the scheme that under the mode of relay output the circuit can output high-level and low-level, and both high level and low level have the ability of electric driving. It can test the high-level and low-level collection circuit of the information acquisition device.

3.4 Design of the Voltage Acquisition Signal Test Circuit

As shown in Fig. 6, the AD5724 DA chip generates four voltage outputs under the control of the CPU. After the output voltage signal passes through the amplifier, a current limiting resistor is added and then the circuit is able to simulate the output of the sensor for testing the voltage signal of the acquisition device. The current limiting resistor in the figure prevents the user from misconnecting and protecting the DA chip.

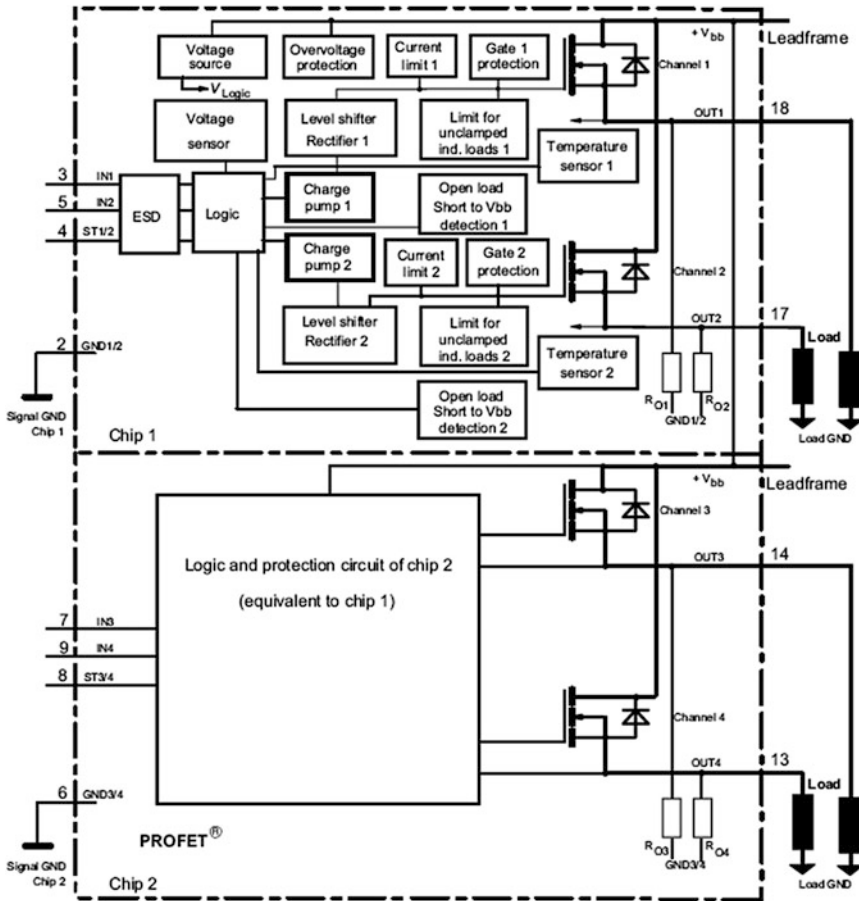


Fig. 3. Scheme of the high-level with high-resistance test circuit

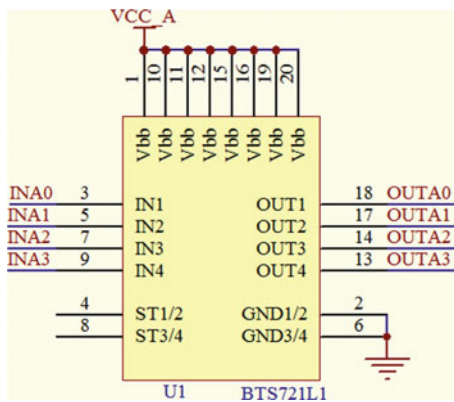


Fig. 4. Package structure of the BTS721L1 chip

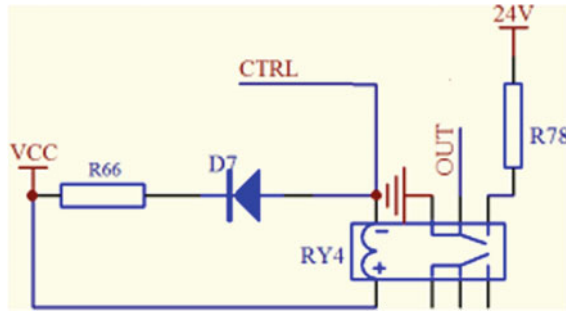


Fig. 5. The principle of high-level and low-level signal test circuit

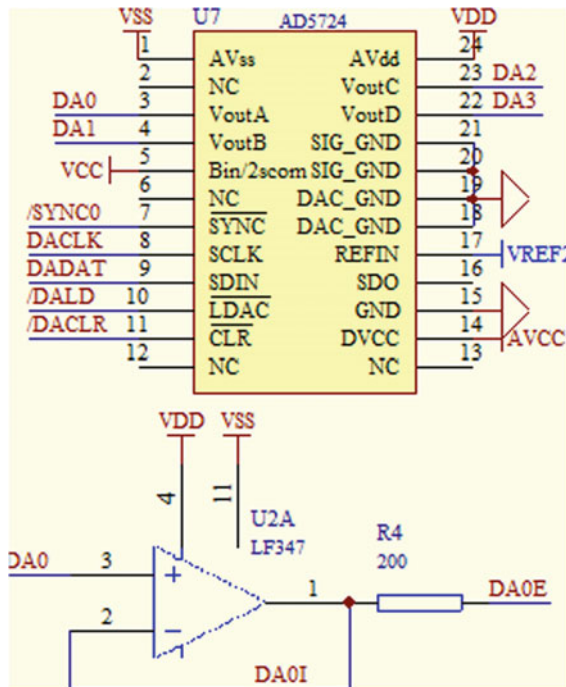


Fig. 6. Voltage signal test circuit

3.5 Design of the Resistance Signal Test Circuit

As shown in Fig. 7, the imported high-performance relay controls the opening and closing of the resistor network so as to achieve the output of the standard resistor and achieve the purpose of simulating the temperature sensor.

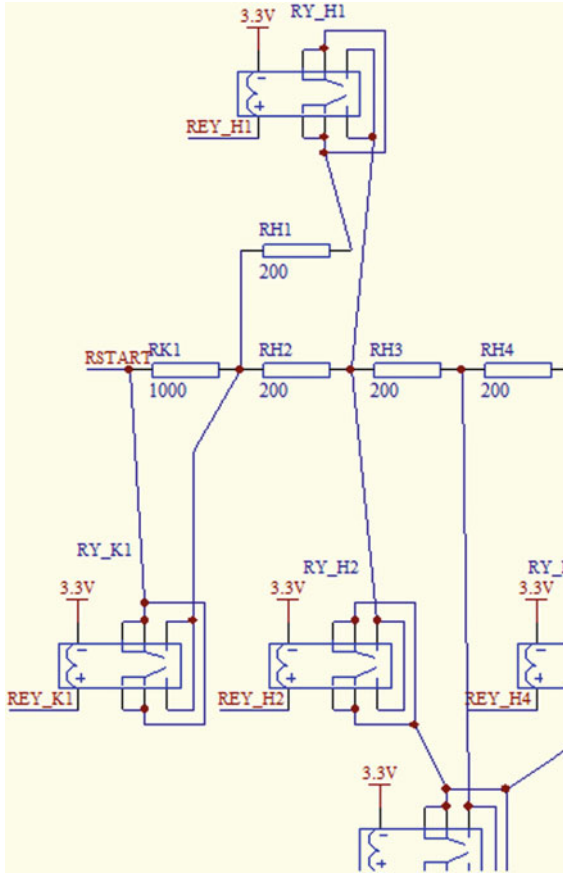


Fig. 7. Resistance signal test circuit

4 Software Design of the Test System

The system uses NI LabVIEW as a development tool, LabVIEW is a kind of graphical programming language, it can provide designers with a very simple and easy graphical design environment, and provides powerful simulation and data processing functions. Also, it is very suitable for the development of information acquisition device test system software and has the advantage to support the future extended use case. In addition, LabVIEW's powerful hardware drivers, graphical interface capabilities, and convenient as well as efficient graphical programming (G language) provide a good solution for process control and industrial applications.

4.1 Operational Interface Design

For different types and quantities of information collected by the acquisition device, it is required to configure the test cases for the acquisition channels from the aspects of the type of the collected data and the valid value scope of the data. The user can configure the acquisition device test cases according to the type and quantity of signals collected by different tanks and armored vehicles. The system has good compatibility, expandability, and versatility, and can meet the user’s need that changes of collection information are also supported for the acquisition device. Thus it can be seen that the system’s basic, and the central feature is the function of configuration.

The test cases of the analog signal acquisition channel need to configure the name of the acquisition channel, the range of the acquisition voltage and resistance (AD value), the range of the physical value and the acquisition accuracy are shown in Fig. 8.

The configuration interface for an analog signal test case consists of the following elements:

- Header:** A text box for "Analog signal test channel name" followed by an empty input field. To the right are two radio buttons: "In use" (which is selected, indicated by a black dot) and "Spare" (which is unselected, indicated by an empty circle).
- Max. value:** A label "Max. value" on the left. To its right are three input fields: "AD", "Physical value", and "Accuracy". The "Accuracy" field is a vertical stack with "+" on top and "-" on the bottom.
- Min. Value:** A label "Min. Value" on the left. To its right are three input fields: "AD", "Physical value", and "Accuracy". The "Accuracy" field is a vertical stack with "+" on top and "-" on the bottom.
- Median value:** A label "Median value" on the left. To its right are three input fields: "AD", "Physical value", and "Accuracy". The "Accuracy" field is a vertical stack with "+" on top and "-" on the bottom.
- Abnormal value:** A label "Abnormal value" on the left. To its right are two input fields: "AD" and "Physical value".

Fig. 8. Configuration interface for analog signal test case

The test cases of the switch signal acquisition channel need to configure the name of the acquisition channel, the type of the signals (high-level, low-level or high-resistance) and the physical implication of all the status as shown in Fig. 9.

IO test channel name	<input type="text"/>	In use	<input checked="" type="radio"/>	spare	<input type="radio"/>	
High level	24V	<input checked="" type="radio"/>	spare	<input type="radio"/>	Physical characteristic	<input type="text"/>
Low level	0V	<input checked="" type="radio"/>	spare	<input type="radio"/>	Physical characteristic	<input type="text"/>
High resistance	In use	<input checked="" type="radio"/>	spare	<input type="radio"/>	Physical characteristic	<input type="text"/>

Fig. 9. Switch signal acquisition test case configuration interface

For the configuration of the acquisition device with FlexRay communication, each acquisition channel needs a time slot corresponding to the FlexRay communication, and it should correspond to the byte and BIT bit of each frame of data. The specific settings are shown in Fig. 10.

Collection channel					
SLOT	<input type="text"/>	OFFSET	<input type="text"/>	Repetition	<input type="text"/>
Starting byte	<input type="text"/>	length	<input type="text"/>	mask	<input type="text"/>

Fig. 10. FlexRay communication configuration interface

For the configuration of the acquisition device with CAN communication, each acquisition channel needs to correspond to the CAN communication data frame, and it should correspond to the byte and BIT bit of the frame data. The specific settings are shown in Fig. 11.

Collection channel					
CANID	<input type="text"/>	Starting byte	<input type="text"/>	length	<input type="text"/>
		mask	<input type="text"/>		<input type="text"/>

Fig. 11. CAN communication configuration interface

4.2 Test Case Execution Flow

The test master control module controls the simulated switch signal output module, the simulated resistance signal output module, and the simulated voltage signal output

module to output signals according to the test cases, and judges the correctness of the test case execution based on the information uploaded by the communication module and fulfills the final test. The execution result of the use case is displayed by the display control module, and the execution flowchart is shown in Fig. 12.

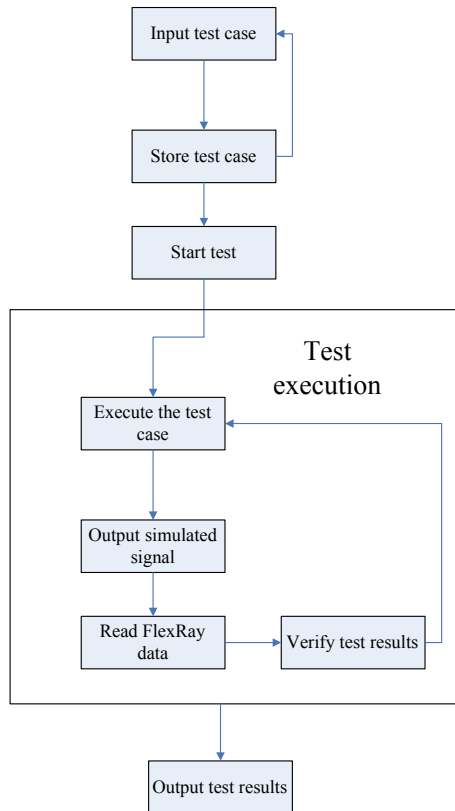


Fig. 12. Test case execution flow

Through the usage of the collection system, according to the actual collection channel of a certain type of the acquisition device, the test case is configured, and the test case is automatically executed to test the device. The results show that the technical requirements are met and the fault is accurately positioned and diagnosed. It has passed the inspection and has already successfully operated on-site. The test case shown by the display control module during the acquisition device test is shown in Fig. 13. The on-site commissioning and stable operation effects show that the proposed test system is reasonably designed with good control effect, short test time and high versatility.

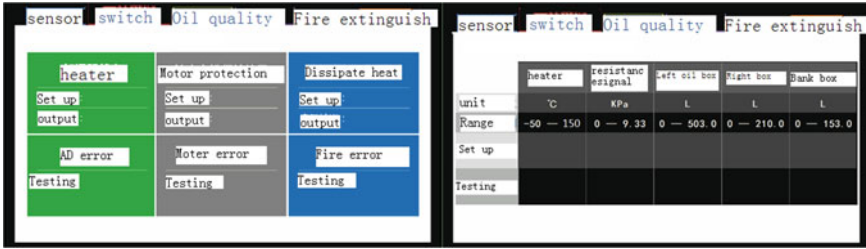


Fig. 13. A type of acquisition device test case output display interface

5 Conclusion

Through the self-configuration of test cases, the proposed configuration-based automatic test system can automatically test multiple acquisition devices through one single system. The system has the features of good scalability, high test speed, stable system performance, friendly man-machine interface, and is easy to operate, as a result, the test efficiency has improved significantly.

References

1. Joint Technical Architecture. Department of Defense, 200104
2. DoD Executive Agent for Automatic Test System Update [EB OL]. <http://140.229.102.110>, 200109
3. DoD Automatic Test System Architecture Guide (1999)



Efficient Sensitivity Analysis of Dynamic Neuro-space Mapping for Transistor Modeling

Lin Zhu¹(✉), Jian Zhao¹, and Wenyan Liu²

¹ Tianjin Chengjian University, Tianjin, China
zeal286@163.com

² Shaanxi University of Science and Technology, Xi'an, China

Abstract. In this paper, an enhanced dynamic Neuro-space mapping (Neuro-SM) method is proposed with emphasis on transistor modeling. By modifying the dynamic voltage relationships in an existing nonlinear model, the proposed Neuro-SM produces a new and more accurate model than the nonlinear model as well as the static Neuro-SM. Compared to the existing dynamic Neuro-SM, a new sensitivity analysis technique is derived to speed up the training of the proposed model with dc, small- and large-signal data. The validity and efficiency of the proposed Neuro-SM method are demonstrated by modeling examples of a GaAs high-electron-mobility transistor (HEMT). Suitable value of time delay parameter which is equal to one divided by 3 or 5 times of the largest frequency considered in simulation is suggested and demonstrated by the modeling example.

Keywords: Neural networks · Neuro-SM · Transistor modeling · Optimization · Simulation

1 Introduction

Recently, a static Neuro-space mapping (Neuro-SM) [1] technique, combining the concept of space mapping and neural networks was introduced which aims to map an existing approximate device model toward an accurate model for transistor modeling. It is proved that the static Neuro-SM model provides better accuracy over the existing equivalent circuit model. The work in [2] considers not only voltage mappings, but also current mappings. Neuro-SM models developed by this method are more accurate than the existing equivalent circuit models and the static Neuro-SM models. However, the gap between the existing model and desired model are not all caused by differences of non-memory devices (such as resistors), partly caused by memory devices (such as capacitors and inductors). In this way, such simple static mappings for voltage and current signals may be still insufficient. In order to further improve the accuracy of transistor modeling, a dynamic Neuro-SM approach was presented by combining dynamic neural network with the existing model [3]. Due to use of dynamic mapping neural network, it can make up any dynamic effects such as capacitor effect and non-quasi-static effect [4] that may be missing in the existing models so that it can produce a more accurate model than static Neuro-SM presented in [1] and [2]. However, the

sensitivity analysis of the dynamic Neuro-SM model under DC, small- and large-signal conditions is not provided in [3].

This paper is a significant expansion and further advance of the work in [3]. The sensitivities of direct current (DC), small-signal S-parameter, and large-signal harmonic balance (HB) responses of the proposed Neuro-SM model with respect to mapping parameters, which are not presented in [3], are derived. This can help the training of the proposed model to be more efficient due to avoiding brute-force perturbation [5] in each training iteration. Suitable value of time delay parameter, which is also not presented in [3], is suggested and demonstrated by application examples of a GaAs high-electron-mobility transistor (HEMT) modeling. The modeling results show that the proposed Neuro-SM approach allows us to produce improved accuracy over existing equivalent circuit model and static Neuro-SM model [1], and to perform faster training over dynamic Neuro-SM presented in [3].

2 Structure of the Dynamic Neuro-SM Model

Suppose the best possible existing model gives only rough approximation of the device behavior, and cannot accurately match the device data. We call this existing model as coarse model. Let the total number of delay buffers of voltages at gate and drain be the same and both equal to N_d . Let the vector v_f and v_c contain the gate and drain voltages of the fine model and coarse model, respectively. Structure of the proposed dynamic Neuro-SM model is shown in Fig. 1. As seen in this figure, the present voltage of the fine model defined as $v_f(t)$, together with its history defined as $v_f(t - \tau)$, $v_f(t - 2\tau)$, ..., and $v_f(t - N_d\tau)$ are mapped into present voltages of the coarse model defined as $v_c(t)$.

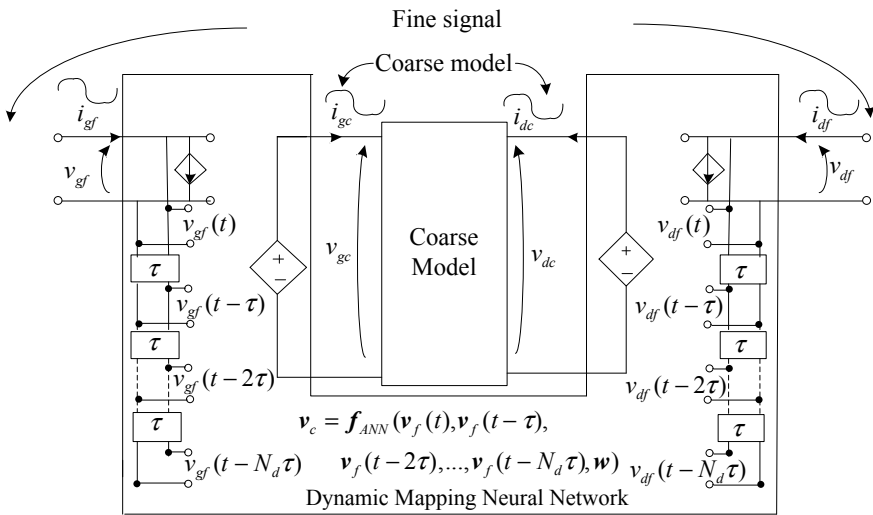


Fig. 1. Structure of the proposed two-port dynamic Neuro-SM nonlinear device mode

In functional form, the voltage signals of the coarse model can be generally described as

$$\mathbf{v}_c(t) = \mathbf{f}_{ANN}(\mathbf{v}_f(t), \mathbf{v}_f(t - \tau), \dots, \mathbf{v}_f(t - N_d\tau), \mathbf{w}) \quad (1)$$

where \mathbf{f}_{ANN} is a neural network to learn and represent the mapping. We use vector \mathbf{w} in (1) to represent the internal synaptic weights of the neural network. Let the vector \mathbf{i}_f and \mathbf{i}_c contain the gate and drain currents of the fine model and coarse model, respectively. The \mathbf{v}_c from (1) is then fed to the coarse model producing \mathbf{i}_c . Finally, \mathbf{i}_c is passed onto external current of the fine model, i.e., \mathbf{i}_f .

3 Proposed Formulation of the Dynamic Neuro-SM Model for Training

The dynamic Neuro-SM model will not be accurate unless the dynamic mapping neural network is trained from transistor data. Dynamic mapping neural network defined in (1) is in the form of instantaneous voltage while the modeling data obtained from simulation or measurement is in DC, bias-dependent S-parameter, and large-signal HB form. In order to formulate an advanced training method for dynamic mapping neural network, the connections of the proposed dynamic mapping neural network with typical types of transistor data, such as DC, S-parameter and HB data, need to be derived.

In DC case, the present voltage signals of the fine model and their delayed signals are all equal and are set to be $\mathbf{V}_{f,DC}$. The dynamic analytical mapping characteristic for DC output current of the fine model can be generally described as

$$\mathbf{I}_f = \mathbf{I}_f(\mathbf{V}_{f,DC}) = \mathbf{I}_c \left| \begin{array}{c} \mathbf{V}_{c,DC} = \mathbf{f}_{ANN}(\overbrace{\mathbf{V}_{f,DC}, \mathbf{V}_{f,DC}, \dots, \mathbf{V}_{f,DC}}^{N_d+1}, \mathbf{w}) \end{array} \right. \quad (2)$$

The small-signal S-parameter can be derived through the dynamic analytical mapping of the \mathbf{Y} matrices between the coarse model \mathbf{Y}_c and the fine model \mathbf{Y}_f , shown as

$$\mathbf{Y}_f(\omega) = \mathbf{Y}_c(\omega) \left| \begin{array}{c} \mathbf{V}_{c,Bias} = \mathbf{f}_{ANN}(\overbrace{\mathbf{V}_{f,Bias}, \mathbf{V}_{f,Bias}, \dots, \mathbf{V}_{f,Bias}}^{N_d+1}, \mathbf{w}) \\ \left(\sum_{k=0}^{N_d} e^{-j\omega k\tau} \cdot \frac{\partial \mathbf{f}_{ANN}^T(\mathbf{v}_f(t), \mathbf{v}_f(t - \tau), \dots, \mathbf{v}_f(t - N_d\tau), \mathbf{w})}{\partial \mathbf{v}_f(t - k\tau)} \right) \Big|_{\mathbf{v}_f(t) = \mathbf{v}_f(t - \tau) = \dots = \mathbf{v}_f(t - N_d\tau) = \mathbf{V}_{f,Bias}} \end{array} \right)^T \quad (3)$$

where the derivative of \mathbf{f}_{ANN} can be obtained using the adjoint neural network method [5], and k represents the index of delay buffers.

For large-signal simulation, the relationship between HB computation and dynamic neural network weights \mathbf{w} needs to be established so that model training can be

performed with HB data. Let the harmonic currents of the dynamic Neuro-SM model and coarse model at a generic harmonic frequency ω_k be $\mathbf{I}_f(\omega_k)$ and $\mathbf{I}_c(\omega_k)$, respectively. The $\mathbf{I}_f(\omega_k)$ can be evaluated as

$$\begin{aligned} \mathbf{I}_f(\omega_k) &= \frac{1}{N_T} \sum_{n=0}^{N_T-1} \left(\mathbf{i}_c(t_n) \Big|_{\mathbf{v}_c(t_n)=f_{ANN} \left(\sum_{k=0}^{N_H} V_f(\omega_k) \cdot W_N^*(n,k), \sum_{k=0}^{N_H} V_f(\omega_k) \cdot W_N^*(n,k) \cdot e^{-j\omega_k t}, \dots, \sum_{k=0}^{N_H} V_f(\omega_k) \cdot W_N^*(n,k) \cdot e^{-j\omega_k N_d t}, \mathbf{w} \right)} \right. \\ &\quad \left. + j\omega_k \mathbf{q}_c(t_n) \Big|_{\mathbf{v}_c(t_n)=f_{ANN} \left(\sum_{k=0}^{N_H} V_f(\omega_k) \cdot W_N^*(n,k), \sum_{k=0}^{N_H} V_f(\omega_k) \cdot W_N^*(n,k) \cdot e^{-j\omega_k t}, \dots, \sum_{k=0}^{N_H} V_f(\omega_k) \cdot W_N^*(n,k) \cdot e^{-j\omega_k N_d t}, \mathbf{w} \right)} \right) \\ &\quad \cdot W_N(n, k) \end{aligned} \quad (4)$$

where the subscript k represents the index of the harmonic frequency, $k = 0, 1, 2, \dots, N_H$, where N_H is the number of harmonics. N_T is the number of time samples, $W_N(n, k)$ is the Fourier coefficient for the n th time sample and the k th harmonic, and superscript* denotes complex conjugate.

4 Proposed Sensitivity Analysis of the Enhanced Neuro-SM Model with Respect to Mapping Neural Network Weights

Let w_i be a generic symbol representing an internal weight of the mapping neural network. The sensitivity of the proposed dynamic Neuro-SM model with respect to w_i gives gradient information needed for efficient training of the proposed dynamic Neuro-SM model. In order to efficiently train the proposed model, gradient information provided by the sensitivity of the model with respect to w_i is needed [6]. If the gradient information is not provided in training process, bruteforce perturbation [5] has to be performed in each iteration of training, which results in slow training speed. In order to improve the training speed and efficiency, the sensitivity formulas of the proposed dynamic Neuro-SM are derived in this paper.

- (1) DC Sensitivity: let the total number of delay buffers of voltages at gate and drain be the same and both equal to N_d . In the dc case, the sensitivity of the output current of the dynamic Neuro-SM model, i.e., \mathbf{I}_f with respect to w_i is

$$\frac{\partial \mathbf{I}_f}{\partial w_i} = \left(\frac{\partial \mathbf{I}_c^T}{\partial \mathbf{V}_{c,DC}} \right)^T \cdot \left(\frac{\partial \mathbf{V}_{c,DC}}{\partial w_i} \right) = \mathbf{G}_c \cdot \left(\frac{\partial f_{ANN}(\overbrace{\mathbf{V}_{f,DC}, \mathbf{V}_{f,DC}, \dots, \mathbf{V}_{f,DC}}^{N_d+1}, \mathbf{w})}{\partial w_i} \right) \quad (5)$$

where \mathbf{G}_c is the dc conductance matrix of the coarse model, and $\partial f_{ANN} / \partial w_i$ is the first-order derivative computed by neural network back-propagation [7].

- (2) **Small-signal Sensitivity:** The sensitivity for Y-parameters with respect to dynamic mapping neural network weight w_i can be derived in Eq. (6). This equation includes two parts. The first part has the second-order derivative of the neural network, which is the differentiation of the Jacobian matrix $\partial \mathbf{f}_{ANN}^T(\mathbf{v}_f(t), \mathbf{v}_f(t - \tau), \dots, \mathbf{v}_f(t - N_d\tau), \mathbf{w}) / \partial \mathbf{v}_f(t - k\tau)$ with respect to w_i , where, k represents the index of delay buffers. This second-order derivative can be achieved by the adjoint neural network sensitivity analysis. The second part is the sensitivity of the coarse model Y-parameter, which is dependent on the mapped DC-bias voltages, and thus the neural network weights. Here, $v_{cj}(t) = f_{ANNj}(\mathbf{v}_f(t), \mathbf{v}_f(t - \tau), \dots, \mathbf{v}_f(t - N_d\tau), \mathbf{w}), j = 1, 2$ presents the voltage signal at the coarse input port 1 (if $j = 1$) or port 2 (if $j = 2$). By converting Y-parameters to S-parameters, sensitivity for S-parameter can be subsequently obtained.

$$\begin{aligned}
 & \frac{\partial \mathbf{Y}_f(\omega)}{\partial w_i} \\
 & = \mathbf{Y}_c(\omega) \bigg|_{\mathbf{v}_c, \text{Bias} = f_{ANN}(\overbrace{\mathbf{V}_{f, \text{Bias}}, \mathbf{V}_{f, \text{Bias}}, \dots, \mathbf{V}_{f, \text{Bias}}}^{N_d+1}, \mathbf{w})} \\
 & \cdot \left(\sum_{k=0}^{N_d} e^{-j\omega k\tau} \cdot \frac{\partial^2 \mathbf{f}_{ANN}^T(\mathbf{v}_f(t), \mathbf{v}_f(t - \tau), \dots, \mathbf{v}_f(t - N_d\tau), \mathbf{w})}{\partial \mathbf{v}_f(t - k\tau) \partial w_i} \bigg|_{\mathbf{v}_f(t) = \mathbf{v}_f(t - \tau) = \dots = \mathbf{v}_f(t - N_d\tau) = \mathbf{V}_{f, \text{Bias}}} \right)^T \\
 & + \left(\sum_{j=1,2} \frac{\partial \mathbf{Y}_c(\omega)}{\partial v_{cj}} \bigg|_{\mathbf{v}_c = \mathbf{V}_{c, \text{Bias}}} \cdot \frac{\partial f_{ANNj}(\mathbf{v}_f(t), \mathbf{v}_f(t - \tau), \dots, \mathbf{v}_f(t - N_d\tau), \mathbf{w})}{\partial w_i} \bigg|_{\mathbf{v}_f(t) = \mathbf{v}_f(t - \tau) = \dots = \mathbf{v}_f(t - N_d\tau) = \mathbf{V}_{f, \text{Bias}}} \right) \\
 & \cdot \left(\sum_{k=0}^{N_d} e^{-j\omega k\tau} \cdot \frac{\partial \mathbf{f}_{ANN}^T(\mathbf{v}_f(t), \mathbf{v}_f(t - \tau), \dots, \mathbf{v}_f(t - N_d\tau), \mathbf{w})}{\partial \mathbf{v}_f(t - k\tau)} \bigg|_{\mathbf{v}_f(t) = \mathbf{v}_f(t - \tau) = \dots = \mathbf{v}_f(t - N_d\tau) = \mathbf{V}_{f, \text{Bias}}} \right)^T
 \end{aligned} \tag{6}$$

- (3) **Large-signal Sensitivity:** The sensitivity of the output current of the proposed dynamic Neuro-SM model at a generic harmonic frequency $\omega_k, k = 0, 1, 2, \dots, N_H$ are shown as

$$\begin{aligned}
 \frac{\partial \mathbf{I}_f(\omega_k)}{\partial w_i} & = \frac{1}{N_T} \sum_{n=0}^{N_T-1} \left(\mathbf{G}_c(t_n) \bigg|_{\mathbf{v}_c(t_n) = f_{ANN} \left(\sum_{k=0}^{N_H} \mathbf{V}_f(\omega_k) \cdot W_N^n(n, k), \sum_{k=0}^{N_H} \mathbf{V}_f(\omega_k) \cdot W_N^n(n, k) \cdot e^{-j\omega_k \tau}, \dots, \sum_{k=0}^{N_H} \mathbf{V}_f(\omega_k) \cdot W_N^n(n, k) \cdot e^{-j\omega_k N_d \tau}, \mathbf{w} \right)} \right. \\
 & \left. + j\omega_k \cdot \mathbf{C}_c(t_n) \bigg|_{\mathbf{v}_c(t_n) = f_{ANN} \left(\sum_{k=0}^{N_H} \mathbf{V}_f(\omega_k) \cdot W_N^n(n, k), \sum_{k=0}^{N_H} \mathbf{V}_f(\omega_k) \cdot W_N^n(n, k) \cdot e^{-j\omega_k \tau}, \dots, \sum_{k=0}^{N_H} \mathbf{V}_f(\omega_k) \cdot W_N^n(n, k) \cdot e^{-j\omega_k N_d \tau}, \mathbf{w} \right)} \right) \\
 & \cdot \frac{\partial \mathbf{f}_{ANN}(\mathbf{v}_f(t_n), \mathbf{v}_f(t_n - \tau), \dots, \mathbf{v}_f(t_n - N_d\tau), \mathbf{w})}{\partial w_i} \cdot W_N(n, k)
 \end{aligned} \tag{7}$$

where, $\mathbf{G}_c(t_n) = (\partial \mathbf{i}_c^T / \partial \mathbf{v}_c)^T$ and $\mathbf{C}_c(t_n) = (\partial \mathbf{q}_c^T / \partial \mathbf{v}_c)^T$ are the nonlinear conductance and capacitance matrices of the coarse model evaluated at the time point t_n at the mapped signal $\mathbf{v}_c(t_n)$.

5 Dynamic Neuro-SM Modeling Example of a GaAs HEMT Trained with Physics-Based Device Data

This example illustrates the dynamic Neuro-SM and the necessity of using sensitivity analysis technique for model training. Training and test data were generated from a physics-based device simulator, Medici [8]. Keysight-ADS internal Angelov model is used as the coarse model. Training was done using DC and bias-dependent S-parameter data at 138 different biases and refined by large-signal harmonic data generated at different fundamental frequencies (2, 3, 4, and 5 GHz) and input power levels (-20 ~ 10 dBm). Time delay parameter of the dynamic mapping neural networks is 0.005 ns.

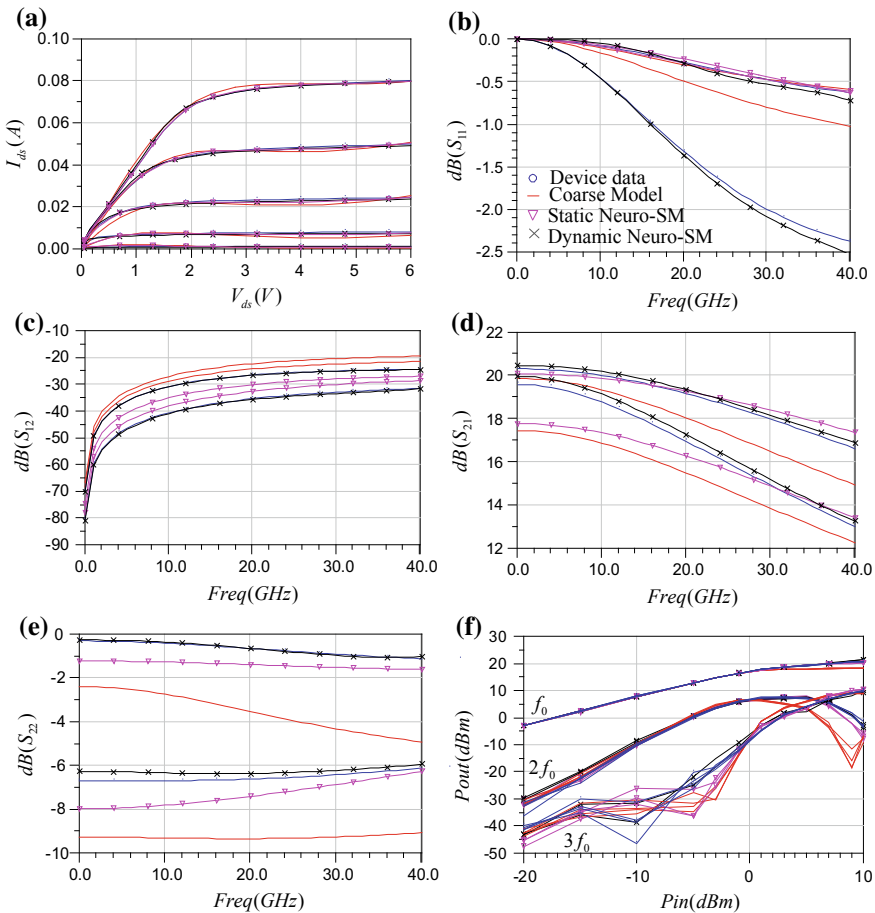


Fig. 2. Comparison between the GaAs HEMT device data, coarse model, static Neuro-SM with 30 hidden neurons, and the proposed dynamic Neuro-SM with 5 delay buffers and 30 hidden neurons, **a** dc. **b ~ e** S-parameter at 2 test bias points (0.7, 2.4 V), and (0.3, 5.2 V). **f** HB at bias points (0.2, 5 V)

The coarse model alone, even after optimizing the parameters in the coarse model, cannot match Medici data, as shown in Fig. 2. In order to improve the model, we first use static Neuro-SM to map the coarse model to device data. Note that although the static Neuro-SM with static mappings is clearly more accurate than the coarse model alone, it's still insufficient to achieve the desired accuracy, as shown in Fig. 2. Hence the proposed dynamic method is used to get a more accurate model. We trained a dynamic Neuro-SM model with 1, 3, and 5 delay buffers. After training, we compare the test error of the proposed Neuro-SM model with the coarse model and the static Neuro-SM model shown in Table 1. It is clearly seen from the table that the proposed dynamic Neuro-SM model with 5 delay buffers is more accurate than the static Neuro-SM and coarse models. To show the results further, we plot the DC, S-parameter and HB responses shown in Fig. 2. As seen in Fig. 2, the proposed dynamic Neuro-SM model can match the device data more closely than the static Neuro-SM and coarse models.

Table 1. Training and test error comparison of coarse model, existing static Neuron-SM model and the proposed dynamic model

Model type	Training error/test error for f_{ANN} with 30 hidden neurons	Training error/test error for f_{ANN} with 40 hidden neurons
Coarse model (%)	10.918/10.664	10.918/10.664
Static Neuro-SM (%)	5.133/4.948	5.035/4.879
Dynamic Neuro-SM ($N_d = 1$) (%)	3.861/3.801	3.611/3.618
Dynamic Neuro-SM ($N_d = 3$) (%)	2.223/2.011	1.995/1.880
Dynamic Neuro-SM ($N_d = 5$) (%)	1.613/1.437	1.576/1.503

Another factor except for number of hidden neurons and delay buffers that impacts the training/testing error is time delay parameter. In order to find out suitable time delay parameters, the training of the proposed Neuro-SM model with the same number of hidden neurons and delay buffers as well as different time delay parameters was done using DC and bias-dependent S-parameter data at 138 different biases. After training, the results are shown in Table 2. As seen in Table 2, too small or large time delay parameter τ may lead to large training/testing error. Hence the choice of suitable τ value is very important to the accuracy of the enhanced Neuro-SM model. In general, $1/(5f_{\max})$ or $1/(3f_{\max})$ are found suitable, where f_{\max} is the maximum frequency considered in transistor modeling. In extreme case of $\tau = 0$, the training/testing error of the model which actual belong to Neuro-SM model with static mapping is still very large. This also explains the necessary of using dynamic mapping in Neuro-SM technique for transistor modeling.

Table 2. Training and testing error comparison of the dynamic Neuro-SM model with five delay buffers and 30 hidden neurons at different time delay parameters

Time delay parameter (ns)	=0	$1/(50f_{\max})$ = 0.0005	$1/(18f_{\max})$ = 0.0014	$1/(5f_{\max})$ = 0.005	$1/(3f_{\max})$ = 0.0083	$1/(2f_{\max})$ = 0.0125	$1/(0.5f_{\max})$ = 0.05
Training error (%)	5.21	2.63	1.93	1.57	1.40	1.58	4.44
Test error (%)	5.12	2.39	1.95	1.41	1.38	1.49	4.10

In this example, we also compared the training time between dynamic Neuro-SM with perturbation in [3] and the proposed dynamic Neuro-SM with sensitivity analysis for 100 training iterations on an Intel i5-4460 3.2-GHz computer as shown in Table 3. Training was also done using DC and S-parameter data at 138 different biases simultaneous. The result shows that the training speed of the dynamic Neuro-SM with sensitivity analysis is much faster than that of the dynamic Neuro-SM with broteforce perturbation presented in [3].

Table 3. Training time comparison between dynamic Neuro-SM with perturbation and the proposed Neuro-SM with sensitivity analysis for 100 training iterations for the GaAs HEMT example

Training time for dynamic Neuro-SM with perturbation in [7]	Training time for proposed dynamic Neuro-SM with sensitivity analysis
1.76 h	218.0 s

6 Conclusion

An enhanced dynamic Neuro-SM technique has been proposed for accurate transistor modeling in this paper. Due to use of dynamic mapping, the proposed technique can make up for any missing dynamic effects in addition to resistive effects in the coarse model so that the proposed dynamic Neuro-SM model can be more accurate than the existing coarse model and the static Neuro-SM model. We have also derived new sensitivity formulations for dc, small-, and large-signal training. The proposed dynamic Neuro-SM has demonstrated much-improved training efficiency over existing dynamic Neuro-SM by a transistor modeling example.

Acknowledgements. This work is supported by Scientific Research Plan Project by Tianjin Education Commission (No. 2016CJ13).

References

1. Zhang, L., Xu, J., Yagoub, M., et al.: Neuro-space mapping technique for nonlinear device modeling and large-signal simulation. *IEEE MIT-S Int. Microw. Symp. Philadelphia, PA*, Jun. 2003, pp. 173–176
2. Zhu, L., Liu, K., Zhang, Q., et al.: An enhanced analytical neuro-space mapping method for large-signal microwave device modeling. *IEEE MIT-S Int. Microw. Symp. Dig. Montreal, QC*, Jun. 2012, pp. 1–3
3. Zhu, L., Zhang, Q., Liu, K., et al.: A novel dynamic neuro-space mapping approach for nonlinear microwave device modeling. *IEEE Microw. Wirel. Compon. Lett.* **26**(2), 131–133 (2016)
4. Long, Y., Guo, Y., Zhong, Z.: A 3-D table-based method for non-quasi-static microwave FET devices modeling. *IEEE Trans. Microw. Theory Tech.* **60**(10), 3088–3095 (2012)
5. Song, Q., Spall, J., Soh, Y., et al.: Robust neural network tracking controller using simultaneous perturbation stochastic approximation. *IEEE Trans. Neural Netw.* **19**(5), 817–835 (2008)
6. Zhang, L., Xu, J., Yagoub, M.C., et al.: Efficient analytical formulation and sensitivity analysis of neuro-space mapping for nonlinear microwave device modeling. *IEEE Trans. Microw. Theory Tech.* **53**(9), 2752–2767 (2005)
7. Zhang, Q., Gupta, K., Devabhaktuni, V.: Artificial neural networks for RF and microwave design: From theory to practice. *IEEE Trans. Microw. Theory Tech.* **51**(4), 1339–1350 (2003)
8. Medici 2013 I-2013.12-0. Synopsys Inc., Mountain View, CA, 2013



Open-Loop Carrier Synchronization Design and Its FPGA Implementation for Short Burst Communication at Low SNR

Wen Che^(✉) and Jinhui Fang

School of Information and Electronics, Beijing Institute of Technology, Beijing
10081, China
1002093408@qq.com

Abstract. The paper presents a design for open-loop carrier synchronization which is suitable for short burst communication at low SNR and it is easy to implement on FPGA platform. The proposed open-loop carrier synchronization is based on V&V and FFT algorithm which estimate both the frequency offset and the phase offset. We also implement a post-processing to solve the problem of phase ambiguity. The simulation results show that our proposed algorithm can work efficiently. Besides, the algorithm is implemented with a Xilinx XC7VX690T FPGA chip which achieves good performance under the condition of large frequency offset and low signal-to-noise ratio.

Keywords: Carrier synchronization · V&V algorithm · FFT · FPGA · Burst communication

1 Introduction

In view of the high dynamic characteristics of burst communications, Doppler shifts are inevitable in the received signals, which will seriously affect the performance of the communication system. By adopting carrier synchronization technology, both carrier frequency offset and phase offset can be corrected. Traditional closed-loop PLLs have a long acquisition time and are not suitable for short burst communication [1]. Open-loop structure avoids the design of feedback loop and can acquire the offset information immediately, thus is very suitable for burst communications.

Due to the high dynamic environment, the carrier synchronization needs to be re-performed in every burst frame. A preamble of unmodulated carrier can indicate the beginning of each frame and also give a rough estimate of carrier frequency offset. However, the capture accuracy is limited by the length of preamble. This paper will show that with just a short length of preamble, the frequency and phase can be estimated accurately without phase ambiguity by using a FFT-based open-loop estimation algorithm.

2 Signal Model

In most burst communication systems, a preamble of unmodulated carrier is provided in each slot which lasts for T_p . T is the symbol duration, T_s the slot duration, T_f is the frame duration. The received signal can be described as

$$x(t) = \begin{cases} ae^{j(2\pi f_c t + \theta_i)} + n(t) & t < T_p \\ ae^{j(2\pi f_c t + \theta_M(t) + \theta_i)} + n(t) & T_p < t < T_s \end{cases} \quad (1)$$

where a is the amplitude of the received signal, f_c is the received carrier frequency, θ_i is the carrier initial phase, $\theta_M(t) \in \{\frac{2\pi m}{4}, m = 0, 1, 2, 3\}$ is the information contained in the signal, $n(t)$ is the additive random noise. And if we absorb the noise term in the signal and perform a rectangular-to-polar transformation, we get (2) where $a'(t)$ is the practical signal amplitude, $\theta_n(t)$ is the equivalent noise phase.

$$x(t) = \begin{cases} a'(t)e^{j(2\pi f_c t + \theta_i + \theta_n(t))} & t < T_p \\ a'(t)e^{j(2\pi f_c t + \theta_M(t) + \theta_i + \theta_n(t))} & T_p < t < T_s \end{cases} \quad (2)$$

Then we implement the quadrature down conversion to the received signal,

$$\begin{aligned} y(t) &= x(t) * e^{-j(2\pi f_o t + \theta_o)} \\ &= \begin{cases} a'(t)e^{j(2\pi(f_c - f_o)t + (\theta_i - \theta_o) + \theta_n(t))} & t < T_p \\ a'(t)e^{j(2\pi(f_c - f_o)t + \theta_M(t) + (\theta_i - \theta_o) + \theta_n(t))} & T_p < t < T_s \end{cases} \quad (3) \\ &= \begin{cases} a'(t)e^{j(2\pi\Delta f t + \Delta\theta(t))} & t < T_p \\ a'(t)e^{j(2\pi\Delta f t + \theta_M(t) + \Delta\theta(t))} & T_p < t < T_s \end{cases} \end{aligned}$$

where f_o, θ_o , are the local carrier frequency and phase, respectively. And $\Delta f = f_c - f_o$ is frequency offset, $\Delta\theta(t) = \theta_i - \theta_o + \theta_n(t)$ is the phase offset.

To simplify the problem, assume that the sample rate is equal to the symbol rate, and the symbol timing is known, so the output of matched filter is sampled at symbol time $t = kT$

$$y(k) = y(kT) = \begin{cases} a'(k)e^{j(\frac{2\pi\Delta f k}{T} + \Delta\theta(k))} & k < \frac{T_p}{T} \\ a'(k)e^{j(\frac{2\pi\Delta f k}{T} + \theta_M(k) + \Delta\theta(k))} & \frac{T_p}{T} < k < \frac{T_s}{T} \end{cases} \quad (4)$$

It is noted that the phase error is made up of two parts, one is the cumulative phase due to the presence of Δf , another is the phase offset. In order to demodulate the received signal correctly, the phase error must be calculate accurately, that is to say, frequency offset and phase offset [2].

3 Algorithm and Implement Description

3.1 Algorithm Description

With the presence of preamble the rough frequency and phase offset can be easily obtain by FFT algorithm. We calculate N_0 -point FFT ($N_0 < \frac{T_p}{2T}$), and search peak point of the FFT amplitude spectrum.

$$\Delta\hat{f}_0 = \begin{cases} \frac{N_{\max}f}{N_0} & N_{\max} < \frac{N_0}{2} \\ \frac{N_{\max}f}{N_0} - f & N_{\max} > \frac{N_0}{2} \end{cases} \quad (5)$$

the rough phase estimate lies in the same location in the FFT phase spectrum.

$$\Delta\hat{\theta}_0 = \varphi(N_{\max}) \quad (6)$$

However due to the limited length of preamble, the estimation can't be that accurate, the capture range is between $-f/2$ and $f/2$, the capture precision is only f/N_0 .

So the compensated phase can be described as below:

$$z(k) = a'(k)e^{j(\frac{2\pi(\Delta f - \Delta\hat{f}_0)k}{f_s} + \theta_M(k) + \Delta\theta(k) - \Delta\hat{\theta}_0)} \quad (7)$$

where $\Delta f - \Delta\hat{f}_0 < \frac{f}{2}$, $\Delta\theta(k) - \Delta\hat{\theta}_0$ is relevant to SNR.

To cope with the high dynamic situation and improve the capture precision we should continue to estimate the phase of the modulated signal.

The V&V algorithm

This algorithm was first proposed by Viterbi and Viterbi [3]. According to this algorithm, first we perform a rectangular-to-polar transformation to the output of the timing recovery

$$z_n = x_n + jy_n = \rho_n e^{j\theta_n} \quad (8)$$

Then multiply phase θ_n by m for m -PSK modulation and perform an arbitrary nonlinear transformation on ρ_n and finally perform a polar-to-rectangular transformation on the result.

$$z'_n = F(\rho_n) e^{jm\theta_n} \quad (9)$$

By calculating the average of the preceding and succeeding $2N + 1$ transformed symbols, we get the estimation of the carrier phase [4] (Fig. 1).

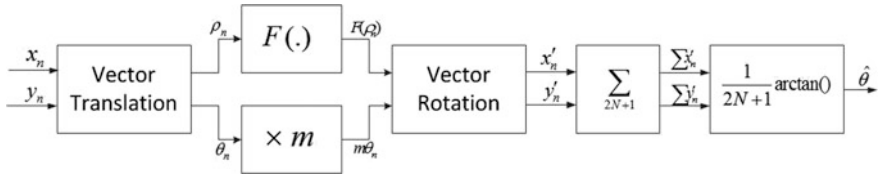


Fig. 1. Block diagram of V&V algorithm

$$\hat{\theta} = \frac{1}{m} \arg \left\{ \frac{1}{2N+1} \sum_{n=-N}^N F(\rho_n) e^{jm\theta_n} \right\} \tag{10}$$

However, this algorithm can't cope with the circumstance where frequency offset can't be negligible. So frequency correction prior phase estimation is then necessary.

Carrier frequency offset estimation FFT algorithm.

As in the V&V algorithm, the term $F(\rho_n)e^{j4\theta_n}$ performs QPSK modulation removal. We use N-point FFT to estimate the frequency offset.

$$Z'(k) = |Z'(k)|e^{j\varphi(k)} = \sum_{n=0}^{N-1} z'(k)e^{-j\frac{2\pi kn}{N}}, k = 0, 1, \dots, N-1 \tag{11}$$

According to the maximum likelihood (ML) estimate, the estimation of the frequency offset lies in the peak point of the FFT amplitude spectrum and accompanying with this, the initial phase estimation is also obtained.

$$\Delta\hat{f} = \begin{cases} \frac{k_{\max}f}{4N} & k_{\max} < \frac{N}{2} \\ \frac{k_{\max}f}{4N} - \frac{f}{4} & k_{\max} > \frac{N}{2} \end{cases} \tag{12}$$

$$\Delta\hat{\theta} = \frac{\varphi(k_{\max})}{4} \tag{13}$$

This time the track range is between $-f/8$ and $f/8$, the track precision is only f/N .

In the V&V algorithm, multiplying the phase by m, the final operation of phase estimation delivered a m-fold ambiguity. We then need to implement a post-processing after that.

Assume the carrier phase estimation at the (k-1)th FFT is $\Delta\hat{\theta}(k-1)$, we may predict the carrier phase estimation at the kth FFT is

$$\Delta\hat{\theta}_{ref}(k) = \Delta\hat{\theta}(k-1) + 2\pi\Delta\hat{f}(k-1)N/f \tag{14}$$

We use $\Delta\hat{\theta}_{ref}(k)$ to modify the $\Delta\hat{\theta}(k)$. Note that $\Delta\hat{\theta}_d(k) = \Delta\hat{\theta}(k) - \Delta\hat{\theta}_{ref}(k)$ lies around the constellation axis $(0, \pi/2, \pi, -\pi/2)$, let $\hat{\theta}_m(k)$ represented the modified phase estimation.

$$\Delta\hat{\theta}_m(k) = \begin{cases} \Delta\hat{\theta}(k) & \left| \Delta\hat{\theta}_d(k) \right| < \frac{\pi}{4} \\ \Delta\hat{\theta}(k) - \frac{\pi}{2} & \frac{\pi}{4} < \Delta\hat{\theta}_d(k) < \frac{3\pi}{4} \\ \Delta\hat{\theta}(k) + \pi & \left| \Delta\hat{\theta}_d(k) \right| > \frac{3\pi}{4} \\ \Delta\hat{\theta}(k) + \frac{\pi}{2} & -\frac{3\pi}{4} < \Delta\hat{\theta}_d(k) < -\frac{\pi}{4} \end{cases} \quad (15)$$

3.2 Implement Description

The main implemental architecture of our proposed algorithm on FPGA platform is shown in Fig. 2. The FFT calculation introduces a long delay, so the signal needs to be buffered, and then output and compensated after the frequency offset and phase offset are estimated. The specific delay is fixed and can be obtained through simulation. After removing the modulation from the initially compensated signal by V&V algorithm, FFT is performed again. The same as before, due to the delay introduced by the FFT calculation, the signal needs to be buffered, in order to prevent the data from being overwritten, a ping-pong operation is used to buffer data alternately with two RAMs.

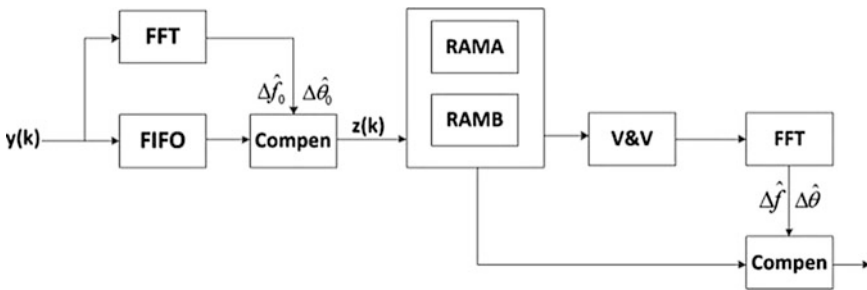


Fig. 2. Main implemental architecture

4 Simulation and Implement Results

The performance of the proposed parallel carrier synchronization algorithm is simulated in MATLAB. The Verilog code version of this algorithm is also simulated in Modelsim with the same source as in MATLAB. The simulation parameters are as follows: QPSK modulation; $T = 1 \mu s$, $T_p = 0.192 \text{ ms}$, $T_s = 0.1 \text{ s}$; Choose $N_0 = 64$, $N = 1024$. $F(\rho) = \rho$ To make the estimate more accurate, insert zeros to the end of the data until the number of FFT points reaching 4096. So the frequency estimate error is less than $\frac{f}{4 \times 4096}$. The results which are shown in Figs. 3, 4, 5, 6 are under the

condition of $E_b/N_0 = 4$. When frequency offset is fixed ($\Delta f = 0.25f$), by comparing the two signal constellations in Fig. 3, we can find frequency offset and phase offset are well compensated. Figure 4 shows the phase offset tracking before and after modification. Figure 5 shows the frequency offset tracking under the condition of frequency offset begins at $0.35f$ and changes with time by 50 kHz/s. Figure 6 compares the BER simulation results of the proposed algorithm with the theoretical value.

We have built a system concept demonstrator using Xilinx XC7VX690T FPGA and A/D converter EV10AQ190. The utilized resources are shown in Table 1.

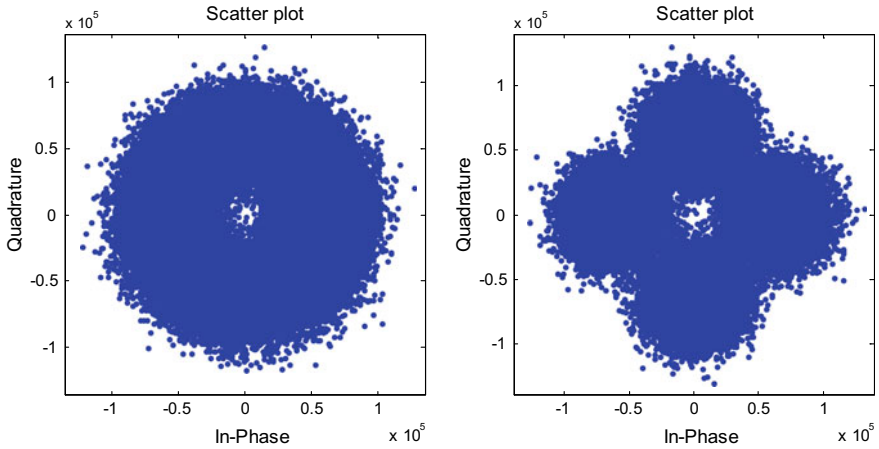


Fig. 3. Signal constellation before and after carrier synchronization

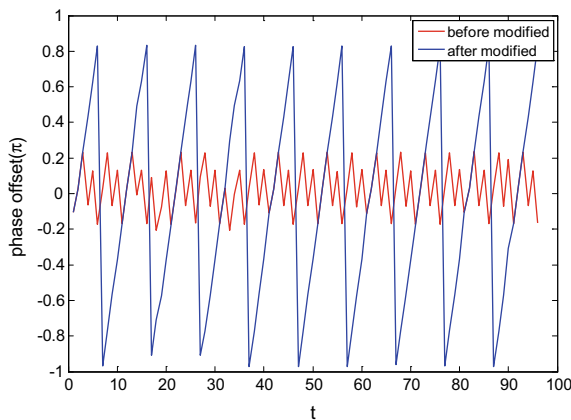


Fig. 4. Phase tracking performance

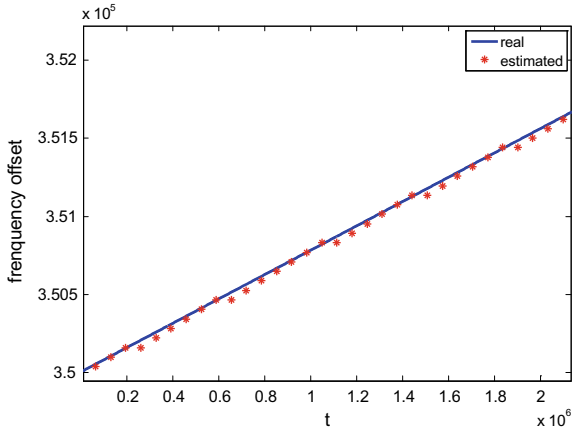


Fig. 5. Frequency tracking performance

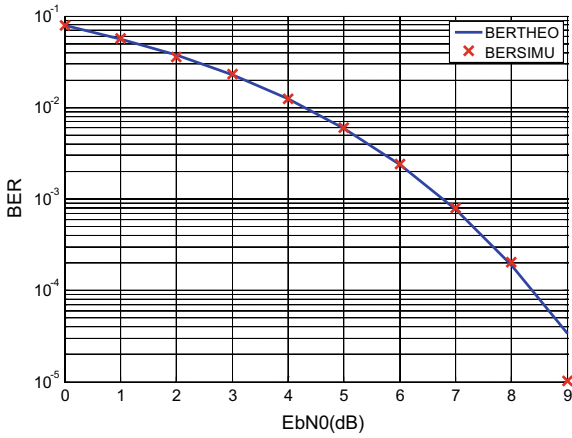


Fig. 6. Performance of the carrier synchronization algorithm

Table 1. Device utilization summary

Logic utilization	Used	Available	Utilization (%)
Slice register	16,045	866,400	1.8
Slice LUTs	11,850	433,200	2.7
Block RAM/FIFO	84	4410	1.9
DSP48	48	3600	1.3

5 Conclusion

In this paper, a open-loop carrier synchronization structure for burst communication is proposed. This algorithm requires a short preamble for capturing and roughly estimating frequency offsets. It can then track changes in frequency offset and adapt to lower signal-to-noise ratios. It can deal with a dynamic range with frequency offset range between $\pm f/2$ and frequency offset variation rate range $\pm f/20$ Hz/s. The simulated BER curve shows that it performs well at low signal-to-noise ratios. The structure is also suitable for FPGA implementation with low consumption of resources.

References

1. Morlet, Catherine, Boucheret, Marie-Laure: Carrier recovery scheme for on-board demodulation suited to low E_b/N_0 . *IEEE Globecom* **6**, 3432–3436 (1998)
2. Luise, M., Reggiannini, R.: Carrier frequency recovery in all-digital modems for burst-mode transmissions. *IEEE Trans. Commun.* **43**, 1169–1178 (1995)
3. Viterbi, A.J., Viterbi, A.M.: Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission. *IEEE trans. Inf. Theory* **29**, 543–551 (1983)
4. Yang, X., Cui, X., Lu, M., Feng, Z.: Carrier recovery using FFT and Kalman filter. In: *Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis*, vol. 2, pp. 1094–1096 (2003)



Design and Simulation of a Sector-Shaped Microstrip Antenna Fed by a T-Shaped Probe

Weiyang Mao and Lihong Song^(✉)

School of Information Science and Engineering, Harbin Institute of Technology
at Weihai, Weihai, China
songlz@hit.edu.cn

Abstract. This paper presents a sector-shaped microstrip patch antenna at an operating frequency of 1.4 GHz. In order to achieve wide beam and the normal radiation pattern, the radiation patch adopts sector. The shape of patch is the main factor affecting the performance of microstrip patch antenna. Besides, our feed network realized by electromagnetic coupling is achieved by a new method—T-shaped probe, which avoids the additional inductance introduced by the usual probe feed. Thus, the radiation efficiency of antenna is improved. In addition, the antenna has the characteristic of miniaturization. The electromagnetic simulation and optimization design of this proposed antenna are carried out using the CST Microwave Studio software package. The simulated VSWR of antenna is less than 1.5 at the working frequency. In this paper, we consider the microstrip antenna applying in some fields.

Keywords: Microstrip patch antenna · Sector-shaped · T-shaped probe · Wide beam

CLC Number: TN821

1 Introduction

Microstrip patch antenna has lightweight, miniaturization, and low profile, and now it has been widely used in mobile communication, aerospace, electronic countermeasures and radar [1]. Due to these highly attractive features, it is important to investigate microstrip antennas. It has that disadvantage of low gain and narrow bandwidth due to the characteristic of the high Q value of the microstrip patch antenna itself [2]. In recent years, in order to overcome the inherent limitations, many scholars have proposed and studied a number of technologies, such as using thick dielectric substrate with low dielectric constant or superimposed parasitic patch, changing the feeding mode, grooving on patch effectively increase working bandwidth [3]. The radiation patch shape of the microstrip patch antenna is various, and the gain of the microstrip patch antenna is different from that of the microstrip patch antenna [4]. Therefore, it is necessary to design the radiation patch shape reasonably. Other than that, as electronic devices become miniaturized, the antenna miniaturization is inevitable. By changing the radiation patch size, the antenna size can be reduced and the antenna miniaturized. Moreover, the miniaturization of antenna can reduce the cost, improve the ability of

antenna to resist natural conditions, and enhance the concealment of antenna. At the same time, wide beam characteristics can effectively improve the coverage area of antenna [5].

This paper designs and simulates a sector-shaped microstrip patch antenna fed by T-shaped probe. The design of the antenna in this paper is to achieve good performances. The designed antenna adopts the sector-shaped patch antenna elements. The simulated results reveal that the performances of the designed microstrip antenna meet the anticipated requirements. The antenna performances are provided and analyzed. In Sects. 2 and 3, the antenna structure and simulation results are described. The conclusion is provided in Sect. 4.

2 Design of the Sector-Shaped Microstrip Antenna

In this section, this paper designs the sector-shaped microstrip patch antenna, which uses T-shaped probe. The structure model of sector and T-shaped probe microstrip antenna is presented in Fig. 1. The employed dielectric substrate is FR-4, and the substrate thickness is 1 mm. The metal copper foil thickness is 0.036 mm. The designed antenna is composed of two dielectric substrates located on the upper and lower sides of the microstrip line, patch and dielectric substrate. The designed antenna is realized with microstrip circuit processing technology. The working performances are simulated and optimized through the full wave electromagnetic simulation software.

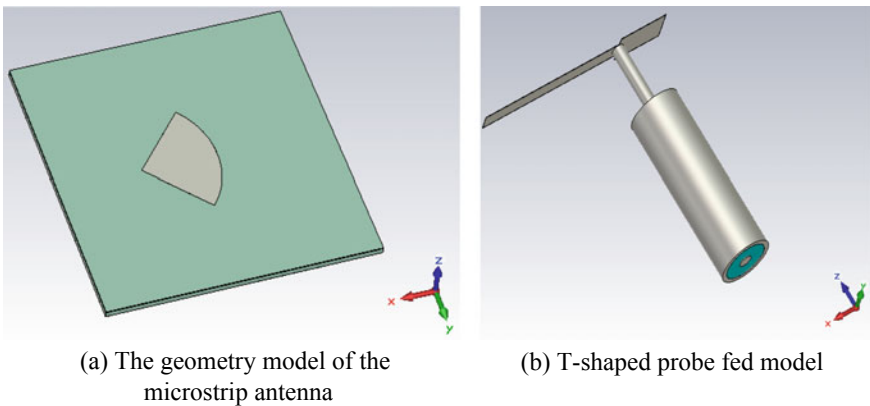


Fig. 1. The proposed antenna model

The engineering drawing of antenna structure are presented in Fig. 2a and b, respectively. The radius of the radiation patch, i.e., w_{patch} , is 82.5 mm. The length of medium plate, i.e., L_{sub} , is 300 mm. the w_{sub} is 300 mm. The engineering drawing

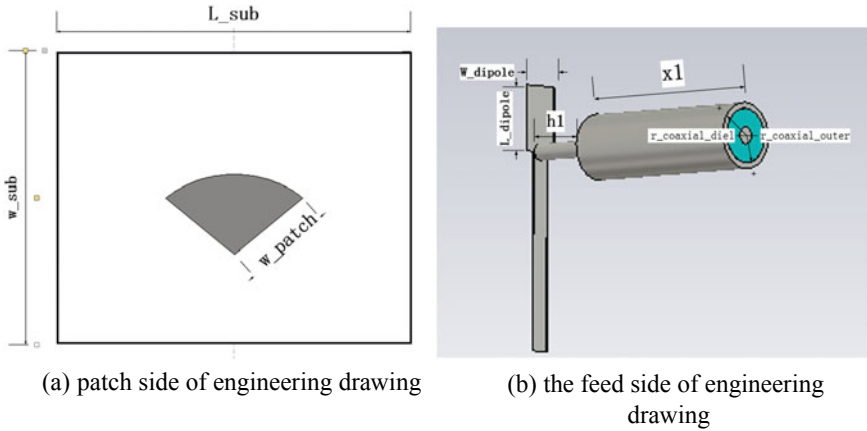


Fig. 2. The engineering drawing of the antenna

of T-shaped probe fed is shown in Fig. 2b, where $r_{coaxial_inner}$ is 3 mm, $r_{coaxial_outer}$ is 5 mm.

Adopting the theory of electromagnetic coupling, it can avoid introducing extra inductance into microstrip antenna. The most prominent advantage of the feed structure is that the coupling between the probe and the patch stimulates the electromagnetic wave radiation from the patch with better efficiency and higher gain.

The following is the formula for calculating the size of the sector-shaped radiation patch [6] fed by T-shaped probe:

$$R = \frac{F}{\left\{ 1 + \frac{c}{\pi \epsilon_r F} \left[\ln\left(\frac{\pi F}{2h}\right) + 1.7726 \right] \right\}^{1/2}} \tag{2.1}$$

$$F = \frac{8.791 \times 10^9}{f_c \sqrt{\epsilon_r}} \tag{2.2}$$

$$\epsilon_{eff} = \frac{1}{2}(\epsilon_r + 1) + \frac{1}{4} \frac{(\epsilon_r - 1)}{\sqrt{1 + \frac{12h}{R}}} \tag{2.3}$$

$$R_e = \frac{1.8412c}{2\pi f_c \sqrt{\epsilon_{eff}}} \tag{2.4}$$

Among them, the R and R_e are respectively the radius and effective radius of the patch.

3 Simulation of the Sector-Shaped Microstrip Antenna

In this paper, the electromagnetic simulation software is used to simulate the radiation performance of presented antenna. According to performance requirements, the antenna structure of sector-shaped microstrip patch antenna fed by T-shaped probe is designed. The return loss is shown in Fig. 3. It can be seen that the return loss is less than -20 dB at 1.4 GHz. The simulated VSWR of the microstrip patch antenna is shown in Fig. 4. The VSWR at the central frequency is lower than 1.5 and the impedance matching performance can meet the requirement.

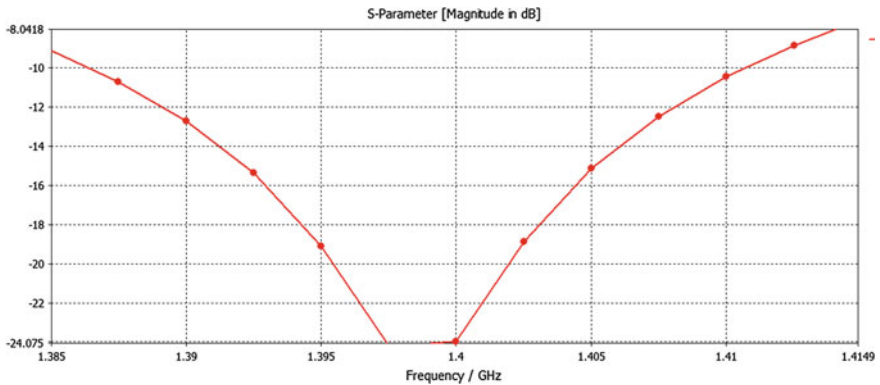


Fig. 3. The simulation results of return loss of the microstrip antenna designed in this paper

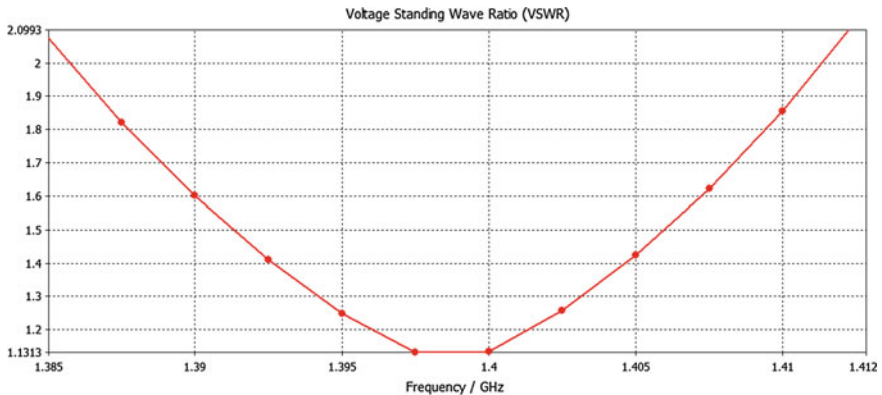


Fig. 4. The simulation results of VSWR of the microstrip antenna designed in this paper

Radiation pattern simulation result of the antenna was presented in Figs. 5, 6, and 7 at 1.39, 1.4, 1.41 GHz. It is obvious that the 3 dB beam of radiation pattern. Each graph gives the radiation pattern at the xoz plane and yoz plane respectively. From radiation pattern, antenna beam of main lobe radiation pattern was wider; beam shape

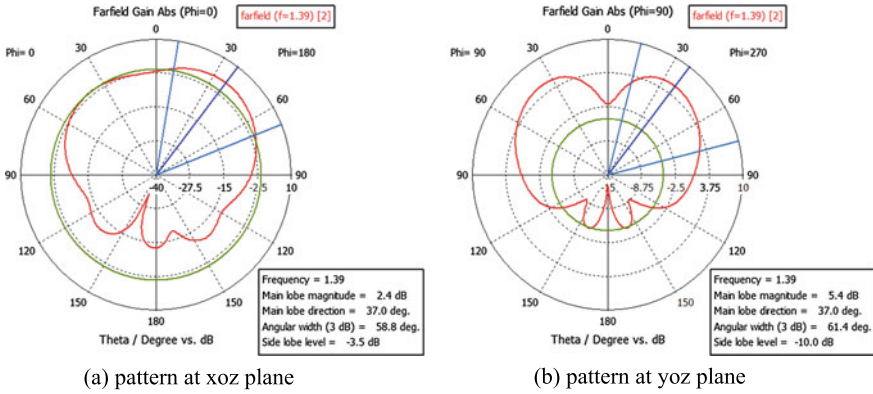


Fig. 5. Radiation pattern at 1.39 GHz

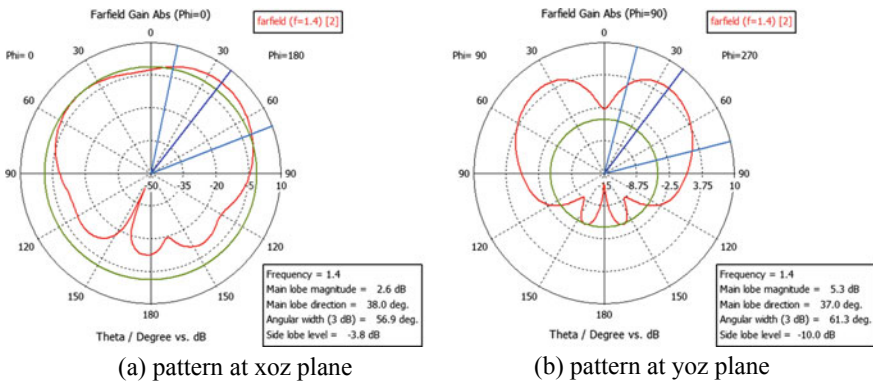


Fig. 6. Radiation pattern at 1.4 GHz

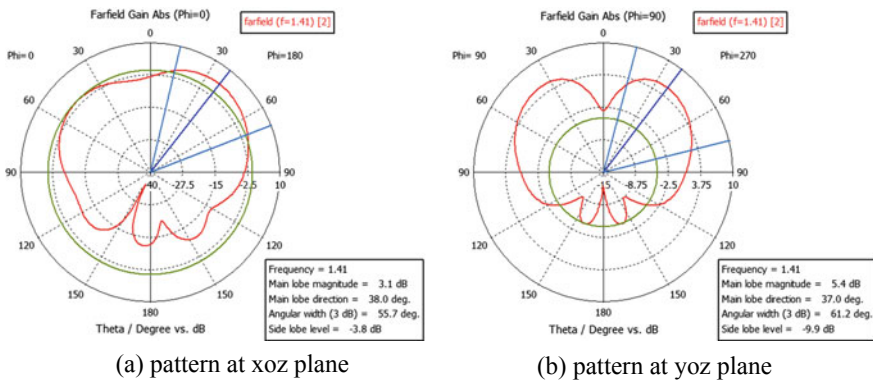
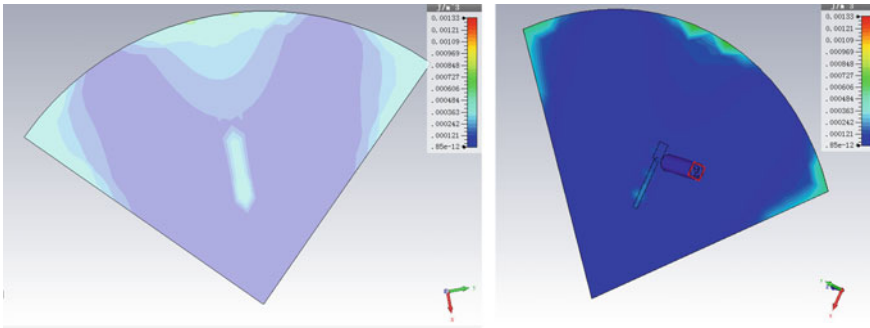


Fig. 7. Radiation pattern at 1.41 GHz



(a) The positive structure of simulated electric energy density at 1.4 GHz

(b) The reverses structure of simulated electric energy density at 1.4 GHz

Fig. 8. Simulated electric energy density at 1.4 GHz

was regular and back lobe was smaller. When the resonance frequency is 1.4 GHz, the 3 dB beamwidth of the xoz plane is 56.9° and the 3 dB beamwidth of the yoz plane is 61.3°. Absorbing material of metal back cavity and other factors caused that gain was slightly lower.

The electric energy density simulation result of the antenna was presented in Fig. 8 at 1.4 GHz. According to electric energy density, energy distribution and current transmission path are obvious.

In this section, we discuss the influence of several key parameters on the performance of the antenna circuit. Figure 9 presents the influence of parameter w_{patch} on the performance of the antenna circuit. The simulation results show that the resonance peak moves to lower frequency as the parameter w_{patch} increases. Besides, compared with other dimensions, the return loss with the parameter w_{patch} of 82.5 mm at the operating frequency point of 1.4 GHz is the best. Obviously, when the value of

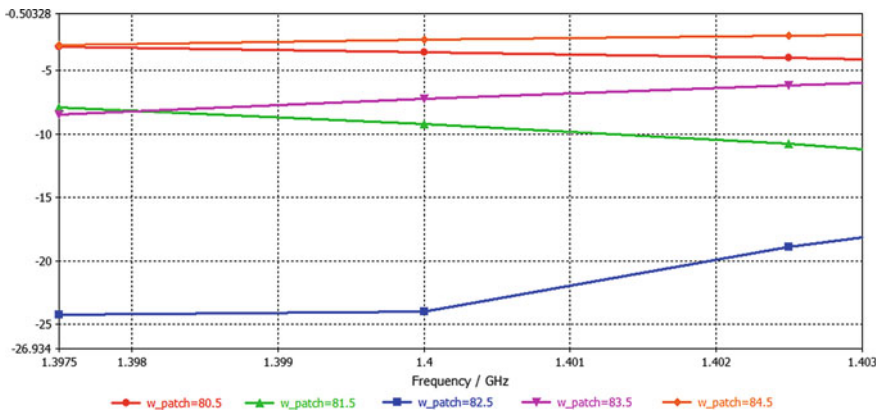


Fig. 9. The influence of parameter w_{patch} on return loss for the antenna

w_patch is 82.5 mm, the matching performance is better at 1.4 GHz. Thus, the optimal result of w_patch is 82.5 mm. Proper selection of the size of the radiation patch facilitates the design of the antenna miniaturization.

4 Conclusions

The wide beam is an important and effective performance for the microstrip antenna, which can save the space and improve the efficiency. The antenna element is fed by electromagnetic coupling. In this paper, the feed network of the antenna is achieved by a new method—T-shaped probe, which improves the radiation efficiency of antenna. For the working frequency, these antenna structures are designed and the parameters are optimized. Besides, the performances of microstrip antenna can meet the requirements of miniaturization and simple structure. The design and simulation work in this paper has certain practical value, which can be used in practical application.

Acknowledgements. This work is sponsored by the National Natural Science Foundation of China (Grant No. 61571154) the Science Foundation of Aeronautics of China (Grant No. 20160177004).

References

1. Hossain, M.I., Faruque, M.R.I., Islam, M.T.: Low SAR microstrip patch antenna for mobile phone. *Frequenz*. **69**(9–10) (2015)
2. Khan, Q.U., Ihsan, M.B.: Higher order mode excitation for high gain microstrip patch antenna. *AEUE—Int. J. Electron. Commun.* **68**(11–15) (2014)
3. Karmakar, N.C., Padhi, S.K.: Very small aperture terminal broadband shared-aperture planar antennas. *Int. J. RF Microw. Comput-Aided Eng.* **13**(3–9) (2003)
4. İmeci, Ş.T.: E-and H-shaped high gain patch antennas. *Microw. Opt. Technol. Lett.* **57**(6–9) (2015)
5. Zhang, Q., Wu, G.C., Wang, G.M., Liang, J.C., Gao, X.J.: Beam scanning antenna with wideband broadside radiation based on multilayered substrate integrated waveguide composite right/left-handed structure. *Frequenz*. **71**(1–2) (2017)
6. Dalli, A., Zenkouar, L., Bri, S.: Comparison of circular sector and rectangular patch antenna arrays in C-band. *J. Electromagn. Anal. Appl.* **04**(11–14) (2012)



Design of a Wideband Receiving Antenna for High-Frequency Ground Wave Radar

Hongbo Li, Yang Song, and Changjun Yu^(✉)

School of Information and Electrical Engineering, Harbin Institute of Technology, Weihai, China
yuchangjun@hit.edu.cn

Abstract. The operating frequency band of high-frequency surface wave radar (HFSWR) is generally 4–9 MHz, which makes the operating frequency band of the antenna a wideband. However, it is difficult for some commonly used antennas to meet this requirement, and it is necessary to widen the antenna. In this paper, a lumped monopole antenna is designed for high-frequency surface wave radar. Genetic algorithm is employed to optimize the loading position and impedance. The operating band of the loaded antenna is wider and can meet the performance requirements of the HFSWR transmitting antenna, and the antenna size is slightly smaller than other antennas in the same band.

Keywords: HFSWR · Lumped loading · Monopole antenna · Wideband

1 Introduction

High-frequency surface wave radar (HFSWR) is widely used in remote sensing and marine surveillance, such as the monitoring of sea wind and current fields, detecting of aircrafts and ships, etc. Its operating frequency is generally 3–30 MHz corresponding to a wide operating frequency band. However, some commonly used receiving antennas such as Yagi antenna and monopole antenna have a narrow operating band and belong to a narrow band antenna, which cannot meet the requirement of normal operation. Therefore, it is necessary to widen the antenna.

The loading technology is widely used in antenna design. Bharath Reddy et al. [1] introduce the design of a miniature slot antenna based on SRR and CSRR loading; A wideband meander line loaded antenna with frequency tunable capability in V/UHF band is introduced in [2]; Karim et al. [3] describe a miniaturized HF-loaded antenna for broadband applications which is designed according to a new loading strategy with lumped elements; a novel antenna design method is proposed, by loading with lumped elements, which is presented in [4] to broaden the operation bandwidth of the monopole antenna.

In this paper, a broadband scheme of HFSWR receiving antenna is proposed. The paper is organized as follows. In Sect. 2, we designed a non-loaded monopole antenna and analyzed its performance. The designed lumped loaded monopole antenna as well as the simulation results is given in Sect. 3. Conclusions are given in Sect. 4.

2 Traditional Monopole Antenna Design

The ideal monopole antenna belongs to the resonant antenna, and its current exhibits a standing wave distribution. At the resonant frequency, it has great performance, low reflection loss, and high efficiency. But as a narrowband antenna, it has a poor performance, low radiation resistance, and large reflection loss when it operates outside the optimal frequency, which causes a mismatch between the feeder and the antenna.

First, a traditional monopole antenna is simulated as a reference for the loaded antenna. A monopole antenna model is established in the HFSS software, the height of the antenna H is set to 7.5 m and the radius R is set to 2.5 cm. The simulation results are shown in Figs. 1 and 2.

Figure 1 shows the simulation results of the traditional monopole antenna's reflection loss. It shows that the return loss decreases slowly in the 4–8 MHz band and decreases sharply in the 8–9 MHz band. The resonance frequency of the antenna is 9 MHz, and the reflection loss gradually increases in the 9–12 MHz band.

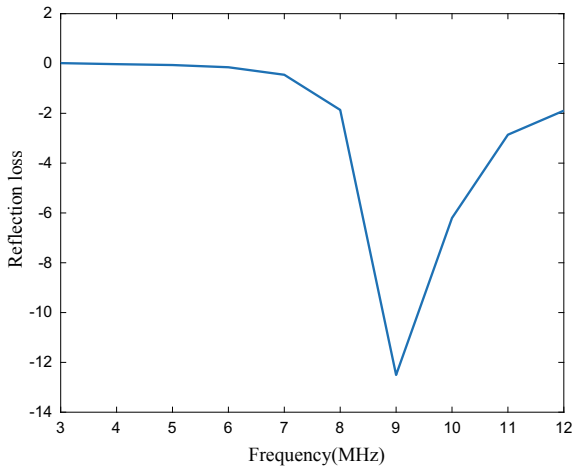


Fig. 1. The reflection loss of the monopole antenna

Figure 2 shows that the voltage standing wave ratio (VSWR) of the antenna is less than 2 in a very narrow frequency band around 9 MHz and relatively high before 8 MHz. There is a serious mismatch between the antenna and the feed end.

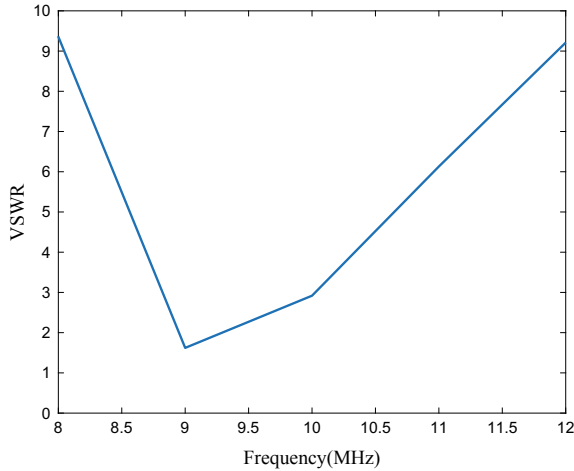


Fig. 2. The VSWR of the monopole antenna

It can be seen that the monopole antenna cannot meet the requirement of the HFSWR receiving antenna in the 4–9 MHz band, and the narrowband characteristics of the monopole antenna cannot be optimized simply by changing the size of the antenna. Document 5 indicates that the antenna loading is a very common and effective technique for widening and miniaturizing the antenna. Therefore, in order to meet the requirements for miniaturization of the receiving antenna, a monopole antenna needs to be loaded.

3 Lumped Monopole Antenna Design

As for the lumped loading, the issue that needs resolving is the loading position and impedance. In this section, the genetic algorithm is used to determine the loading position and impedance.

3.1 Fixed Position Loaded Antenna

In order to observe whether the impedance characteristic curve will change smoothly and whether the resonant frequency point will decrease after the antenna loading element, we first load the antenna with a fixed position. According to the conclusion of Document 6, the impedance of a single-loaded antenna changes drastically, while the impedance of a double-loaded antenna changes moderately. Therefore, the parallel RLC circuit is selected to be loaded at two fixed positions. The height of the antenna is still 7.5 m, and two loading positions are selected as $H_1 = 2.1875$ m and $H_2 = 3.5$ m. The antenna structure is shown in Fig. 3.

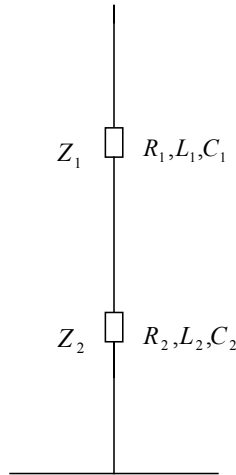


Fig. 3. Fixed position loaded monopole antenna schematic

At the two positions H1 and H2, RLC parallel circuits are, respectively, loaded, and the component parameter values are optimized using genetic algorithm. The optimization range and results of the parameters are shown in Table 1. The optimized values are brought into the antenna structure and the antenna is simulated by HFSS software. The simulation results are analyzed in the next section to compare with the optimized position loaded antenna.

Table 1. The optimization range and result of loading element

Variable	Optimization ranges	Optimization results
R (Ω)	0–1500	1361.3
$L1$ (μH)	0–100	4.39
$L2$ (μH)	0–100	9.11
C (pF)	0–20	2.44

3.2 Optimized Position Loaded Antenna

In the previous section, antenna loading can improve the VSWR in the low-frequency part. However, the fixed loading position limits the freedom of the parameter change, so the performance of antenna is not ideal. In order to improve the performance of antenna, in this section, the antenna height, the loading position, and the parameter values of the parallel RLC components are used as optimization variables. In this way, the number of variables is increased to nine. Optimization goal of antenna is setted as $VSWR < 5$ and $G > -6\text{dB}$. The optimization range and results of the parameters are shown in Table 2. Simulating the optimized antenna in the HFSS software, the results are shown in Figs. 4, 5 and 6.

Table 2. The optimization range and result of the variables

Variable	Ranges	Results	Variable	Ranges	Results
H (m)	7–8	7.83	$L1$ (μH)	0–100	0.64
$H1$ (m)	1–3.4	1.43	$L2$ (μH)	0–100	10.35
$H2$ (m)	3.5–6.4	5.99	$C1$ (pF)	0–20	133.35
$R1$ (Ω)	0–1500	321.8	$C2$ (pF)	0–20	9.31
$R2$ (Ω)	0–1500	326.45			

Figure 4 shows the VSWR of the fixed position loaded antenna and the optimized position loaded antenna.

For the fixed position loaded antenna, the VSWR is smallest at 5 MHz, which indicates that the resonant frequency has decreased after loading.

The VSWR of the optimized position loaded antenna is around 5 in the operating frequency band, which is lower than the fixed position loaded antenna. Documents 7 and 8 indicate that when the VSWR changes smoothly, the VSWR can be further reduced to be less than 2 by adding a matching network, which can meet the performance requirements of the receiving antenna.

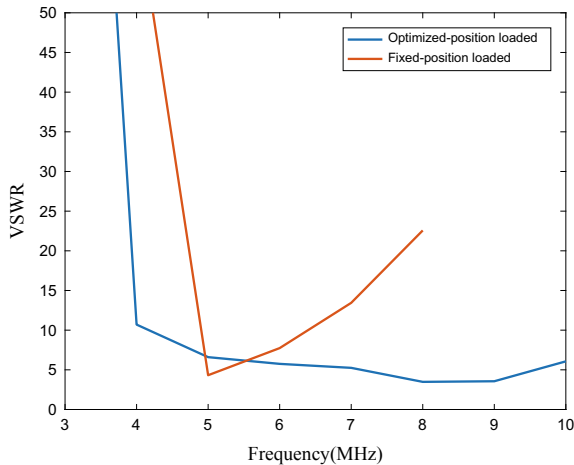
**Fig. 4.** The VSWR of the loaded antenna

Figure 5 shows the gain of the fixed position loaded antenna and the optimized position loaded antenna.

For the fixed position loaded antenna, in the 4–8 MHz band, the gain of the antenna is relatively low, which is due to the large reflection loss and energy loss caused by the loaded impedance element.

As the frequency increases, the gain of the optimized position loaded antenna gradually increases, indicating that the energy loss is high in the low-frequency band. It can be seen that the gain of the optimized position loaded antenna is higher than the fixed position loaded antenna.

Figure 6 shows the efficiency of the fixed position loaded antenna and the optimized position loaded antenna.

For the fixed position loaded antenna, it can be seen that the antenna has the highest radiation efficiency at 5 MHz, and as the frequency deviates from 5 MHz, the efficiency gradually decreases, which shows that the loaded impedance element will introduce losses, making the overall efficiency of the antenna low.

The efficiency of the optimized position loaded antenna gradually increases first and then decreases slightly. However, the overall efficiency is low due to the loss caused by the loading element. Document 9 indicates that the receiving antenna does not require too high efficiency to operate normally when the external noise is greater than the internal noise of the receiver. Therefore, it is possible to obtain a wider frequency band by sacrificing a certain efficiency of antenna.

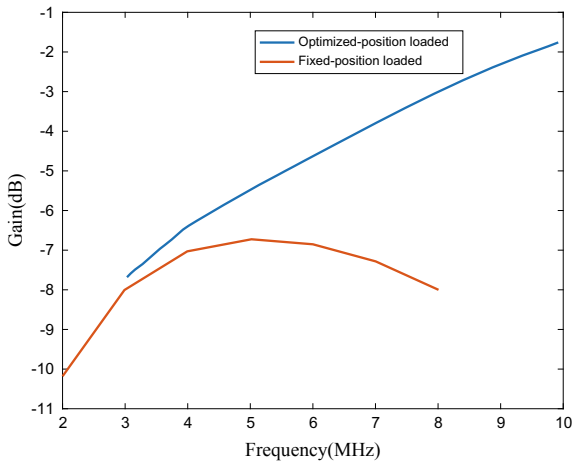


Fig. 5. The gain of the loaded antenna

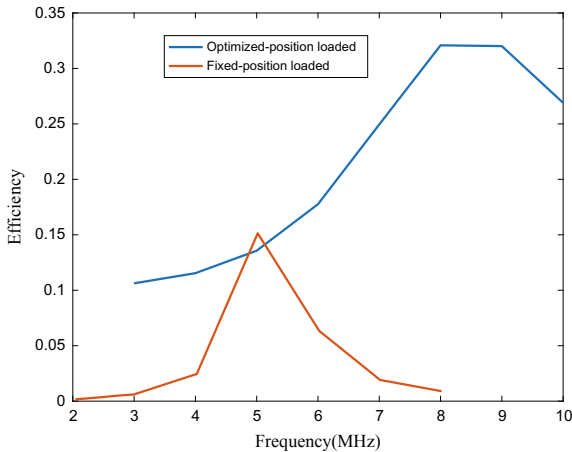


Fig. 6. The efficiency of the loaded antenna

The above analysis shows that the optimized position loaded antenna can work normally in the 4–9 MHz band. The monopole antenna designed in this way can meet the broadband requirement of the receiving antenna. Based on this, adding matching network can further optimize the performance of the receiving antenna.

4 Conclusion

This paper analyzes the electrical characteristics of the monopole antenna in the 4–9 MHz band under the background of the application of the HFSWR receiving antenna. The monopole antenna is loaded by lumped loading, and the loading positions and component parameters are optimized by genetic algorithm. A monopole antenna of about 7.5 m is designed. Its impedance and radiation characteristics are simulated. The results show that it can meet the specifications of the HFSWR receiving antenna. Moreover, the size of this antenna is smaller than that of other antennas at the same frequency, and the antenna operating band has also been broadened.

Acknowledgements. This work is supported by National Natural Science Foundation of China under Grant No. 61571159 and the National Marine Technology Program for Public Welfare under Grant No. 201505002.

This work was supported in part by the National Key R&D Program of China under Grant 2017YFC1405202, in part by the National Natural Science Foundation of China under Grant 61571159 and Grant 61571157, and in part by the Public Science and Technology Research Funds Projects of Ocean under Grant 201505002.

References

1. Bharath Reddy, G., et al.: Miniaturization of microstrip slot antenna using SRR and CSRR loading. In: 2018 3rd International Conference on Microwave and Photonics (ICMAP), pp. 1–2 (2018)
2. Molaee-Ghaleh, F., et al.: Wideband tunable meander-line loaded antenna with dual-mode capability in V/UHF frequencies. In: 2017 IEEE Asia Pacific Microwave Conference (APMC), pp. 437–439 (2017)
3. Lucas, C., et al.: Design of electrically loaded bifolded monopole antenna using genetic algorithm. In: 2006 12th International Symposium on Antenna Technology and Applied Electromagnetics and Canadian Radio Sciences Conference, pp. 1–4 (2006)
4. Weiping, C., et al.: Study and design of ultra-broadband loaded monopole antenna. In: 2016 11th International Symposium on Antennas, Propagation and EM Theory (ISAPE), pp. 453–455 (2016)



Optimal Design of S Band with Graphene Frequency Multiplier

Huili Chen^(✉), Yong Fang, Xiaoling Zhong, Qingyan Song, Xueshi Hou, and Haoxuan Sheng

College of Information Science & Technology, Chengdu University of Technology, Chengdu 610059, China
2074811788@qq.com

Abstract. A graphene frequency multiplier is proposed in this paper. The nonlinear characteristics of graphene are similar to the reverse parallel diode. In order to improve the efficiency of GFM, we use the method of recovering the fundamental and fifth harmonic wave which utilizes the way of branch recovery. According to this mechanism, an S band graphene frequency multiplier is designed. During the working frequency between 500 MHz and 3 GHz, the minimum conversion loss of -20.81 dBm can be obtained at 830 MHz and the input power is 18 dBm.

Keywords: Graphene · Nonlinear electromagnetic response · Frequency multiplier · High efficiency · Branch recovery

1 Introduction

Microwave and millimeter wave signal sources are widely used. Including wireless routers, satellite communications, radar, and other electronic systems, frequency multiplier is one of the methods of signals processor. In the microwave and millimeter wave band, Schottky diode is usually used to design frequency multiplier. Compared with high-frequency active oscillator, Schottky diode is the ideal choice for making odd harmonic multipliers.

In recent years, graphene as a two-dimensional material has attracted much attention of many researchers due to its unique electrical characteristics. This makes signal source applications possible [1]. 2007, and S. A. Mikhailov put forward to the proposal.

Graphene has a strong nonlinear electromagnetic response. In 2008, 2014, and 2015, S. A. Mikhailov further proved its theory. He pointed out that the output harmonic current of the graphene nonlinear device became slower with the decreased harmonic orders of frequency, which is compared with the traditional nonlinear dual port devices, such as the Schottky diode. In addition, under the stimulation of electromagnetic waves, the output current of the graphene circuit contains only fundamental and odd harmonic waves (such as 1ω , 3ω , 5ω , 7ω , etc.). Graphene circuits have the characteristics of suppressing even harmonic waves. It is suitable perfectly for making nonlinear devices (such as frequency multipliers, etc.).

In 2010, utilizing the nonlinear electromagnetic properties of graphene, M. Dragoman developed a graphene multiplier based on CPW which is working at the range of millimeter wave. He verified the nonlinear theory [2] of graphene for the first time. In 2011, R. Cambor developed a three-frequency multiplier [1] with microstrip gap which is loaded with multilayer graphene. When the input power is 20 dBm, the minimum conversion loss of the three-frequency multiplier is -25.32 dBm.

In 2017, graphene frequency tripler design using reflector networks as frequency tripler based on graphene nonlinear electromagnetic response is developed. The non-linearity of graphene is similar to that of anti-parallel diodes. The output frequency of this tripler can cover 12–30 GHz. The harmonic signal recovery using the reflector networks can significantly reduce the conversion loss [3].

From the above analysis, it can be seen that due to the strong nonlinear electromagnetic response of graphene, the output signals of graphene circuits include only basic harmonics and odd harmonics. Graphene is very suitable for frequency multiplier. However, the efficiency of the reported graphene frequency multiplier is low. This paper optimizes the design of graphene multiplier. Finally, an optimized graphene frequency multiplier was manufactured and tested.

2 Explanation of the GFT

2.1 Analysis of the Nonlinear Characteristics of GFM

The AC current of the graphene's electron nonlinear behavior can be expressed in Eq. (1) which is given below:

$$j_x(t) = en_s V \operatorname{sgn}(\sin \omega t) = en_s V \frac{4}{\pi} \left(\sin \omega t + \frac{1}{3} \sin 3\omega t + \frac{1}{5} \sin 5\omega t + \dots \right) \quad (1)$$

It can be seen from Eq. (1) that the output signal of the graphene circuit contains only fundamental and odd harmonics. The power ratio of output harmonic signal is shown in Fig. 1. For the output signal, the power ratio of fundamental signal accounts for about 100%, the power ratio of the third harmonic signal accounts for about 11.11%, and the power ratio of the fifth harmonic signal accounts for about 4%. The third and fifth harmonic signals account for about 15% of the total power. Therefore, in order to improve the efficiency of the frequency multiplier, it is an effective way to recover the low-order harmonic signal [3].

2.2 Optimal Design of GFM Transform Loss

The design method of graphene odd harmonic frequency multiplier is that it uses the way of the graphene odd harmonic frequency multiplier to determine the recovered signal frequency component according to the input signal and frequency doubling times; the input signal is the fundamental wave and the frequency doubling time is N , in which the output signal is Nf_0 after frequency doubling, and the signal needed to be recovered of the frequency component is $(2n + 1)f_0$, $N = 1, 2, 3, 4, \dots$ and $(2n + 1)N$.

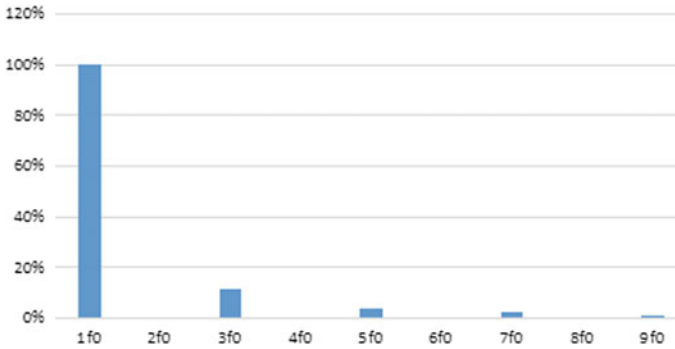


Fig. 1. The power ratio of output harmonic signal

The input reflection network is the matching state with the frequency f_0 and the frequency component of the recovered signal is the grounding state.

The output reflection network is matched state with the frequency Nf_0 , and the fundamental wave and the harmonic components which are needed to be recovered are fully reflected, and the full reflection means that the basic wave and the recovered harmonic components are reflected to the input reflection network through the graphene.

2.3 The Work of Graphene Frequency Multiplier

The graphene multiplier is connected with the excitation signal source and it generates the fundamental wave. After the fundamental wave stimulates the graphene for the first time through the input reflection network, the fundamental waves and odd harmonic components are generated at the back end of the graphene. The odd harmonic components include the output signal Nf_0 , and the signal frequency component $(2n + 1)f_0$ which needs to be recovered, $n = 1, 2, 3, 4, \dots$ and $(2n + 1) \neq N$.

The output reflection network matches the output signal, and it makes the fundamental wave and the recovered harmonic component reflect to the front end of the graphene.

The signal frequency component that needs to be recovered is grounded by the input reflection network. The fundamental wave is reflected back to the front end of the graphene. When it passes through the graphene, it can stimulate the graphene for the second time and produce the odd harmonic components at the back end of the graphene again, and the fundamental wave itself continues to be grounded through the excitation signal source.

In order to recover the fundamental and fifth harmonic power signals, the topology of the proposed GFM is shown in Fig. 2a. GFM consists of three parts: input signal reflection network, graphene, and output signal reflection network. The reflected network of the input signal mainly contains the third and fifth harmonic signals and the fundamental wave is grounded through the microwave source. The reflected network of the output signal is mainly to restore the fundamental and fifth harmonic signals, and it matches the impedance of the third harmonic signal at the same time. According to

Formula (1) and combined with the nonlinear electromagnetic response of graphene, there are two ways to improve the efficiency of GFM.

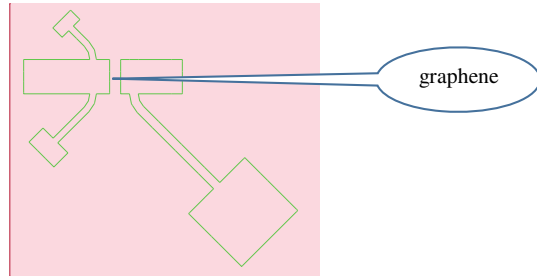


Fig. 2. The topology of the GFM

GFM can be equivalent to anti-parallel diode pair multiplier circuit, and it is shown in Fig. 2b. Then, the output signal of the GFM contains only the frequency $(mwf_0 \pm nwfs)$, and $(M + n)$ is an odd number. For example, $M = 2$ and $N = 1$, and the total current of GFM contains the third harmonic signal.

By multiplier mode and mixer mode, GFM gains the third harmonic signal. Therefore, output reflector network enhances the output power of the third harmonic signal. Figure 3 shows the scattering parameters of input and output reflector networks which are simulated using Ansoft HFSS [4].

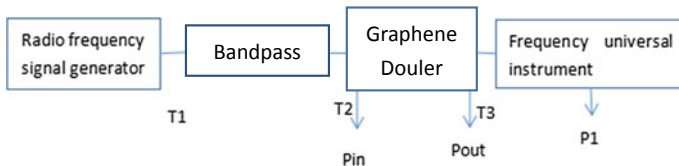


Fig. 3. Three-frequency doubling test flowchart

3 The Test of Frequency Multiplier

First, the third frequency multiplier test requires complete circuit structure. It includes the output of the radio frequency of fundamental signal, bandpass filter, frequency multiplier, and it needs to observe the spectrum and record the data in a spectrum analyzer. We can clearly see the whole test process through Fig. 3.

The output power of RF signal generator is 19 dBm, T1 is the coaxial line connecting the signal generator and filter. T2 is the coaxial line connecting the output end

of the filter and the input end of the frequency multiplier cavity, T3 is the coaxial line connecting the output end of the cavity and the spectrum instrument, Pin is the power of the microwave signal when inputting cavity and Pout is the power for the signal to output of the cavity, and P1 is the power of the signal recorded by the spectrum analyzer.

The test range is from 750 MHz to 1.1 GHz. We take 830 MHz test as an example, as shown in Fig. 4. The output frequency is 830 MHz and the output power is set to 19 dBm on the N9310A RF generator.

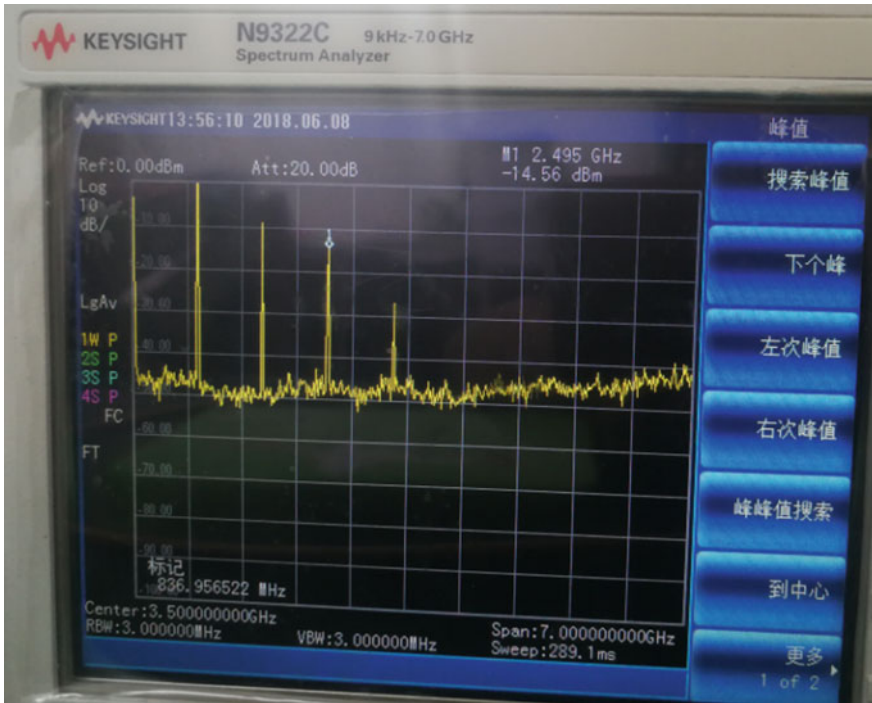


Fig. 4. 830 MHz three harmonic spectrum diagram

From the diagram, we can observe that, except for the fundamental wave of the experimental operation, the frequency spectrum of 2.49 GHz is very obvious, and the spectrum purity is very high and its power is -14.56 dBm.

In addition to the 900 MHz fundamental wave signal, multiple sets of signals are also tested. As shown in Fig. 5, a total of 15 frequency bands from 750 MHz to 1.1 GHz are tested. The basic power settings are 19 dBm, and the data include first, second, third, and fourth harmonic frequency spectrums, as shown in Fig. 5.

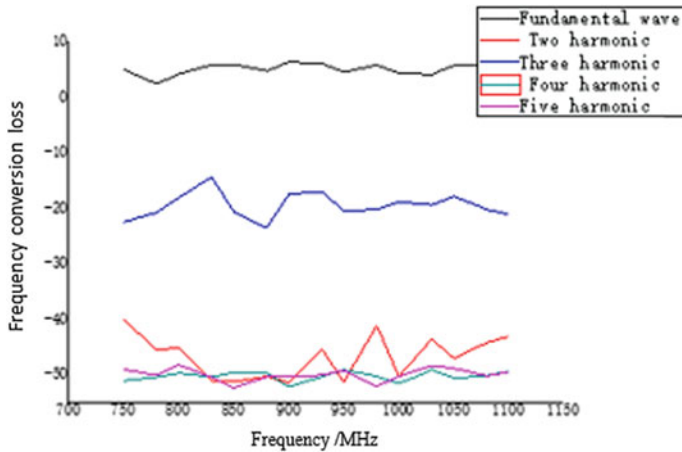


Fig. 5. Test results of graphene doubler

4 Conclusion

A frequency multiplier based on nonlinear electromagnetic response of graphene is studied in this paper. The nonlinearity of graphene is similar to that of anti-parallel diode. The use of branch recovery can significantly reduce conversion loss. With the increase of input power, the conversion loss gradually decreases. The graphene doubler is optimized to a large extent.

Acknowledgements. This work is supported financially by the National Natural Science Foundation of China under Grants 41574137.

References

1. Cambor, R., et al.: Microwave frequency tripler based on a microstrip gap with graphene. *J. Electromagn. Waves Appl.* **25**, 1921 (2011). <https://doi.org/10.1163/156939311798072090>
2. Dragoman, M., et al.: Millimeter-wave generation via frequency multiplication in grapheme. *Appl. Phys. Lett.* **97**, 093101 (2010). <https://doi.org/10.1063/1.3483872>
3. Fang, Y.: Graphene frequency tripler design using reflector networks. <https://doi.org/10.1587/elel.15.20171190>
4. High Frequency Structure Simulator, HFSS, vol. 13. Ansoft Corporation



Intelligent Street Lamp Management System

Shimin Wang^(✉) and Yongjie Yang^(✉)

School of Electronics and Information, Nantong University, Nantong 226019,
China

1171794408@qq.com, yang.yj@ntu.edu.cn

Abstract. In order to actively respond to the national strategy of green lighting, this paper designs and implements a smart street lamp management system based on solar street lamps, which uses multilayer distributed structure. With the embedded ARM microprocessor as the core, the bottom terminal combines with NB-IOT transmission technology, sensor technology, GPS positioning technology, electricity collection technology, and so on. The system is equipped with a good human–computer interaction platform, and collects information to the top-level database terminal with the network transmission technology of narrow band, realizing the effective collection and storage of city streetlight information. Through scientific and effective control and management of urban street lighting facilities such as line control, point control, and spot survey, the intelligent street lamp management system saves energy and improves work efficiency, and improves the modern management level and scientific means of urban lighting facilities. The system which has good stability and can effectively manage the solar streetlights was tested on the spot at Nantong University, so it has a good value of application and popularization.

Keywords: Street lamp management · ARM · Database server · Power collection · NB-IOT

1 Introduction

With the rapid development of urban construction and the continuous improvement of road lighting quality, the government and people pay more and more attention to street lighting and urban image [1]. With the rapid development of road construction in urban areas, the number of street lamps is increasing. However, the management departments still use the original manual operation and timing control to complete the task much heavier than before. This cannot solve the inherent contradiction between lighting rate and energy conservation [2]. Therefore, it has become an urgent need for urban lighting management to study a set of advanced urban road centralized monitoring and management methods to reduce labor intensity, save resources, reduce costs, and improve the level of informatization, digitalization, high efficiency, and energy saving of urban road lighting management [3].

2 System Architecture

The architecture of the intelligent street lamp management system is shown in Fig. 1. The multilayer distributed structure is divided into monitoring layer, transmission layer, and management layer from bottom to top. The transport layer is the communication bridge of the system. The stability of its working condition directly determines the quality of the data transmission of the whole system [4]. Considering that the traffic capacity of urban street lamps is large and the overabundance of street lamps is too dense to cause the individual to communicate, the NB-IOT communication mode is used to access the terminal server terminal. The underlying monitoring layer is based on the monitoring terminal, and each monitoring terminal sets an NB-IOT node, and each node has a unique identification (IMEI). Some nodes do data transmission with the server terminal. NB-IOT is not real-time online, and the server cannot communicate with each other in real time [5]. This is the possibility of packet loss. There must be a retransmission and verification mechanism. So the system uses HUAWEI's CoAP mechanism to avoid packet loss and conflict, and ensure that the underlying monitoring terminal communicates well with the server. All the streetlights are distributed in distributed control. The information of light, electricity, temperature, humidity, and location of the ground floor is uploaded to the server by network [6]. The upper computer can query and handle the information of the database by reading the information of the database. The location information of the street lamp and the possible damage information are printed to the maintenance workers, and the maintenance workers can get the position of the street lamp for the first time, so as to carry on the

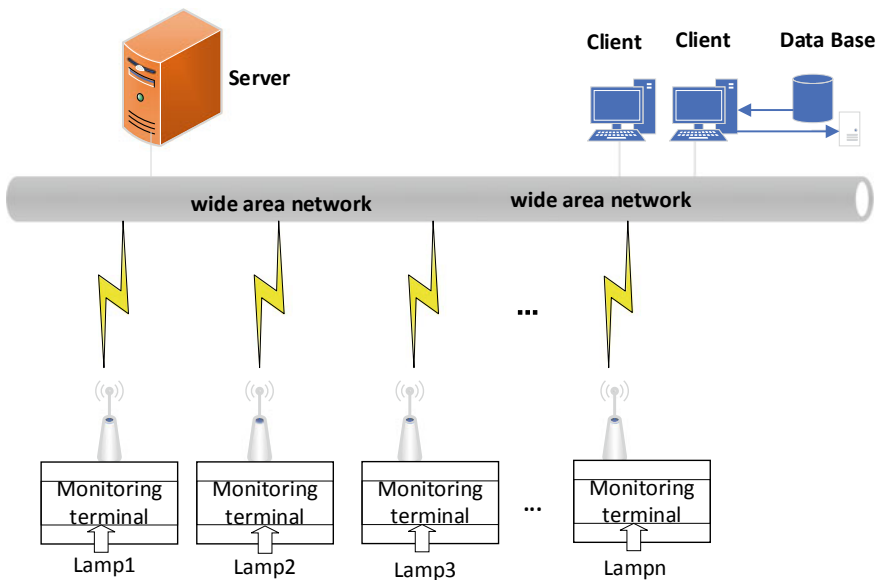


Fig. 1. The system architecture

effective and rapid response and deal with the problem of the street lamp in time, realizing the intellectualized management of the street lamp and the intelligent management of the city.

3 System Terminal Hardware Design

The hardware design is mainly concentrated on the monitoring layer of the system. Taking a monitoring terminal as an example, it introduces the monitoring of a street lamp. The design of monitoring terminal is more complicated. The design of many modules and multiple functions plays a key role in the street lamp management system. The monitoring terminal can be divided into two parts: data acquisition and electricity management. The two parts can achieve stable data upload and dispatch through reasonable task arrangement and scheduling [7].

3.1 Main Hardware Design of Data Acquisition Part of Monitoring Terminal

According to the functions required, the core processor of the acquisition part adopts ST’s enhanced 32-bit RISC core chip -STM32F103C8T6 based on Cortex-M3 architecture. The chip has a maximum operating frequency of 72 MHz, a built-in high-speed memory (up to 512 KB of memory and 64 KB SRAM), a rich I/O port and peripherals connected to two APB buses, working temperature range from -40 to +105 °C, and a series of power saving modes that fully meet the design requirements of the system. The collection of peripheral hardware mainly includes power acquisition, illumination acquisition, GPS location acquisition, fault collection, and so on. The interface design between the main chip and the peripheral modules is shown in Fig. 2.

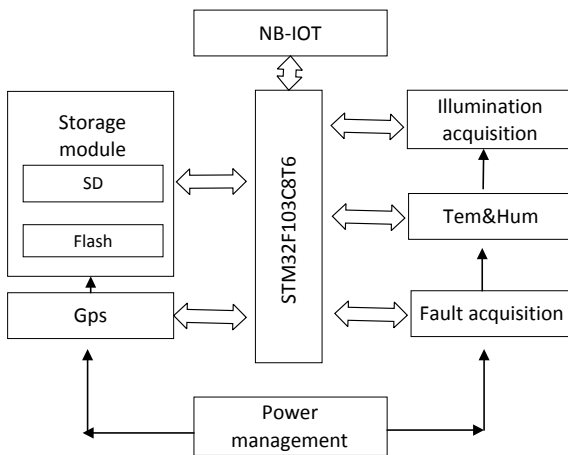


Fig. 2. Hardware structure diagram of data acquisition

The data acquisition part based on STM32F103C8T6 mainly implements the following functions:

1. Data acquisition function: Using photosensitive resistor, SHT20 temperature, and humidity sensor to achieve environmental lighting and temperature and humidity collection. Through the L70-R GPS module based on ROM version, the specific location of the street lamp is obtained, so that the maintenance personnel can quickly realize the street lamp positioning. The fault detection section will be involved in the power management section [8].
2. Interaction and transmission function: Data collected from the data collection section, the data is sent to the top layer server through NB-IOT, the data format is CJJSON format, and the communication baud rate is 115,200 bit/s. In view of the instability of the network and the problem of packet conflict, when the above problems arise, the local data is temporarily present in the SD card, and all the data stored in the SD card will be reuploaded to the next transmission cycle.
3. Other functions: Briefly explain the power supply function in the power management section here. The system power supply is from the solar charging pool, the battery voltage is 12 V, through a LM2596 power chip to 5 V DC, and then a AMS1117-3.3 V power chip is used to obtain 3.3 V DC, providing power supply by the main chip [9]. The maximum current is less than 1A of the maximum output current of the power chip. The hardware implementation and function of the power management part of the monitoring terminal will be introduced in the next section.

3.2 Main Hardware Design of Power Management Part of Monitoring Terminal

The last section introduces the hardware components and related functions of the data acquisition part, and the hardware design of power management is the core part of the whole system.

In the intelligent street lamp management system, the power management part uses the BQ27542-G1 power management chip, uses the impedance track algorithm to monitor the electric quantity, and provides the battery residual quantity (mAh), the charging state (%), the endurance time, the battery voltage (MV), the temperature (centigrade), and so on, and supports the battery with the highest capacity of 14,500 mAh [10]. Detection can also provide detection function for internal short circuit or battery terminal disconnection. In order to obtain accurate electricity state, it is necessary to make the learning cycle of the BQ27542-G1 power management chip and use the TI company to provide the Battery Management Studio software for the cycle learning of the battery, and communicate with the software system by the IIC interface. The cycle learning steps are shown in Fig. 3.

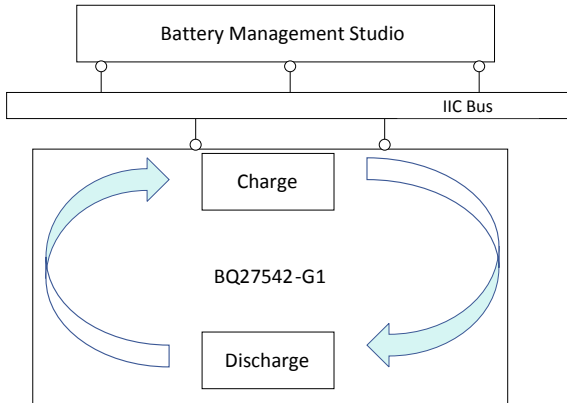


Fig. 3. Periodic learning

The learning cycle starts with a discharge and a relaxed battery (the battery voltage is stable and is in a low voltage state), usually in 3.0–3.3 V. Then, the battery is charged to the full (battery management software must be detected to be full of electricity). Once the charge is stopped, the battery voltage is at a high voltage, usually between 4.1 and 4.2 V, and then the battery is discharged again. Then, the battery management software uses another voltage measurement and determines the state of the charge it represents, through two known types. The charging state and the maximum volume of the battery are calculated through the Kulun gauge [11].

4 System Software Design

The system software is mainly composed of two parts: the bottom monitoring terminal and the top-level management host computer software. This paper mainly introduces the design of management software.

4.1 Software Design of Monitoring Terminal

The monitoring terminal has many functions, such as complex logic and time sequence, and the idea of sampling and blocking and subfunction calling when programming. The sub-functions of the monitoring terminal include SD card, FLASH module, GPS module, illumination module, and temperature and humidity module. The main structure of the software design of the monitoring terminal is shown in Fig. 4. The main program can directly call the sub-functions to realize the expected function.

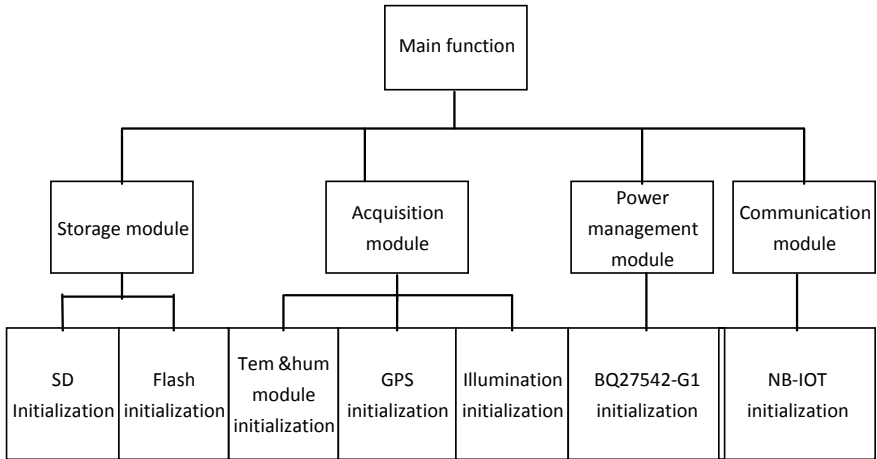


Fig. 4. Software design for monitoring terminal

4.2 The Design of the Management Software of the Upper Computer

Visual Studio software is used as the development platform and written in C# language. The management software can be divided into data query, personnel management, and common tools from function division. According to the use authority, the management software can be divided into non-internal management personnel and internal management personnel, non-internal management personnel only does inquiry and other basic works. The right to use, and the authority of the internal manager can use all the functions. Management software plays an important role in the whole intelligent street lamp management system [12]. On one hand, the management software uses UDP protocol and the server to carry out data transmission. Management software receives data from the monitoring terminal, parses data, and stores it in the database. On the other hand, the management software gets the number from the SQL SERVER database. According to it, the management software communicates with the monitoring terminal in a similar way of broadcasting, and the monitoring terminal will receive instructions from the management software at the same time. Only when the ID is consistent with its own device, the instruction will be stored in the monitoring terminal, and the action is executed according to the instruction. The data lattice is used in the CJSON format. The design of the management software is shown in Fig. 5.

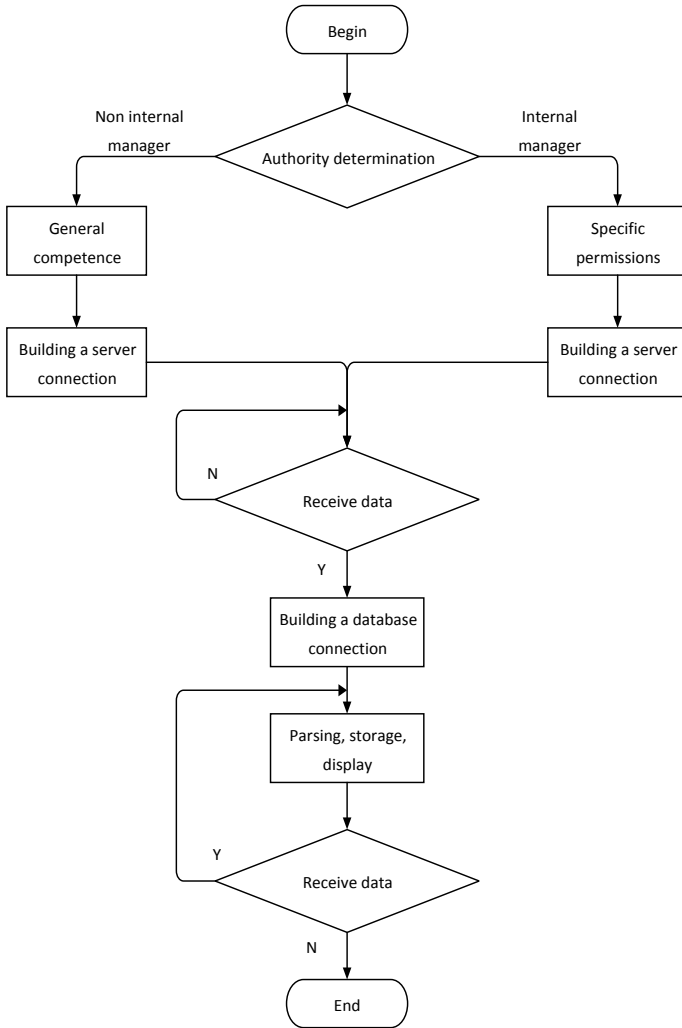


Fig. 5. Management software design

5 Field Test Results

The system has been tested at Nantong University. From the field test results, the modules of the acquisition system can work steadily without interference between them. It can collect information such as power, illumination, GPS location, and upload to the server according to the expected requirements. The success rate of collecting and uploading the streetlight information is 98.23%. The good feedback mechanism can realize the effective control of the street lamp and realize the monitoring and

management of the city street lamp. Through the upper computer software, the manager can see the streetlights in each area of the city. Compared with the traditional manual inspection, it can realize the intelligent management of the streetlights in the city, which

Table 1. Street lamp status information

Reg	No.	Hum (%RH)	Lig	Tem (°C)	Pow	Gps
ch	1	40.2	62.1	26.1	80.3	Get
ch	2	45.3	61.5	25.3	82.4	Get
ch	3	44.5	63.0	27.5	86.1	Get
ch	4	45.6	65.4	28.6	83.5	Get
ch	5	48.7	56.6	29.5	84.6	Get

is in line with the street lamp management requirements of the intelligent city. Table 1. is the street lamp status information obtained by the upper computer software.

From the data statistics of the field test, the distributed street lamp management system can achieve the desired effect better. The street lamp status information can be uploaded to the host computer management software in time and effectively, and it is obtained by the management. In view of packet loss in information uploading, the antenna quality will be changed in the later stage, and the anti-collision work of uploading data can be optimized. The success rate of data uploading can reach 100%.

6 Conclusions

The design of intelligent street lamp management system based on ARM Cortex M3F is proposed and realized for the first time. The system integrates embedded technology, network technology of narrow band, and database technology to realize the intelligent management of urban street lamps, and has taken a great step to the process of intelligent city [13]. The concept proposed in this paper has changed the current situation of urban streetlight management, which not only can solve the problems of urban street lamp management well, but also have a very valuable benchmarking on other infrastructure management in smart cities [14].

Acknowledgements. This work was supported by the open research project of Nantong Intelligent Information Technology Research Center (KFKT2017B05) of Nantong University and First phase project of Jiangsu university brand specialty construction project (PPZY2015B135). In addition, it has also been supported by the graduate student innovation program of Nantong University. The authors thank the three anonymous reviewers for their helpful suggestions.

References

1. Gong, Y., Han, P.: Research on energy-saving scheme based on led street lamp management-system. *Appl. Mech. Mater.* **448–453**, 2850–2855 (2014)
2. Zhou, Y.: Design of intelligent street lamp management system under the background of wisdom city: a GIS-based exploration. *J. Shunde Polytech.* (2017)
3. Daely, P.T., Reda, H.T., Satrya, G.B., et al.: Design of smart LED streetlight system for smart city with web-based management system. *IEEE Sens. J.* **99**, 1–1 (2017)
4. Jamaluddin, A., Aini, A.N., Adhitama, E., et al.: Assessment of LiFePO4 battery performance in stand alone photovoltaic street light system ☆. *Procedia Eng.* **170**, 503–508 (2017)
5. Bellido-Outeiriño, F.J., Quiles-Latorre, F.J., Moreno-Moreno, C.D., et al.: Streetlight control system based on wireless communication over DALI protocol. *Sensors* **16**(5), 597 (2016)
6. Ramadhani, F., Bakar, K.A., Hussain, M.A., et al.: Optimization with traffic-based control for designing standalone streetlight system: a case study. *Renew. Energy* **105**, 149–159 (2017)
7. Ożadowicz, A., Grela, J.: Energy saving in the street lighting control system—a new approach based on the EN-15232 standard. *Energy Effi.* **10**(3), 1–14 (2017)
8. Tannous, S., Manneh, R., Harajli, H., et al.: Comparative cradle-to-grave life cycle assessment of traditional grid-connected and solar stand-alone street light systems: a case study for rural areas in Lebanon. *J. Clean. Prod.* (2018)
9. Shahzad, K., Čuček, L., Sagir, M., et al.: A case study for developing eco-efficient street lighting system in Saudi Arabia. *Chem. Eng. Trans.* **52**, 1141–1146 (2016)
10. Mahoor, M., Salmasi, F.R., Najafabadi, T.A.: A hierarchical smart street lighting system with brute-force energy optimization. *IEEE Sens. J.* **17**(9), 2871–2879 (2017)
11. Ding, Q., Sun, B., Zhang, X.: A traffic-light-aware routing protocol based on street connectivity for urban vehicular ad hoc networks. *IEEE Commun. Lett.* **20**(8), 1635–1638 (2016)
12. Ramakrishnareddy, C.K., Shunmugam, P., Vishwanathan, N.: Soft switched full-bridge light emitting diode driver configuration for street lighting application. *IET Power Electron.* **11**(1), 149–159 (2018)
13. Wu, L., Li, M., Pan, Z., et al.: Based on the internet of street lamp illumination by automatic control system. *Int. J. Smart Home* **9**(12), 57–64 (2015)
14. Osman, S.R., Rahim, N.A., Selvaraj, J., et al.: Single sensor charging system with MPPT capability for standalone streetlight applications. *J. Power Electron.* **15**(4), 929–938 (2015)



Control System Design of Reactor Robot for Object Salvaging Underwater

Xiaochen Huang^{1,2}, Lingyu Sun^{2(✉)}, Xiaojun Zhang², Manhong Li²,
and Minglu Zhang²

¹ School of Mechanical Engineering, Tianjin University of Technology and Education, Dagunan Road No. 1310, Hexi District, Tianjin 300222, China
hxc_tj@hotmail.com

² School of Mechanical Engineering, Hebei University of Technology, Xiping Road No. 5340, Beichen District, Tianjin 300401, China
sunlyu78@126.com

Abstract. A reactor robot for object salvaging underwater was designed to reduce the labor intensity of operators under the nuclear radiation environment. The control system of the robot, composed of control system hardware and control system software, was designed to be compact and radiation resistant. The control algorithm of the robot based on current, velocity, and position feedback of each motor makes the robot achieve the goal of salvaging exactly and rapidly. In order to verify the stability of the control system, the salvaging foreign matter experiment was tested in nuclear base. The result shows that the control system of the robot is stable enough.

Keywords: Reactor robot · Object salvaging underwater · Control system · Control algorithm

1 Introduction

As the development of nuclear power plant automation, the robot technology is widely used in this field. The robot can replace the operators to accomplish some tasks, avoid the hazards caused by long-term exposure of operators to the nuclear environment. Therefore, using robot to replace the operator has become the consensus [1–7].

In recent years, especially after the accident of Fukushima Nuclear Power Station, many robots have been applied in the nuclear plant. “Packbot 510” is a robot developed by American iRobot company. Fin-shaped support wheel is mounted on both sides of this robot’s underpan which uses caterpillar structure. The size of the support wheel is 889 mm in length, 521 mm in width, and 178 mm in height. There are two models of 4-DOF manipulator matching for the robot, and their maximum grasp ability is 13.6 kg and 6.8 kg, respectively. Different kinds of sensors can be mounted on the robot. This robot was the first batch of the rescue robot which entered in the accident scene of Fukushima Nuclear Power Station. It moved in the wreckage and returned the data of the nuclear accident site [8–10]. “Rosemary” is a robot which is developed based on the robot named “Quince” by Chiba Institute of Technology in Japan. It consists of a crawler mobile platform and a manipulator fixed on the mobile platform. All the key

components are replaced by the radiation-resistant ones on the basis of “Quince”. After 11 main components have been tested, the result of the irradiation experiment proved that the robot could work for at least 400 h in the radiation environment [11–16]. In addition, for different demands, the researchers in China, Germany, France, England, Canada, South Korea, and other countries have done majority researches and developed a series of specialized purposed robots for nuclear power stations [17–20].

During the normal operation of the nuclear power plant, underwater inspection is a routine task. Once foreign matter is found underwater, operators need to use salvage equipment for underwater salvaging operations. To solve this problem, a reactor robot for object salvaging underwater was designed to reduce the labor intensity of operators under the nuclear radiation environment.

2 The System Composition

According to the functional requirements, the robot is divided into robot mechanical body and robot control console. Figure 1 shows the system composition of the robot. The robot mechanical body, including wheeled mobile platform, 4-DOFs manipulator, and camera system, is used as an executive mechanical structure for underwater foreign matter salvage. The robot control console, including control instruction interactive system, data acquisition system, video capture system, and other hardware, “Robot real-time control system” software, is used as control core unit to remote control the robot. The control instructions are issued through the man–machine interaction interface, while the sensor information installed in the robot mechanical body is received and displayed in the robot control console.

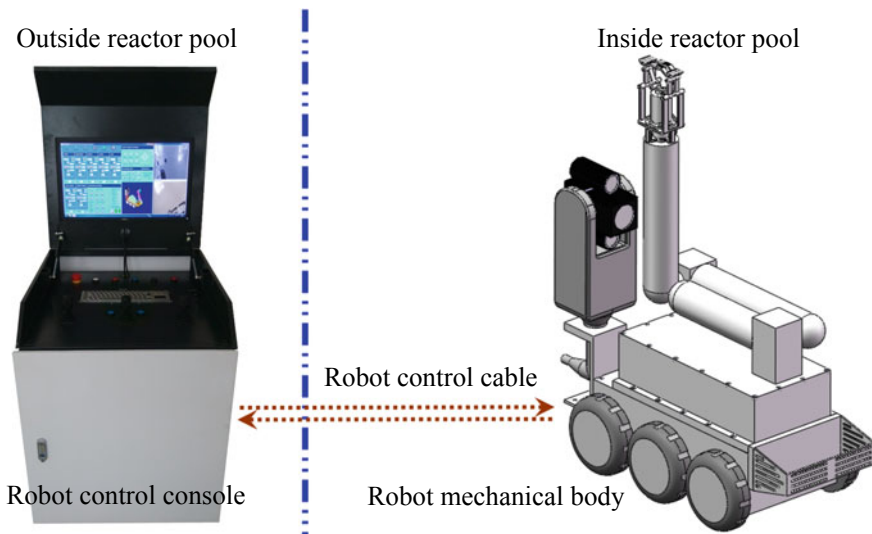


Fig. 1. The system composition of the robot

3 The Robot Control Console

Considering the robot running in the RX plant where the nuclear power plant reactor pressure vessel is placed, the robot control system is integrated into a control console as much as possible due to the high safety requirements of the system. Through the study of the working area, robot uses the mechanical body to realize the design target. And the control system ensures the safety of the operation, the accuracy of the location, and the speed of the grabbing. According to the requirements of the function, the robot chooses the reasonable hardware equipment and the reasonable control strategy to optimize the foreign matter salvaging operation.

3.1 The Hardware of the Control System

In order to provide sufficient operability, the Advantech industrial computer is used as the main control unit which provides adequate hardware interface to facilitate the integration of control instruction interactive system, data acquisition system, and video capture system. In the control instruction interactive system, each joint uses Copley driver and Maxon servomotor as the driving unit. In order to minimize the connected cable weight between robot mechanical body and control console, and provide reliable instruction interaction at the same time, each node and the main control unit uses CAN bus as the control protocol. In the data acquisition system, analog data acquisition card is chosen to collect analog information of three joysticks. Number one joystick controls the manipulator end executor of MRR along the X, Y, Z three-axis linear motion and wrist joint rotation movement in the Cartesian coordinate. Number two joystick controls the wheeled mobile platform of MRR front, rear movement, and pivot steering movement. Number three joystick controls the main camera PTZ horizontal, vertical rotation, and camera lens focal length zoom. In the video capture system, the main camera controller uses P150R camera controller of Ahlberg Electronics Corporation. The output signal port connects the MRAD-S color radiation-resistant camera through a radiation-resistant cable. The input signal port connects I150USB camera signal converter of Ahlberg Electronics Corporation via a dedicated cable. Operating personnel realized the control of the main camera through this series of hardware connections. Compared with the main camera, auxiliary camera has fewer functions. Only using a separate controller fixed to the dedicated mounting holes of control console connected N29F radiation-resistant camera via a dedicated cable, operating personnel realizes the control of the auxiliary camera. In addition, in order to realize the video display of the main camera and the auxiliary camera, the main camera controller and auxiliary camera controller are designed with a dedicated image signal output port which is connected to the video capture card. Figure 2 shows the structural composition of the control system hardware.

3.2 The Software of the Control System

To ensuring safe, controllable, and efficient control of the robot, the control system software of the robot is designed to provide reliable multimode motion control of the robot. The control command for the desired action is issued in the human-computer

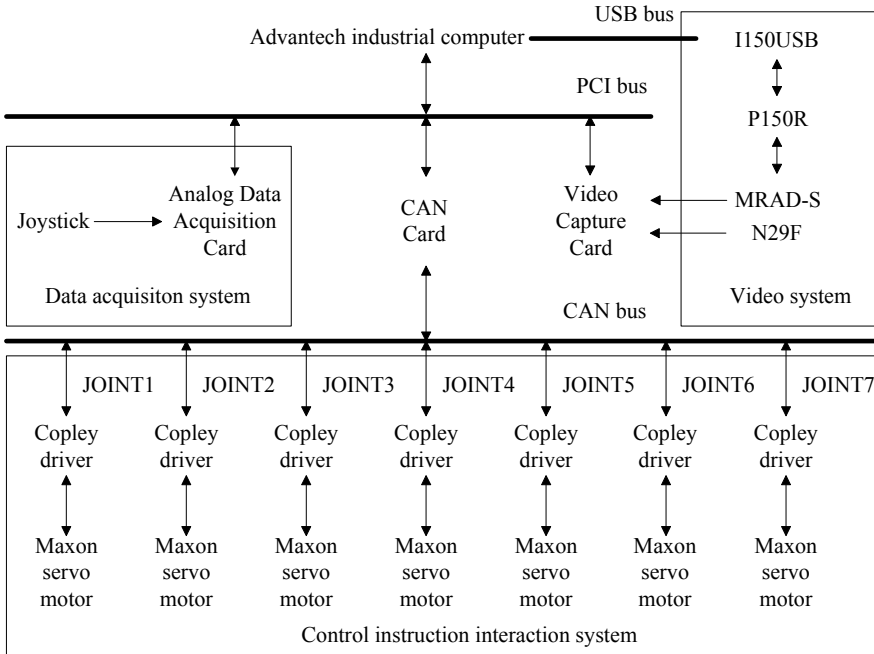


Fig. 2. The structural composition of the control system hardware

interaction interface through the mouse, keyboard, and joystick. Meanwhile, the human–computer interaction interface in real time displays the sensor information installed in the robot mechanical body. The response of the front panel of the control system software is operated through the man–machine interaction interface. According to the function, the control system software can be divided into seven modules, namely, three-dimensional model real-time simulation module, parameter configuration module, single-axis control module, multi-axis linkage module, teleoperation module, main camera display module, and auxiliary camera display module. Figure 3 shows the control interface of the software for robot’s remote control system.

4 The Control Algorithm of Salvaging Foreign Matter

The control algorithm of salvaging foreign matter based on current, velocity, and position feedback of each motors uses the combination mode of semiautomatic on a massive scale and manual operation on a small scale. This method can make the robot achieve the goal of salvaging exactly and rapidly. Specifically, it designs “pre-set setting zone” in the control system software. There are five pre-set setting zones which are “calibration location”, “pre-working location”, “grabbing location”, “place location”, and “zero location” in the human–computer interaction interface. In the process of operation, the fixed algorithm is written on those five pre-set setting zones, with



Fig. 3. The control interface of the software for robot's remote control system

semiautomatic being used in the operation. Then, the operator operates relevant position via one-click way which is a button in the human-computer interaction interface.

5 Experiment

To test the comprehensive performance of robot and verify the control algorithm of salvaging foreign matter, we were being the experiment of salvaging foreign matter in a nuclear base. The way of the experiment is that operator controls robot to search and salvaging foreign matter. First, when the robot is in the initial state, the operator uses the main camera pan-tilt to search environment and check the position of foreign matter. Then, the operator controls the robot to salvaging foreign matter and puts the foreign materials into the container. Finally, the operator makes the robot get back to the zero position.

In this experiment, the operator issued the order to the human-computer interaction interface by mouse, keyboard, or analog stick, and robot completed the test of salvaging foreign matter successfully. Hence, it can verify that the control system of the robot is stable enough.

6 Conclusion

In order to reduce the labor intensity of operators under the nuclear radiation environment, a reactor robot for object salvaging underwater was designed. The robot control system is integrated into a control console as much as possible due to the high safety requirements of the system. The control system of the robot, composed of control system hardware and control system software, was designed to be compact and

radiation resistant. The control algorithm of salvaging foreign matter based on current, velocity, and position feedback of each motor uses the combination mode of semi-automatic on a massive scale and manual operation on a small scale. The salvaging foreign matter experiment was tested in nuclear base and the result shows that the control system of the robot is stable.

Acknowledgements. This work was supported by Tianjin Enterprise Science and Technology Commissioner Project under grant 18JCTPJC66000, University Foundation of Tianjin University of Technology and Education under grant KJ1702 and Scientific Research Foundation of Tianjin University of Technology and Education under grant KYQD1807.

References

1. Ruimin, M., Jian, Z., Xueliang, Y.: China's approach to nuclear safety—from the perspective of policy and institutional system. *Energy Policy* **76**(1), 161–172 (2015)
2. Sophie, G., Staffan, J.S., Carl, H., et al.: New perspectives on nuclear power-generation IV nuclear energy systems to strengthen nuclear non-proliferation and support nuclear disarmament. *Energy Policy* **73**(10), 815–819 (2014)
3. Masataka, M., Randeep, S., Thang, N., et al.: Heat pipe based passive emergency core cooling system for safe shutdown of nuclear power reactor. *Appl. Therm. Eng.* **73**(1), 697–704 (2014)
4. Liu, Q., Wang, G., Dong, Y., et al.: A novel nuclear station inspection robot. In: 2014 4th IEEE International Conference on Information Science and Technology, 26–28 April 2014, Shenzhen, China (2004)
5. Robert, B.: Robots in the nuclear industry: a review of technologies and applications. *Ind. Robot* **38**(2), 113–118 (2011)
6. Richard, B.: How do you decommission a nuclear installation? Call in the robots. *Ind. Robot* **37**(2), 134–136 (2010)
7. Keiji, N., Seiga, K., Yoshito, O., et al.: Emergency response to the nuclear accident at the Fukushima Daiichi Nuclear Power Plants using mobile rescue robots. *J. Field Robot.* **30**(1), 44–63 (2013)
8. Park, J.-Y., Cho, B.-H., Lee, J.-K.: Trajectory-tracking control of underwater inspection robot for nuclear reactor internals using time delay control. *Nucl. Eng. Des.* **239**(11), 2543–2550 (2009)
9. Dexter, D., Brandon, S., Robin, M.: Run the robot backward. In: 2013 IEEE International Symposium on Safety, Security, and Rescue Robotics, 21–26 Oct 2013, Linköping, Sweden (2013)
10. Yamauchi, B.M.: PackBot: a versatile platform for military robotics. *Proc. SPIE (Unmanned Ground Vehicle Technology VI)* **5422**, 228–237 (2004)
11. Shinji, K., Mineo, F., Takashi, O.: Emergency response by robots to Fukushima-Daiichi accident: summary and lessons learned. *Ind. Robot* **39**(5), 428–435 (2012)
12. Gunn, J.E., Carr, M., Smee Stephen, A., et al.: Detectors and cryostat design for the SuMIRe Prime Focus Spectrograph (PFS). In: The International Society for Optics and Photonics, 9 Oct 2012, Amsterdam, Netherlands (2012)
13. Keiji, N., Seiga, K., Yoshito, O., et al.: Redesign of rescue mobile robot Quince—toward emergency response to the nuclear accident at Fukushima Daiichi Nuclear Power Station on March 2011. In: IEEE International Symposium on Safety, Security, and Rescue Robotics, Kyoto, Japan (2011)

14. Tomoaki, Y., Keiji, N., Satoshi, T., et al.: Improvements to the rescue robot quince toward future indoor surveillance missions in the Fukushima Daiichi nuclear power plant. In: Springer Tracts in Advanced Robotics, Matsushima, Japan (2014)
15. Keiji, N., Seiga, K., Yoshito, O., et al.: Gamma-ray irradiation test of electric components of rescue mobile robot Quince. In: IEEE International Symposium on Safety, Security, and Rescue Robotics, 1–5 Nov 2011, Kyoto, Japan (2011)
16. Eric, R., Kazunori, O., Tomoaki, Y., et al.: Integration of a sub-crawlers' autonomous control in Quince highly mobile rescue robot. In: IEEE/SICE International Symposium on System Integration, 21–22 Dec 2010, Sendai, Japan (2010)
17. Gang, W., Xi, C., Xiufen, Y.: Research on anti-rollover stability for crablike robot. In: 2013 IEEE International Conference on Mechatronics and Automation, 4–7 Aug 2013, Takamastu, Japan (2013)
18. Karydis, K., Poulakakis, I., Tanner, H.G.: A switching kinematic model for an octapedal robot. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, 7–12 Oct 2012, Vilamoura, Portugal (2012)
19. Andrews, H.L., Bakkali Taheri, F., Barros, J., et al.: Longitudinal profile monitors using coherent Smith-Purcell radiation. *Nucl. Instrum. Methods Phys. Res. A* **740**, 212–215 (2014)
20. Oka, K., Shibamura, K.: Development of a radiation-proof robot. *Adv. Robot.* **16**(6), 493–496 (2002)



Application of EWT and PSO-SVM in Fault Diagnosis of HV Circuit Breakers

Bing Li, Mingliang Liu^(✉), Zijian Guo, and Yamin Ji

HLJ Province Key Lab of Senior-Education for Electronic Engineering,
Heilongjiang University, Harbin 150080, China

{mll_0608, Jiyamin1225}@163.com, {1248572466,
345404027}@qq.com

Abstract. In order to improve the recognition rate of mechanical vibration signals of high voltage circuit breakers, a feasible new fault diagnosis method is proposed in this paper. Firstly, the empirical wavelet transform (EWT) is adopted to decompose the original multi-component signals into a series of intrinsic mode functions (IMF). Secondly, the envelop energy entropies of these IMF components are calculated as signal features. Finally, establishing the optimal support vector machine (SVM) classifier by particle swarm optimization (PSO) method. Using this EWT-PSO-SVM model to identify the unknown samples, the results show that the EWT method can effectively reduce modal aliasing problem, and the recognition rate of EWT-PSO-SVM model is higher than EMD-PSO-SVM model, these results verify the feasibility and superiority of the proposed EWT-PSO-SVM fault diagnosis method.

Keywords: Empirical wavelet transform · Particle swarm optimization · Support vector machine · Fault diagnosis

1 Introduction

With the increasing demand for electricity, People's requirements on reliability of High voltage circuit breakers (HVCBs) are getting higher and higher. Therefore, the operation state of HVCBs is highly valued by engineers and technicians, and the researches on fault diagnosis of HVCBs have been carried out in succession.

The vibration signals produced by HVCBs is a transient non-stationary time series, and usually, these signals are analyzed by time-frequency analysis method. Empirical mode decomposition (EMD) [1] is a widely used time-frequency analysis method in the field of mechanical fault diagnosis, it can decompose the signals into a series of intrinsic modal functions (IMF) with the frequencies from high to low in order. For example, EMD was used for fault diagnosis of HV Circuit Breaker by Huang et al. [2] in 2011. But with the in-depth studies of EMD, it found that the EMD method has serious mode aliasing problem [3]. In order to solve this problem, the EWT method was proposed [4] in 2013. This method has been widely used in recent years [5], and some achievements have been obtained. Just because the EWT method can effectively reduce the modal aliasing problem. Therefore, in this paper, The EWT method is used

for signal decomposition. Finally, support vector machine (SVM) is used for vibration signal recognition in this paper, while the parameters in SVM are optimized by particle swarm optimization (PSO) method.

2 EWT and Feature Extraction Method

2.1 EWT

Gilles combined the theoretical framework of wavelet transform and the self-adaptation of EMD, a new signal processing method EWT was proposed in 2013 [6]. Firstly, the frequency axis of the original signal Fourier spectrum is normalized to $[0, \pi]$. The normalized frequency axis of Fourier spectrum is then adaptively divided into N sub-bands, $N + 1$ boundaries are needed. Let w_n to be the boundaries of the adjacent interval, and w_n are chosen from the midpoint of two adjacent maxima of Fourier spectrum, $n = 0, 1, 2, \dots, N$, and define $w_0 = 0, w_N = \pi$. Then each sub interval can be represented as: $A_\lambda = [w_{\lambda-1}, w_\lambda], \lambda = 1, 2, \dots, N$, that is $[0, \pi] = \bigcap_{\lambda=1}^N A_\lambda$. The detailed segmentation is shown in Fig. 1. The shadow parts are defined as the transition section T_n with w_n as centre and $2\tau_n$ as width. Where $\tau_n = \gamma w_n, \gamma$ is coefficient and $0 < \gamma < 1$.

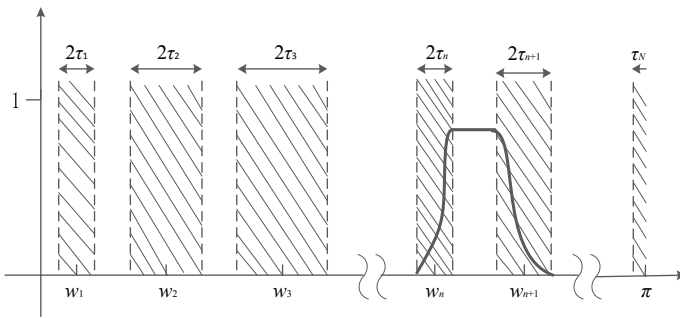


Fig. 1. Sensor network with sensors, mobile sink and data routing

After determining the division intervals A_λ , the empirical wavelet is defined as a band-pass filter on each A_λ interval, According to the construction method of Meyer wavelet, the empirical scaling function $\hat{\phi}_n(w)$ and the empirical wavelet function $\hat{\psi}_n(w)$ are defined as:

$$\hat{\phi}_n(w) = \begin{cases} 1, & |w| \leq w_n - \tau_n \\ \cos \left[\frac{\pi}{2} \beta \left(\frac{1}{2\tau_n} (|w| - w_n + \tau_n) \right) \right], & w_n - \tau_n \leq |w| \leq w_n + \tau_n \\ 0, & \text{other} \end{cases} \quad (1)$$

$$\hat{\psi}_n(w) = \begin{cases} 1, & w_n + \tau_n \leq |w| \leq w_{n+1} - \tau_{n+1} \\ \cos \left[\frac{\pi}{2} \beta \left(\frac{1}{2\tau_{n+1}} (|w| - w_{n+1} + \tau_{n+1}) \right) \right], & w_{n+1} - \tau_{n+1} \leq |w| \leq w_{n+1} + \tau_{n+1} \\ \sin \left[\frac{\pi}{2} \beta \left(\frac{1}{2\tau_n} (|w| - w_n + \tau_n) \right) \right], & w_n - \tau_n \leq |w| \leq w_n + \tau_n \\ 0, & \text{other} \end{cases} \quad (2)$$

where $\beta(x)=x^4(35 - 84x + 70x^2 - 20x^3)$, $\tau_n = \gamma w_n$. According to the construction method of classical wavelet transform, we can construct empirical wavelet transform. The detail coefficients are defined as the inner product of signal $f(t)$ and empirical wavelet function $\psi_n(t)$:

$$w_f^e(n, t) = \langle f(t), \psi_n(t) \rangle = \int f(\tau) \overline{\psi_n(\tau - t)} d\tau = \mathcal{F}^{-1} \left[f(w) \hat{\psi}_n(w) \right] \quad (3)$$

The approximate coefficient is defined as the inner product of signal $f(t)$ and empirical scaling function $\phi_1(t)$:

$$w_f^e(0, t) = \langle f(t), \phi_1(t) \rangle = \int f(\tau) \overline{\phi_1(\tau - t)} d\tau = \mathcal{F}^{-1} \left[f(w) \hat{\phi}_1(w) \right] \quad (4)$$

where the Fourier transforms of $\psi_n(t)$ and $\phi_1(t)$ are $\hat{\psi}_n(w)$ and $\hat{\phi}_1(w)$ respectively. The complex conjugates of $\overline{\psi_n(\tau - t)}$ and $\overline{\phi_1(\tau - t)}$ are $\psi_n(\tau - t)$ and $\phi_1(\tau - t)$ respectively. Signals reconstruct from detail coefficients and approximate coefficients, give the result as follows:

$$\begin{aligned} f(t) &= w_f^e(0, t) * \phi_1(w) + \sum_{n=1}^N w_f^e(n, t) * \psi_n(t) \\ &= \mathcal{F}^{-1} \left[\widehat{w}_f^e(0, w) \widehat{\phi}_1(w) + \sum_{n=1}^N \widehat{w}_f^e(n, w) \widehat{\psi}_n(w) \right] \end{aligned} \quad (5)$$

where $*$ represents convolution, the Fourier transforms of $w_f^e(0, t)$ and $w_f^e(n, t)$ are $\widehat{w}_f^e(0, w)$ and $\widehat{w}_f^e(n, w)$ respectively, according to the mathematical form of reconstructed signal, empirical mode f_k ($k = 0, 1, 2, \dots, N - 1$) is defined as:

$$f_k(t) = \begin{cases} w_f^e(0, t) * \phi_1(t), & k = 0 \\ w_f^e(k, t) * \psi_k(t), & k = 1, 2, \dots, N - 1 \end{cases} \quad (6)$$

2.2 Feature Extraction Method

Hilbert transform is often used to extract the envelope of signals in mechanical fault diagnosis field. The Hilbert transform of signal $x(t)$ is defined as:

$$\hat{x}(t) = \frac{1}{\pi t} * x(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(\tau)}{t - \tau} d\tau \tag{7}$$

while the analytical signal of signal $x(t)$ is defined as:

$$g(t) = x(t) + j\hat{x}(t) \tag{8}$$

and the amplitude of $g(t)$ is: $A(t) = \sqrt{x(t)^2 + \hat{x}(t)^2}$, here, $A(t)$ is envelope of signal $x(t)$. According to information entropy theory, the characteristic entropies of signals can be calculated as follows: let's the envelope signal divided into N segments along time axis, then the segmented energy can calculate by following formula:

$$Q(i) = \int_{t_i}^{t_i + \frac{T}{N}} |A(t)|^2 dt \tag{10}$$

where $i = 1, 2, 3, \dots, N$; T is total time of envelope, t_i and $t_i + \frac{T}{N}$ are start and end time of the i -th segment respectively. While in this paper, $Q(i)$ are normalized as follows:

$$\varepsilon(i) = Q(i) / \sum_{i=1}^N Q(i) \tag{11}$$

and the characteristic entropy of the signal $x(t)$ is defined as:

$$H_j = - \sum_{i=1}^N \varepsilon(i) \lg \varepsilon(i) \tag{12}$$

In summary, the steps of feature extraction are as follows:

1. EWT decomposition, According to Eq. 6, the IMF components $f_k(t)$ are obtained;
2. Calculate the characteristic entropy of each IMF component by Formula 12, and in this paper, each IMF component is divided into 16 segments along the time axis;
3. Consist eigenvectors: $T = [H_1, H_2, \dots, H_N]$.

3 SVM and PSO

3.1 SVM Classification Principle

For non-linear separability samples, the optimal classification hyperplane can be calculated by solving the optimal problem with the following constraints:

$$\begin{cases} \min \Phi(\mathbf{w}, \varepsilon) = \frac{\|\mathbf{w}\|^2}{2} + C \sum_{i=1}^l \varepsilon_i \\ \text{s.t. } y_i[\mathbf{w} \cdot \mathbf{x}_i + \mathbf{b}] \geq 1 - \varepsilon_i, \\ (i = 1, 2, 3, \dots, l) \end{cases} \quad (13)$$

where ε_i ($\varepsilon_i \geq 0$) is relaxation variable, $\Phi(\mathbf{w}, \varepsilon)$ is objective function, and C is penalty coefficient. And the optimal hyperplane means that the hyperplane can not only divide the two types of samples correctly, but also make the classification interval $2/\|\mathbf{w}\|$ of the two categories reach the maximum, as shown in Fig. 2.

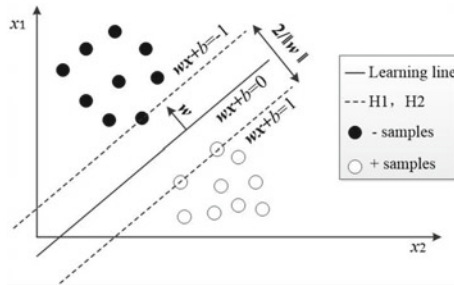


Fig. 2. Optimal classification hyperplane

For non-linear SVM, the proper inner product function $K = (x_i, x_j)$ is used to transform the nonlinear problem into the above linear problem, and the most commonly used kernel function is RBF kernel function, and it defined as:

$$K(x_i, x) = \exp\left(-\frac{\|x_i - x\|^2}{2g^2}\right) \quad (14)$$

where g is kernel parameter, the corresponding classification decision function is:

$$f(x) = \text{sgn}\left(\sum_{j=1}^l a_j \cdot y_j K(x_i, x) + b\right) \quad (15)$$

The kernel parameter g and penalty coefficient C jointly affect the performance of SVM, In this study, PSO optimization algorithm is used to select parameters C and g .

3.2 PSO

PSO is a widely used model parameter optimization algorithm, it has a strong global search capability. Assume that in a D dimensional search space, the population $Z = (Z_1, Z_2, Z_3, \dots, Z_s)$ consists of S particles (each solution of the optimization problem is called a particle), among them, the i th particle represents a D dimensional

vector $\mathbf{Z}_i = (Z_{i1}, Z_{i2}, \dots, Z_{iD})^T$, \mathbf{Z}_i is the position of i th particle in D dimensional space, according to the objective function, the fitness value corresponding to each particle position Z_i can be calculated, velocity of particle i is $\mathbf{V}_i = (V_{i1}, V_{i2}, \dots, V_{iD})^T$, individual extreme value of population is $\mathbf{P}_i = (P_{i1}, P_{i2}, \dots, P_{iD})^T$, while global extremum is $\mathbf{P}_g = (P_{g1}, P_{g2}, \dots, P_{gD})^T$. The position and velocity of a particle is updated by the following formula:

$$\mathbf{V}_{id}^{k+1} = w\mathbf{V}_{id}^k + c_1r_1(\mathbf{P}_{id}^k - \mathbf{Z}_{id}^k) + c_2r_2(\mathbf{P}_{gd}^k - \mathbf{Z}_{id}^k) \quad (16)$$

$$\mathbf{Z}_{id}^{k+1} = \mathbf{Z}_{id}^k + \alpha\mathbf{V}_{id}^{k+1} \quad (17)$$

where k is iteration times, w is inertia weight, $d = 1, 2, \dots, D$; $i = 1, 2, \dots, s$; c_1 and c_2 are learning factor, and they are greater than 0; r_1 and $r_2 \in [0,1]$ is random number; α is a constraint factor to control the speed weight.

4 Results and Discussion

4.1 Test Environment and Pre-treatment

In this study, the type of HVCB is ZW32-12G/630-20. Three kinds of vibration signals are collected, they are normal state (Normal), pedestal screw loose state (Fault I), time delay state caused by insufficient lubrication (Fault II). Each group collects 30 sets of data, 20 sets of data are used to train SVM, and the other 10 sets of data are used to test SVM classifier model. According to the recognition rate of SVM classifier model, it can verify the superiority of the proposed fault diagnosis method.

The data acquisition platform is built by piezoelectric acceleration sensor and NI USB6002 data acquisition card, the sampling frequency f_s is set to 40 kS/s, and the sampling point n is set to 8000, using the EXCEL table to store the sample points. Figure 3a is a normal state time domain waveform which is collected by the acquisition platform in 0.2 s, it has not been de-noised yet. Taking into account the smoothness of

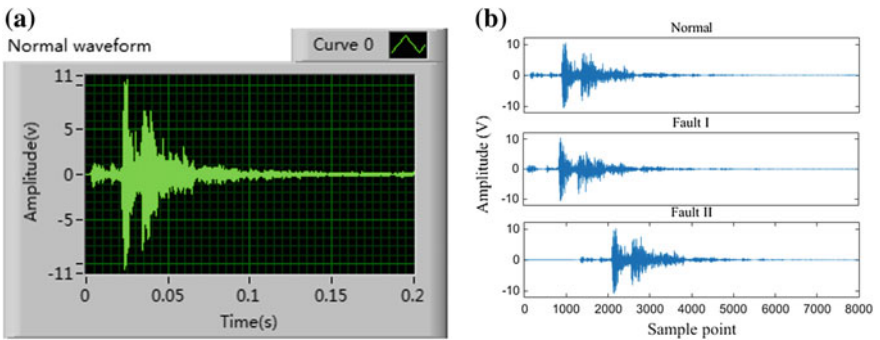


Fig. 3. **a** Time domain waveform of normal state ($f_s = 40\text{KS/s}$); **b** Three kinds of signals after de-noising

the de-noised signal, and try to make the signal to noise ratio as large as possible, in this paper, using wavelet soft thresholding method to reduce noise of each signal, and the value of threshold thr is 0.2034 V, Fig. 3b is the de-noised signals of three types of original vibration signals.

4.2 EWT Simulation of Measured Signals

The vibration signal is analyzed by EWT toolbox. EWT method decomposes signals into a series of IMF components according to the frequency component of the signals, that's to say, we can get the sub events $f_k(t)$ of the signals. In order to compare it with the EMD method, take the normal signal as an example, as shown in Fig. 4, the same normal state signal is decomposed by EWT and EMD respectively, when the number of segments of EWT is 5, it can be found that under the same condition, the number of IMF components decomposed by EMD are 14, it is more than those decomposed by EWT. However, a large number of studies have shown that the number of main IMF components of HVCB is not so much [5, 7, 8], this phenomenon shows that the IMF components decomposed by the EMD method have spurious modes, that is, the same mode may contain two or more main frequency components, or the two modes share the same dominant frequency component simultaneously. In terms of instantaneous non-stationary signals generated by the HVCBs, it shows that the EMD method is not suitable for the decomposition of HVCBs mechanical vibration signals.

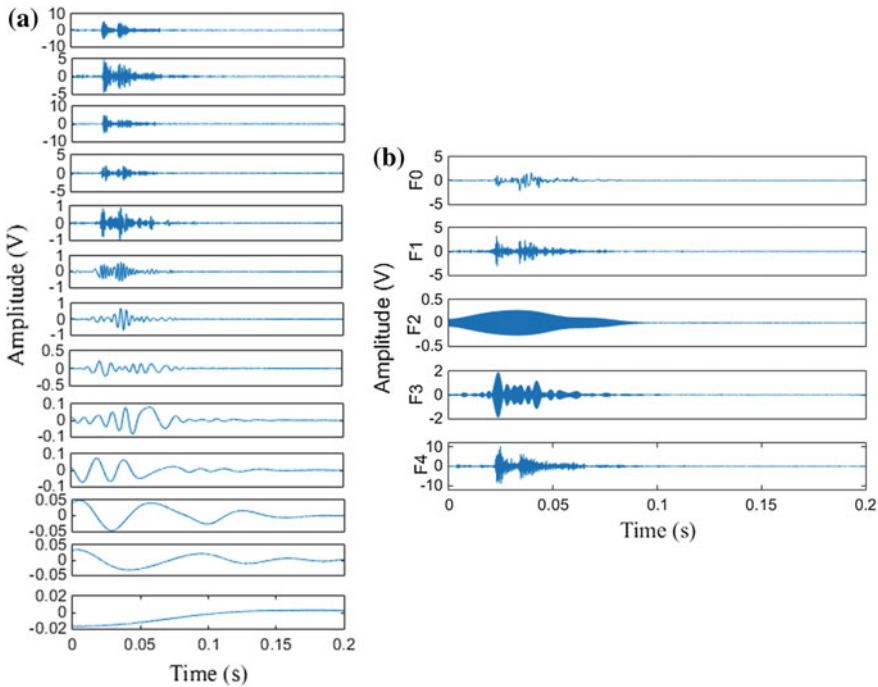


Fig. 4. Comparison of IMFs of EWT and EMD methods **a** EMD decomposition result **b** EWT decomposition result

It can be seen from Fig. 4a that the energies of IMF components are mainly concentrated in the first 6 orders IMFs. That is to say, the information of the original signal is mainly reflected in the first 6 order IMFs, therefore, in this paper, these IMF components are chosen to extract the characteristics of the signals. Signals with different mechanical conditions are different in energy distribution of each IMF component, however, the energy distribution of each IMF component fluctuates little in the same mechanical condition. According to Eq. 12, firstly, each IMF component is divided into 16 segments along the time axis, and Tables 1 and 2 list a set of entropies by using EWT-information entropy and EMD-information entropy methods respectively, here, $T = [H_1, H_2, H_3, H_4, H_5]$ as the eigenvector obtained by using the EWT-information entropy method. While $T = [H_1, H_2, H_3, H_4, H_5, H_6]$ as the eigenvector obtained by using the EMD-information entropy method. It can be seen from the table that the features of signals in different states are very different, especially Table 1.

Table 1. Entropies obtained by using EWT-information entropy

		H_1	H_2	H_3	H_4	H_5
EWT	Normal	1.0778	1.1278	1.0089	2.0679	2.7175
	Fault I	1.3695	1.5496	1.9139	2.1117	2.4553
	Fault II	1.3005	1.2534	1.0324	1.0078	0.4059

Table 2. Entropies obtained by using EMD-information entropy

		H_1	H_2	H_3	H_4	H_5	H_6
EMD	Normal	1.2886	1.2563	2.1249	1.1447	2.0987	1.0303
	Fault I	1.3240	1.1450	2.2204	1.1894	2.0473	1.0426
	Fault II	1.3514	1.3397	1.9775	1.0592	2.0297	0.8116

4.3 Analysis of Recognition Results

In order to compare the pros and cons of EWT and EMD methods, we input the eigenvectors obtained by the EWT and the EMD methods into SVM (which has been trained) respectively, we can determine the quality of the methods by comparing the recognition rate. Use test set data for PSO parameter optimization, the initial parameters of PSO are as follows: $c_1 = 1.5$, $c_2 = 1.7$, termination generation is 200, initial population size is 20. As shown in Figs. 5a and 5b, for EWT-information entropy features, the best parameter c and the best parameter g of SVM classifier are 0.1 and 0.38582 respectively, and the recognition rate of optimal SVM classifier is 100%. While Figs. 6a and 6b are the results of the EMD-information entropy features, a group of Fault I sample is mistakenly recognized as Normal state, the classification results of each state (each state has 10 groups of test data) are shown in Table 3. Compared with EMD-information entropy method, the features of signals obtained by the EWT-information entropy method are more significant, it because that the EWT method can effectively decompose the signals into a series of real vibration events, while EMD method could not.

Table 3. Recognition rate comparison

	Normal (%)	Fault I (%)	Fault II (%)
EWT-PSO-SVM	100	100	100
EMD-PSO-SVM	90	100	100

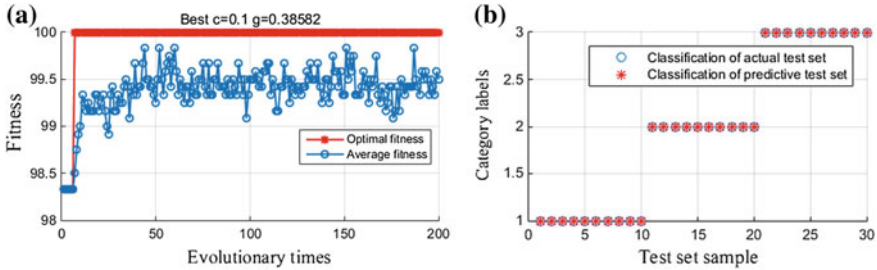


Fig. 5. **a** Parameter optimization of PSO algorithm (EWT), **b** recognition results (EWT)

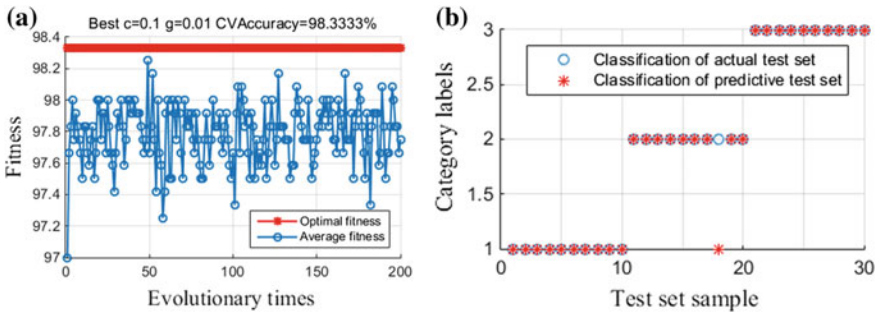


Fig. 6. **a** Parameter optimization of PSO algorithm (EMD), **b** recognition results (EMD)

5 Conclusions

Both the EWT-PSO-SVM and EMD-PSO-SVM methods can achieve a certain effect in HVCBs fault diagnosis. Under the same condition, the proposed EWT-PSO-SVM method is better than the EMD-PSO-SVM method in recognition rate. The quality of signal decomposition directly affects the accuracy of fault diagnosis, choosing an effective signal decomposition method is very important, the EWT method can effectively reduce the modal aliasing shortcomings of EMD method, and can effectively decompose vibration signals into real sub events, and then the obtained characteristic parameters are obvious. The EWT method is a new signal decomposition method, and it was proposed in 2013, it has been widely used in recent years. In this study, by comparing with EMD method, experimental results verify the feasibility and superiority of the EWT method in fault diagnosis filed.

Acknowledgement. This work is supported by natural science foundation of Heilongjiang province of China (No. E201233), science and technology innovation research team in higher educational institutions of Heilongjiang province (No. 2012TD007), Graduate Innovation Research Project of Heilongjiang University (No. YJSCX2018-052HLJU).

References

1. Huang, N.E., Shen, Z., Long, S.R., et al.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. A* **454**(1971), 903–995 (1998)
2. Huang, J., Hu, X.G., Gong, Y.N.: Machinery fault diagnosis of high voltage circuit breaker based on empirical mode decomposition. *Proc. CSEE* **31**(12), 108–113 (2011)
3. Hu, A.J., Sun, J.J., Xiang, L.: Mode mixing in empirical mode decomposition. *J. Vib. Meas. Diagn.* **31**(4), 429–434 (2011)
4. Wu, Z.H., Huang, N.E.: Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **1**(1), 1–41 (2009)
5. Li, B., Liu, M.L., Guo, Z.J., Ji, Y.M.: Mechanical fault diagnosis of high voltage circuit breakers utilizing EWT-improved time frequency entropy and optimal GRNN classifier. *Entropy* **20**(6), 448 (2018)
6. Gilles, J.: Empirical wavelet transform. *IEEE Trans. Signal Process.* **61**(16), 3999–4010 (2013)
7. Chang, G., Wang, Y., Wang, W.: Mechanical fault diagnosis of high voltage circuit breakers utilizing zero-phase filter time-frequency entropy of vibration signal. *Proc. CSEE* **33**(3), 155–162 (2013)
8. Huang, N.T., Zhang, S.X., Cai, G.W., et al.: Mechanical fault diagnosis of high voltage circuit breakers utilizing empirical wavelet transform and one-class support vector machine. *Chin. J. Sci. Instrum.* **36**(12), 2773–2781 (2015)



A Dynamic Programming Track-Before-Detect Algorithm with Adaptive State Transition Set

Hao Xing^(✉), Jidong Suo, and Xiaoming Liu

Dalian Maritime University, Dalian, China
sjddmu@dlmu.edu.cn

Abstract. Due to the use of a fixed-size state transition set, the traditional dynamic programming Track-Before-Detect (DP-TBD) algorithm significantly reduces the detection and tracking performance of maneuvering targets. This paper proposes a DP-TBD algorithm with an adaptive state transition set (ASTS-DP-TBD). The algorithm improves the search efficiency of the maneuvering target by introducing Kalman filtering and target state transition probability in the traditional algorithm. In addition, this paper also optimizes the termination decision strategy of the algorithm, which significantly improves the detection performance. Simulation results show that the proposed algorithm in this paper has better detection and tracking results than traditional algorithms for maneuvering targets.

Keywords: Track-Before-Detect (TBD) · Dynamic programming (DP) · State transition set · State transition probability

1 Introduction

TBD technology is an effective method for detecting and tracking weak targets in a low SNR environment [1]. TBD is a multi-frame signal accumulation technique. Compared with the traditional detection method, TBD does not detect the target by setting a threshold for each frame. Instead, after accumulating multiple frames of data, the target trajectory is given at the same time when the detection result is obtained.

The principle of DP-TBD algorithm is clear and its performance is excellent. It is a research hotspot in recent years. The basic idea of DP-TBD algorithm is to convert the target detection from a multistage decision-making process to multiple single-stage problems. Through each stage, the merit function is optimized to obtain a global optimal solution. The DP-TBD algorithm was originally used for optical image processing and this is the first application of DP in TBD [2]. The DP-TBD algorithm is divided into probability density accumulation and energy accumulation [3, 4]. The first one is suitable for maneuvering targets, but need to know clutter prior distribution (CPD) information. The second one does not require CPD information, and constructs the stage function directly with the target amplitude or energy, but it is only applicable to weak maneuvering targets with approximate trajectories of the motion trajectories. From then on, mountains of work have been done to improve the performance of DP-TBD [5–8].

This paper proposes an improved DP-TBD algorithm that can effectively achieve the detection and tracking of maneuvering targets. The improved algorithm introduces the acceleration component into the target state vector, which can optimize the prediction of the position and velocity in the Kalman filtering, so that the state transition set can be timely adjusted according to the change of the target speed. In addition, the introduction of state transition probability enables the target to more accurately accumulate energy along the true trajectory. At last, this paper optimizes the termination decision strategy by improving the setting method of the threshold. And the detection performance of maneuvering target is obviously improved.

2 Traditional Energy Accumulation Algorithm Model

2.1 System Model

In this paper, we assume the target is moving radially relative to the radar, and the radar obtains a $M \times N$ size measurement sequence for each full scan, which is called a frame and observes a total of K frames. The radar scanning interval is T and the measurement data at frame k can be expressed as

$$Z_k = \{z_k(x, y) | x = 1, 2, \dots, M; y = 1, 2, \dots, N; 1 \leq k \leq K\} \tag{1}$$

Where $z_k(x, y)$ is the measurement recorded in cell (x, y) , which is given by

$$z_k(x, y) = \begin{cases} A_k + n_k(x, y), & \text{no target in cell } (x, y) \text{ at frame } k \\ n_k(x, y), & \text{target in cell } (x, y) \text{ at frame } k \end{cases} \tag{2}$$

where A_k is the target amplitude, n_k is the measurement noise which obeys normal distribution.

$X_k = [s_x(k), v_x(k), s_y(k), v_y(k)]$ represents the position and velocity of the target in the x -direction and y -direction at frame k . The target trajectory sequence from the first frame to the K -th frame is given by

$$X(k) = \{X_1, X_2, \dots, X_K\} \tag{3}$$

$Z(k) = \{Z_1, Z_2, \dots, Z_K\}$ is a set of measurement sequences obtained by the target trajectory. $\hat{X}(k)$ is the target estimation sequence. It is hoped that $\hat{X}(k)$ is most likely to come from a trajectory of a real target.

2.2 Traditional DP-TBD Algorithm

(1) Initialization: For all states X_1

$$\begin{cases} I_1(x, y) = |z_1(x, y)| \\ \Psi_1(x, y) = 0 \end{cases} \tag{4}$$

where $I(\cdot)$ denotes merit value function, $\psi(\cdot)$ stores states transition records.

(2) Recursion: For $2 \leq k \leq K$, for all states X_k

$$\begin{cases} I_k(x, y) = \max_{(x^*, y^*) \in J_k(x, y)} \{|z_k(x, y)| + I_{k-1}(x^*, y^*)\} \\ \Psi_k(x, y) = \arg \left\{ \max_{(x^*, y^*) \in J_k(x, y)} (I_{k-1}(x^*, y^*)) \right\} \end{cases} \quad (5)$$

where $J_k(x, y)$ denotes state transition set, which is a set of all possible positions of the target from frame $k - 1$ to frame k . Determined by the target's maximum speed v_x^{\max} and v_y^{\max} , as follows:

$$J_k(x, y) = \{(a, b) | x - v_x^{\max} \leq a \leq x + v_x^{\max}, y - v_y^{\max} \leq b \leq y + v_y^{\max}\} \quad (6)$$

(3) Judgment and termination: For threshold V_T , find

$$\{\hat{X}_K\} = \{X_K : I_K(x, y) > V_T\} \quad (7)$$

(4) Backtracking: for $k = K - 1, K - 2, \dots, 1$

$$\hat{X}_k = \Psi_{k+1}(x, y) \quad (8)$$

Traditional DP-TBD algorithm is only a multi-frame accumulation of measured data, ignoring the state-characteristic relationship between measurement data frames. If the size of the state transition set is not suitable for the algorithm, it will reduce the detection and tracking of the target, especially the radar maneuvering target. Therefore, the selection of the state transition set is very important for the performance of the algorithm.

3 The ASTS-DP-TBD Algorithm

When facing with maneuvering targets, the traditional DP-TBD with constant size of transition set can hardly detect the targets. An improved DP-TBD algorithm is proposed in this paper. In the DP, there exists an energy accumulation path for each present resolution cell (i, j) at frame k , so we can take advantage of positions of the plots included in the path to estimate the target state vector. According to the estimated state vector, we can predict the target state vector at frame $k + 1$ by the one-step state prediction in the Kalman filtering.

On the one hand, the acceleration component is introduced into the target state vector. This can optimize the prediction of the position and velocity in the Kalman filtering, so that the state transition set can be changed according to the target speed.

On the other hand, in the process of target transfer, it is interfered by various noises, which will reduce the probability of energy accumulation of the target on the real trajectory. In this paper, the state transition probability is introduced in the energy accumulation strategy. It can make the target more accurately accumulate energy along the true trajectory and reduce the interference of process noise. Based on these, this paper also optimizes the termination decision strategy of the algorithm by improving the setting method of the threshold. The new threshold can be better adapted, and the detection performance of maneuvering target is obviously improved.

(1) The introduction of acceleration component in target state vector

The improved target state vector $x(k)$ and measurement state vector $z(k)$ are as follows:

$$x(k) = [s_x(k), v_x(k), a_x(k), s_y(k), v_y(k), a_y(k)] \tag{9}$$

$$z(k) = [\bar{s}_x(k), \bar{v}_x(k), \bar{a}_x(k), \bar{s}_y(k), \bar{v}_y(k), \bar{a}_y(k)] \tag{10}$$

where $s_x(k), v_x(k), a_x(k)$ is the position, velocity, acceleration on direction x respectively, and $s_y(k), v_y(k), a_y(k)$ is the counterparts on direction y.

In fact, we only have the measurements of target position, as to the measurements of velocity and acceleration are given by

$$[\bar{v}_x(k), \bar{v}_y(k)]^T = [\bar{s}_x(k), \bar{s}_y(k)] - [\hat{s}_x(k-1), \hat{s}_y(k-1)], \quad k \geq 2 \tag{11}$$

$$[\bar{a}_x(k), \bar{a}_y(k)] = [\tilde{a}_x(k), \tilde{a}_y(k)], \quad k \geq 2 \tag{12}$$

By introducing the acceleration component in the target state vector, the state transition set can be adjusted timely according to the change of the target position and speed. The state transition set is given by

$$J_k(x, y) = \{(a, b) | x - \tilde{v}_x(k) \leq a \leq x + \tilde{v}_x(k), y - \tilde{v}_y(k) \leq b \leq y + \tilde{v}_y(k)\} \tag{13}$$

(2) The introduction of state transition probability

In order to make the target more accurately accumulate energy along the true trajectory, this paper introduces the state transition probability in the energy accumulation strategy. In the Kalman filtering process, the probability of the transition from $x_{k-1}(x', y')$ to $x_k(x, y)$ can be calculated as

$$D_{k-1}(x', y') = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{d^2}{2\sigma^2}) \quad k \geq 2 \tag{14}$$

where $d = \sqrt{(x'' - x)^2 + (y'' - y)^2}$ denotes the distance between (x'', y'') and (x, y) . And (x'', y'') is the predicted value of (x, y) at frame k.

(3) ASTS-DP-TBD algorithm steps

Step 1: Initialization. For all states X_1

$$\begin{cases} I_1(x, y) = |z_1(x, y)| \\ \Psi_1(x, y) = [0, 0]^T \end{cases} \tag{15}$$

$$\hat{x}(1) = [\bar{s}_x(1), \bar{v}_x(1), 0, \bar{s}_y(1), \bar{v}_y(1), 0] \tag{16}$$

$$\hat{P}_1 = [\sigma_w, \sigma_w/T, \sigma_w/T^2, \sigma_w, \sigma_w/T, \sigma_w/T^2]^* \tag{17}$$

$$[\sigma_w, \sigma_w/T, \sigma_w/T^2, \sigma_w, \sigma_w/T, \sigma_w/T^2]$$

where P is predicted state error covariance matrix.

Step 2: State prediction.

$$\tilde{x}(k) = F\hat{x}(k - 1), \quad k \geq 2 \tag{18}$$

where F is state transition matrix and is given by

$$F = \begin{bmatrix} 1 & T & T^2/2 & 0 & 0 & 0 \\ 0 & 1 & T & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & T & T^2/2 \\ 0 & 0 & 0 & 0 & 1 & T \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{19}$$

Step 3: Recursion. For $2 \leq k \leq K$, for all states X_k

$$\begin{cases} I_k(x, y) = \max_{(x^*, y^*) \in J_k(x, y)} \{|z_k(x, y)| + D_{k-1}(x^*, y^*) + I_{k-1}(x^*, y^*)\} \\ \Psi_k(x, y) = \arg \left\{ \max_{(x^*, y^*) \in J_k(x, y)} [I_{k-1}(x^*, y^*)] \right\} \end{cases} \tag{20}$$

where $J_k(x, y)$ is state transition probability, which is determined by (13). And $D_{k-1}(x^*, y^*)$ is determined by (14).

Step 4: Calculation of predicted state error covariance matrix

$$\tilde{P}_k = F\hat{P}_{k-1}F^T + C_w, \quad k \geq 2 \tag{21}$$

$$G_k = \tilde{P}_k(\tilde{P}_k + C_v)^{-1}, \quad k \geq 2 \tag{22}$$

$$\hat{P}_k = (I - G_k)\tilde{P}_k, \quad k \geq 2 \tag{23}$$

where G_k is filtering gain matrix. C_w, C_v is process noise covariance matrix and measurement noise covariance matrix respectively.

Step 5: State estimation.

$$\hat{x}(k) = \tilde{x}(k) + G_k[z(k) - \tilde{x}(k)] \quad k \geq 2 \quad (24)$$

Step 6: Judgment and termination: For threshold V_T , find

$$\{\hat{X}_K\} = \{X_K : I_K(x, y) > V_T\} \quad (25)$$

$$V_T = -b_n \cdot \ln[-\ln(1 - p_d)] + a_n \quad (26)$$

$$a_n = \mu + \sigma[(2 \lg n)^{1/2} - \frac{1(\lg(\lg n) + \lg(4\pi))}{(2 \lg n)^{1/2}}] \quad (27)$$

$$b_n = \frac{(2 \lg n)^{1/2}}{\sigma}, \quad n = M \times N \times num \quad (28)$$

where p_d is the detect probability, M and N are the number of distance units in the x and y directions; num is the number of distance units the target may span between two frames; μ and σ are the mean and variance of the merit function obtained by accumulating k frames of the target trajectory.

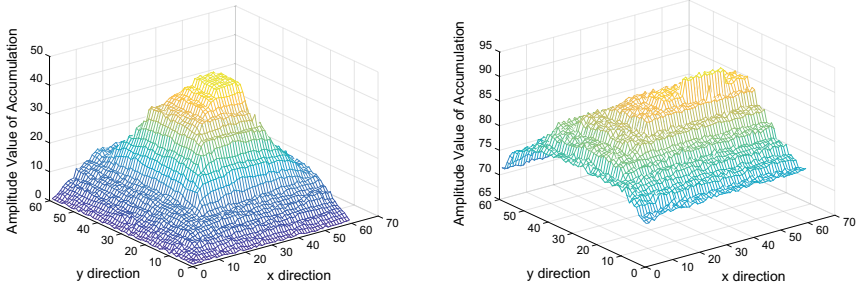
Step 7: Backtracking: for $k = K - 1, K - 2, \dots, 1$

$$\hat{X}_k = \Psi_{k+1}(x, y) \quad (29)$$

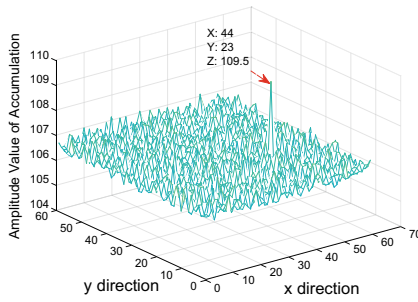
4 Simulations and Result Analysis

To verify the performance of the improved DP-TBD algorithm, the ASTS-DP-TBD algorithm is compared with other DP-TBD algorithms. Assume that the size of the radar data received per frame is 70×60 . There are a total of 22 frames of received data, and the interval $T = 1$. The target with initial state $x_1 = [1, 1.6, 0, 8, 2, 0]^T$ is to make steering movements that change in both magnitude and direction in the observation area. In addition, we assume that measurement noise obeys a Gaussian distribution. This paper uses target detect probability (p_d) and track probability (p_k) to verify the performance of the algorithm. (1) Detect probability: the probability of maximal accumulation value at last frame exceeds the threshold and error between the estimated target position and the real target position is no more than two units at the last frame. (2) Track probability: the probability of error between the estimated target position and the real target position is no more than two units at every frame. In the simulation we performed 100 Monte Carlo experiments.

Figure 1 shows a comparison of merit functions obtained after 22 frames of accumulation for various DP-TBD algorithms with an SNR of 7 dB. Figure 1a shows the merit function of traditional DP-TBD algorithm. Figure 1b shows the merit function of ASTS-DP-TBD algorithm with no state transition probability (NASTS-DP-TBD). Figure 1c shows the merit function of ASTS-DP-TBD algorithm. The first two algorithms have agglomeration effects so that the merit function amplitude value is not highlighted from clutter. We can see that ASTS-DP-TBD algorithm can well overcome the agglomeration effect and detect the true state of the target more accurately.



(a) Merit function of traditional DP-TBD Algorithm (b) Merit function of NASTS-DP-TBD Algorithm



(c) Merit function of ASTS-DP-TBD Algorithm

Fig. 1. Comparison of merit functions

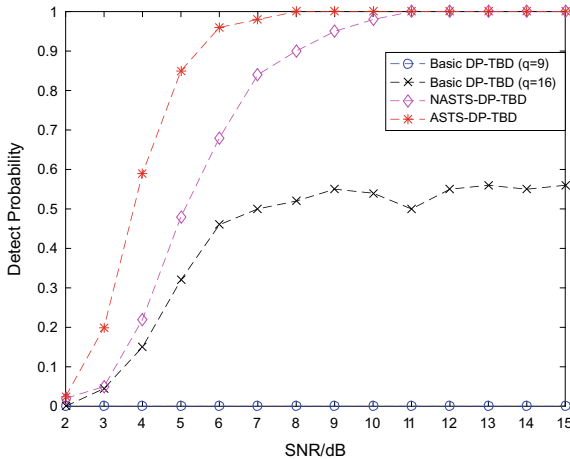


Fig. 2. Detect probability

Figure 2 shows the detect probability of basic DP-TBD algorithm, NASTS-DP-TBD algorithm and ASTS-DP-TBD algorithm. When the state transition number $q = 9$, the p_d of basic DP-TBD is almost 0. When the state transition number $q = 16$,

the performance of p_d is still not ideal. This is because the mismatch between state transition set and target velocity. From Fig. 2, we can see that the NASTS-DP-TBD and ASTS-DP-TBD have better performance, further more, when considering the state transition probability, the performance of ASTS-DP-TBD can obtain a further escalation. As what Fig. 3 shows, when the track probability closes to 1, the demanded SNR of ASTS-DP-TBD is less more 3 dB compare to the NASTS-DP-TBD. This is because the proposed algorithm uses the measurement data and target motion character jointly to confirm the real target. From Fig. 4, we can see that the NASTS-DP-TBD can effectively track the maneuvering target compare to other algorithms.

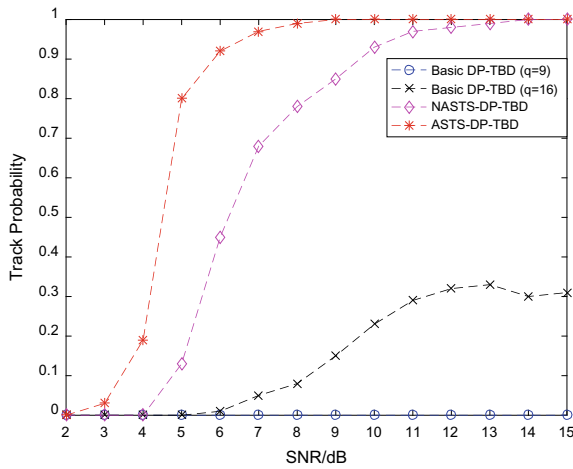


Fig. 3. Track probability

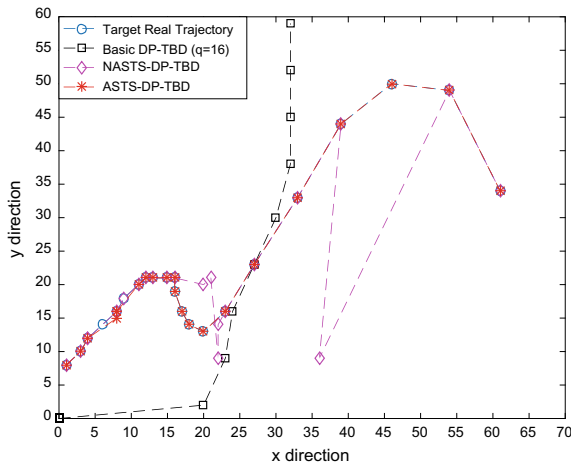


Fig. 4. Comparison of detection and tracking results

5 Conclusions

Traditional DP-TBD with constant size of state transition set can hardly detect and track the maneuvering targets because the mismatch between target velocity and transition set, and ignoring of the state transition probability leads to a loss to the algorithm's performance. This paper proposes a new method that introduces Kalman filtering and target state transition probability in the traditional algorithm. In addition, this paper also optimizes the termination decision strategy of the algorithm, which significantly improves the detection and tracking performance.

References

1. Davey, S.J., Rutten, M.G., Cheung, B.: A comparison of detection performance for several track-before-detect algorithms. In: 2008 11th International Conference on Information Fusion, pp. 1–8, Cologne, (2008)
2. Barniv, Y.: Dynamic programming solution for detecting dim moving targets. *IEEE Trans. Aerosp. Electron. Syst.* **AES-21**(1), 144–156 (1985)
3. Jiang, H., Yi, W., Kong, L., et al.: Tracking targets in G0 clutter via dynamic programming based track-before-detect. In: 2015 IEEE Radar Conference (RadarCon), pp. 0356–0361, Arlington, VA, (2015)
4. Tonissen, S.M., Evans, R.J.: Target tracking using dynamic programming: algorithm and performance. In: Proceedings of 1995 34th IEEE Conference on Decision and Control, vol. 3, pp. 2741–2746, New Orleans, LA (1995)
5. Johnston, L.A., Krishnamurthy, V.: Performance analysis of a dynamic programming track before detect algorithm. *IEEE Trans. Aerosp. Electron. Syst.* **38**(1), 228–242 (2002)
6. Liu, S., Chen, X., Zeng, T., et al.: New analytical approach to detection threshold of a dynamic programming track-before-detect algorithm. *IET Radar Sonar Navig.* **7**(7), 773–779 (2013)
7. Yue, S., Kong, L., Yang, J., et al.: A Kalman filtering-based dynamic programming track-before-detect algorithm for turn target. In: 2010 International Conference on Communications, Circuits and Systems (ICCCAS), pp. 449–452, Chengdu (2010)
8. Li, X., Wang, S., Zheng, D.: A DP-TBD algorithm with adaptive state transition set for maneuvering targets. In: 2016 CIE International Conference on Radar (RADAR), pp. 1–4, Guangzhou (2016)



New Processing Method Based on Intelligently Manufacturing Blade with Multiple Space and Compound Angles

Yuanyuan Wang^(✉), Yadi He, Pengli Zhu, Lidong Zhang,
and Feichao Zhao

Northwest Institute of Mechanical and Electrical Engineering, Shanxi 712099,
China
49231216@qq.com

Abstract. The paper mainly analyzes the development status and the development trend of the intelligent manufacturing technology in the machining field at present, presents the necessity of development of intelligent manufacturing, and creatively processes a new method of a blade with multiple space and compound angles so as to provide reference for development of intelligent manufacturing in the machining field in China.

Keywords: Intelligent manufacturing · Machining field · Digital manufacturing · Industry 4.0

1 Introduction

As the computer technology, the information technology and the automation technology develop rapidly, market competition is fiercer; intelligent manufacturing becomes the core of the high-end manufacturing industry at present and is the base and leading edge of the current manufacturing development, so that the machining and manufacturing technology faces more and more challenges. From the view of time development, intelligent manufacturing is the trend of integrating integrative development of information technology and manufacturing technology in the new century under the condition of technological development and market requirement; and particularly, intelligent manufacturing technology must be applied in the machining field of enterprises if they need to make breakthroughs and develop.

The paper mainly represents the developing situation and the developing trend of intelligent manufacturing and analyzes the application advantages of intelligent manufacturing in the field of mechanical processing at the same time; and on this basis, application of the intelligent manufacturing technology in the machining field is realized through innovation of mechanical processing method of a blade, so that the foundation is laid for the follow-up development.

2 Development of Intelligent Manufacturing Technology

2.1 Conception of Intelligent Manufacturing

The conception of “intelligent manufacturing” is first proposed by the writer of an American book named *Manufacturing Intelligence* and is that an intelligent robot can finished small batches of production by adopting technologies such as integrated knowledge engineering, robot vision, software manufacturing systems and the like under the condition of no human intervention, and the production process is called intelligent manufacturing [1].

From the twenty-first century, new information technology promotes innovation of knowledge society; conversely, knowledge innovation 2.0 promotes development and morphological transformation of the new information technology, i.e., the production of Internet Plus [2]. In short, Internet Plus is the new development form produced by integration of Internet and traditional industries; Internet Plus manufacturing is industry 4.0, of which the essential core is intelligent manufacturing, is a new conception proposed by Germany and is called as Made-in-China 2025.

2.2 Development Situation of Intelligent Manufacturing in the Field of Mechanical Processing

At present, US, Japan, EU, especially Germany with very high development level of manufacturing industry all pay great attention to the development of intelligent manufacturing technology; various researches about intelligent manufacturing were carried out in 1990s and were brought into the national strategic planning of technological development; therefore, intelligent manufacturing became the important direction of manufacturing development of the developed countries and became commanding heights of development of advanced manufacturing industry of every country [3–6]. Intelligent manufacturing of China started late, but in our country, policy researches are carried out actively and investment is increased to intelligent manufacturing industry [7].

As science and technology of our country develop increasingly, the machining field of our country develops in an intelligent and diversified manner, and machining intelligence has been realized [8]. In addition, the numerical control technology is introduced to the machining field in the continuous development process of manufacturing industry in our country, so that the purposes that the machining technology of China develops sustainably and the machining technology keeps up with the times to satisfy requirement of times can be achieved.

Machining intelligence of China achieves certain development and still has the following problems compared with that of foreign countries: The first, a small amount of money is invested to intelligent manufacturing technology, and the importance of intelligent manufacturing technology is not understood correctly; the second, machining intelligence is not understood correctly and the intensity of input and use is small; At last, the traditional production mode cannot be truly integrated with intelligent manufacturing technology. Therefore, the application of intelligent manufacturing technology in mechanical processing is restrain to some extent.

2.3 Development Trend of Intelligent Manufacturing in the Field of Mechanical Processing

Intelligent manufacturing technology is an integrated technology formed by permeating and interweaving of manufacturing technology, automation technology, system engineering, artificial intelligence and other technologies, and is driven to generate and develop by market requirements and science and technology, namely, driven by constant change of market requirements, the production scale of manufacturing develops in a flexible direction of composite variety and scalable batch; meanwhile, under the impetus of the development of information science technology, the resource allocation of manufacturing has developed in an information intensive direction. Intelligent manufacturing is not only a manufacturing process method or a product design any more but an integrated activity and system from the product concept system to the final product, and is a functional system and an information processing system. For example, from the global perspective, a digital manufacturing technology represented by 3D printing exists, namely, a computer design is taken as blueprint, specially made powdered or liquid metal material is taken as raw material, and a 3D printer is taken as the tool, so that a required product can be printed out directly through the manner that material is added in hierarchy of the material. Creation and practical application of intelligent manufacturing technology can be applied throughout the whole manufacturing process; and advanced manufacturing technologies are integrated quickly so as to make design, production, management, service and other processes of manufacturing more and more intelligent, and the manufacturing is networked, systematized, intellectualized, and globalized.

For China, the development of manufacturing plays an important role in improving the production efficiency of equipment, increasing transformation and upgrading of national manufacturing industry, lowering energy consumption, and realizing intellectualization of equipment manufacturing. In such an environment, applying intelligent manufacturing technology in the field of mechanical process becomes inevitable trend of development of manufacturing.

3 Advantages of Application of Intelligent Manufacturing Technology in the Field of Mechanical Process

3.1 Improving the Working Efficiency Through Man–Machine Integration

Man–machine integration can be realized when the intelligent manufacturing technology is applied to mechanical processing, so that the workload of traditional working staff can be reduced, the working efficiency can be improved, and the modernization development of an enterprise can be promoted. Through man–machine integration, the dominant position of workers can be embodied in the manufacturing process effectively, and the intelligent manufacturing technology can release potential abilities continually through manual operation.

3.2 Autonomous Learning and Automatic Perfection

The intelligent manufacturing technology takes the original knowledge storage as the basis and can have relates data perfected automatically through unceasing practice, use and unceasing learning in the practical manufacturing process, so that better operation can be realized. Besides, the system can restore automatically when out of order.

3.3 Improving the Production Quality

As the market competition is fiercer and fiercer, the production quality must be improved so as to improve the enterprise market competition. The intelligent manufacturing technology is of high accuracy, namely, products of related specification can be manufactured accurately in the manufacturing process, so that the production quality can be improved effectively [9].

4 Application Examples in the Field of Mechanical Processing Based on Intelligent Manufacturing

4.1 Analysis of Technological Characteristic of a Blade

The purpose of improving the firing accuracy of a correction projectile of a project is achieved by correcting the balistique through an electric steering engine on the head of the projectile. The blade is the operative part of the steering engine and is an important part of the steering engine; lateral force vertical to the projectile shaft can generate during flying by adjusting the angle between the blade and the projectile shaft, so that the flying direction of the projectile can change to as to achieve the purpose of correcting the flying posture and track of the projectile. During flying, the force that air applies on the blade must be stable and balance, so that the blade is in for form of bevels forming multiple space and compound angles; and the relative position and angle of the bevels can directly influence the flying posture of the correction projectile.

The structure of the blade is shown as Fig. 1; it can be seen from Fig. 1 that the bevels form at one end of the part and form space compound angles; the thickness direction is symmetrically vertical to the center line; a rotating body is formed at the other end; and the center of the rotating body coincides with the symmetrical center line of the bevels at the left end. The total length of the part is 82.5 mm, the thinnest part of the part is only 2 mm in thick; the angle tolerance of the two surfaces is plus or minus 10'; and the thinnest part is 0.2 mm in the thickness direction.

4.2 The Original Processing Method

The traditional processing method of the part is as follows:

Step 1 a lathe worker processes one end of the rotating body and leaves a cylindrical technical holder at the other end, a center hole is formed in an end surface of the technical holder by a center drill;

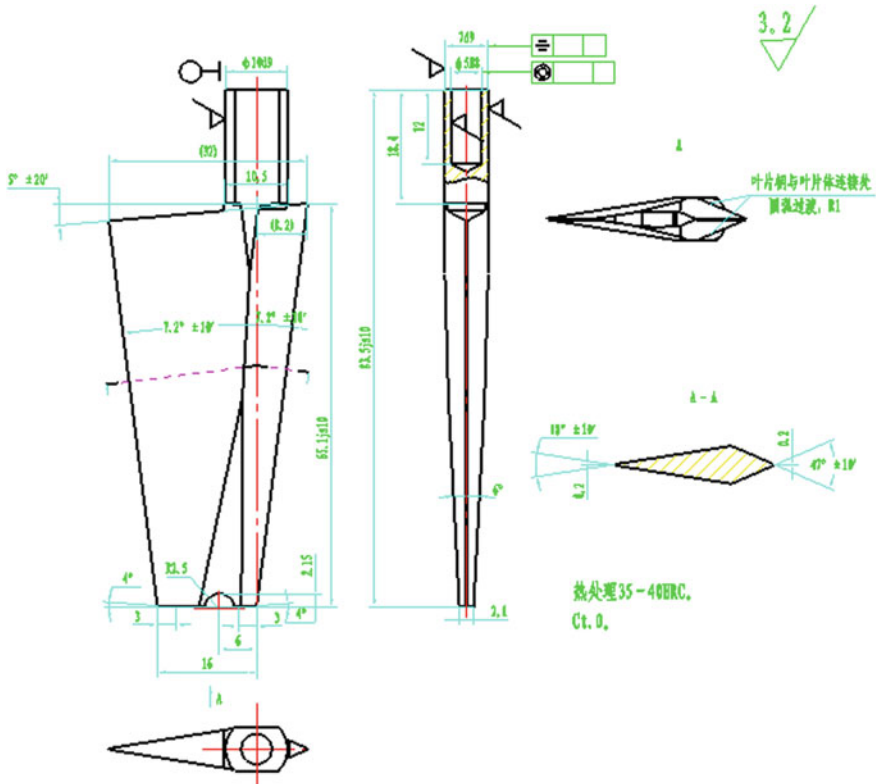


Fig. 1. The structure of the blade

- Step 2 one end of the rotating body is clamped on a linear cutting machine, so that the outline of the part can be formed, and the technical holder is still remained; and the index head is applied.
- Step 3 the common index head is used to clamped a tail seat of the rotating body and pushes up the processing centers of five shafts and the technical center hole at the other, and the shaft X, the shaft Y, the shaft Z, the shaft A, the shaft B, and the Shaft C moves simultaneously, so that processing of the bevels is finished.

The disadvantage of the processing method is that clamping is carried out twice, the error of repeated clamping causes the problem that the relative position of the part is lowered, the process is complicated, the cost is high, the cutting stress is high, and the processing period is long; and in addition, the size of a processed part is not stable, and the residual deformation is great, so that the percent of pass is only 20–30%.

As the processing volume increases, the disadvantages of the processing is more and more prominent during the processing process, and the processing plan and the processing cost are influenced seriously, so that a new processing method is required to be adopted to finish blade processing.

4.3 New Processing Method

4.3.1 Processing Technology Principle

Motor Function of an Extension Machine Realizing One-Time Cutting and Clamping of All Bevels

An index head is fixed on the linear cutting machine for clamping a blade, the shaft X, the shaft Y, and the shaft A can move through rotation of the index head, and one-time clamping of the part can be realized, so that the accuracy of relative positions of the bevels can be ensured, and cutting stress and residual deformation can be reduced.

Performing Calculation and Check Assembly Through Assembly Module Function of UG

Excel is used to calculate that the indexing accuracy can meet the requirements of a drawing (Table 1) when a 57-hole code disc is adopted; UG software is used to performing three-dimensional simulation check assembly; when the part is clamped at a certain angle, the space compound bevels can transform into a plane inner bevel; the plane inner bevel forms an included angle with a workbench, and the plane inner bevel can be directly cut by a machine tool, so that the processing difficulty is greatly reduced. Therefore, we determine that the processing think is as follows: the rotating body is still formed through turning, the technical holder is not remained at the other end anymore; the linear cutting machine is adopted to finish processing of a compound bevel; and the linear cutting machine can cut a great amount of remain, and then the processing stress is smaller, so that residual deformation of a processed part can be reduced, and the size stability of the processed can be ensured.

Table. 1. Calculation table of indexing accuracy of index head

Number of dividing head holes		66	62	59	58	57	54
The degree of one hole		0.136364	0.145161	0.152542	0.155172	0.157895	0.166667
The theory number of holes		96.06667	90.24444	85.87778	84.42222	82.96667	78.6
The actual number of holes		96	90	86	84	83	79
The degree of error		0.009091	0.035484	-0.01864	0.065517	-0.00526	-0.06667
The minimum error	0.005263	0.009091	0.035484	0.018644	0.065517	0.005263	0.066667
The actual degree		13.09091	13.06452	13.11864	13.03448	13.10526	13.16667
The target degree of drawing	13.1						

4.3.2 Processing Process

The processing process is as follows:

Step 1 the included angle between symmetric shafts vertical to the blade is calculated, and the calculation result is 13° 6’;

- Step 2 an index head is fixed on the linear cutting machine, after the length direction is found, the part is inclined for $13^{\circ} 6'$ through rotating of the index head, at the moment, molybdenum wire is parallel to one compound bevel; then programming is performed to make the molybdenum wire work on the coordinate X and the coordinate Y in a diagonal manner, so that two bevels can be cut; then the direction is rotated for $9^{\circ} 1'$ oppositely, so that other two parallel bevels can be cut;
- Step 3 the direction is rotated to 90° , and then the left end surface and a 3° bevel can be finished; Fig. 2 (small bevel at the right end);
- Step 4 the direction is rotated to 0° , and then the processing of outline can be finished; (Fig. 3).

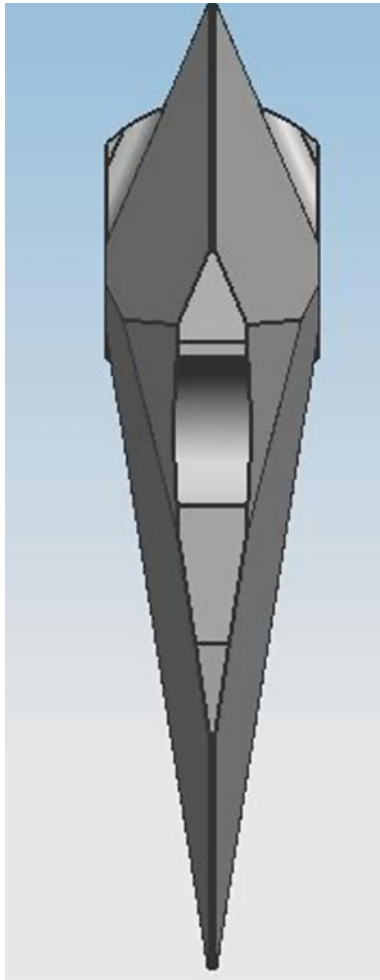


Fig. 2. $13^{\circ} 6'$ Location map

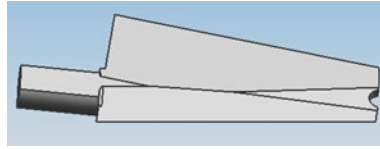


Fig. 3. Position map of cut shape fashion clip

A blade processed through the above method is proved that the use requirements can be met completely through tests and experiments, and the size stability of processed parts is good, so that the processing efficiency and the product quality can be improved greatly, the processing period is decreased to 20% of that of the original parts, and the qualification rate is increased to 100%.

5 Conclusion

In the processing process of a blade with multiple space and compound angles, a data analysis mode combining UG assembly module and Excel table is applied first, so that accuracy and speediness of calculating data are ensured. The processing method has the advantages of compensating the defect that parts cannot rotate or move during processing of a linear cutting machine, enlarging the processing range of the linear cutting machine, lowering the technical capacity requirements and workload of an operator and the processing cost and providing a new processing thinking and method for like parts, and has promotional value.

A new machining method of the part is researched to complete transition traditional mechanical processing to intelligent manufacturing; meanwhile, the advantages of intelligent manufacturing technology in the machining field is realized, and the intelligent manufacturing technology is applied scientifically, so that the machining field of manufacturing industry of China is promoted to break through to some extent and develop constantly.

References

1. Wright, P.K., Bourne, D.A.: *Manufacturing Intelligence*. Addison-Wesley (1988)
2. Huateng, M.: "Internet plus" activates more information energy. In: *Guangming Daily* (2015)
3. Sandler, J.U.: *Industry 4.0 the Coming Fourth Industrial Revolution*. Machinery Industry Press
4. Nai-Ming, L.: *Industry 4.0 the Practice of Digitizing the Equipment Manufacturing Industry*. Machinery Industry Press, Beijing (2015)
5. Jie, Z., Lv, Y.-L.: Current situation and development trend of intelligent manufacturing. *High Technol. Ind.* **11**(3), 42–47 (2015)

6. Jian-zhong, F.U.: Development status and trend of intelligent manufacturing equipment. *Mech. Electr. Eng.* **31**(8), 959–962 (2014)
7. Chang, J.Y.: On core value and new characteristics of intelligent manufacturing. *Ind. Technol. Innov.* 1031–1033 (2016)
8. Li, H.-L.: The application of intelligent manufacturing in the field of machine adding. *Technol. Equip.* 147–148 (2017)
9. Li, X.-H., Gao, B.-B.: The development of intelligent manufacturing technology and its research and application in defense abroad. *Natl. Def. Manuf. Technol.* (4), 9–12 (2015)
10. Yang, X.-P.: Discussion on status and development tendency of intelligent manufacture technology. *Intern. Combust. Engine Parts* 132–133 (2015)



Design of Three-Phase AC-Voltage Regulator Energy-Saving System by SOPC

Tu Ya^{1,2(✉)}, Jialu Du¹, and Peng Cui³

¹ School of Marine Electrical Engineering, Dalian Maritime University, Dalian, Liaoning 116026, China

tuya@neusoft.edu.cn, dujl66@163.com

² Dalian Neusoft University of Information, Dalian, Liaoning 116023, China

³ Liaoning Vocational College Light Industry, Dalian, Liaoning 116100, China
18532450@qq.com

Abstract. Nios II is applied in three-phase AC-voltage regulator energy-saving system as the core microprocessor of FPGA system in this paper. A complete voltage regulating main topology circuit is designed. Only three groups of triggering signals are required to control the six thyristors' conducting angles. A SOPC system is created by Qsys and Quartus II. The PID control algorithm is used to adjust the output voltage in real time according to the load change to save energy. The generation of high-precision triggering signals and PID control regulator functions are realized by Nios II, and simulated in MATLAB/Simulink. Experiments and in-field tests have shown the feasibility of the proposed scheme. The whole system's analysis, synthesis, simulation and configuration are completed by Quartus II. Experiments show that the method is simple and reliable, and achieves a better energy-saving effect.

Keywords: Thyristor · Three-phase AC-voltage regulator · Energy-saving system · SOPC

1 Introduction

Nowadays, energy issues have become the focus of attention of governments around the world. Research and application of energy-saving technology can reduce industrial power consumption, save energy costs for enterprises and reduce environmental pollution. If the output voltage of the voltage regulator can be adjusted with the change of the load, and the power factor is kept constant in a certain range, the purpose of saving energy is realized. Three-phase AC-voltage regulator is widely used in industry, electricity, transportation, and other fields. At present, most of the thyristors are used to realize AC-voltage regulator. Its advantage is that the control circuit is simple and the power capacity is large. But the drawback is that the power factor will decrease when the control angle increases. The traditional method is to generate trigger pulses to control thyristors by microprocessor [1]. However, trigger pulse generation algorithm

has some difficulties for microprocessor. With the appearance of FPGA/CPLD, an effective way to control thyristor is provided. FPGA/CPLD is highly capable of handling logic operations and numerical operations. As the development of modern SOPC (System on Programming Chip) technology, the application and design based on Nios II and VHDL/Verilog HDL hardware description language have new design methods and design process [2].

In this paper, the complex control algorithm is realized by Nios II processor and HDL language. It not only reduces the development cycle, but also embodies the design ideas for users and applications. Finally, this SOPC system is implemented by Quartus II and Qsys. This control system is tested on the three-phase AC-voltage regulator energy-saving system Results show this system hardware and software design is reasonable and has good energy-saving effect.

2 Principle and Simulation

2.1 The Principle of Phase-Controlled Voltage Regulator

The phase-controlled AC-voltage regulator changes the supply voltage by controlling the phase angle of the switching device, so as to effectively control the load voltage. Unidirectional AC-voltage regulator with resistance load circuit as shown in Fig. 1.

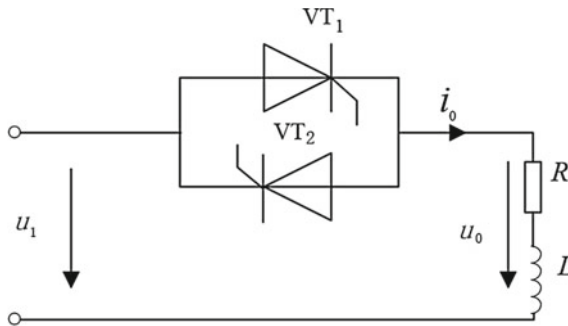


Fig. 1. Unidirectional AC-voltage regulator with resistive load

The property of the load has a great influence on the circuit. The output voltage of the phased the effective value U_o of the load voltage is

$$U_o = \sqrt{\frac{1}{\pi} \int_{\alpha}^{\pi} (\sqrt{2}U_i \sin \omega t)^2 d\omega t} = U_i \sqrt{\frac{1}{2\pi} \sin 2\alpha + \frac{\pi - \alpha}{\pi}} \tag{1}$$

U_i is the effective value of AC-voltage regulator input. The effective value of load current is $I_o = U_o/R$, and its factor is

$$\cos \varphi = \frac{P}{S} = \frac{U_o I_o}{U_i I_o} = \frac{U_o}{U_i} = \sqrt{\frac{1}{2\pi} \sin 2\alpha + \frac{\pi - \alpha}{\pi}} \tag{2}$$

The phase shift range of the α is $0 \leq \alpha \leq \pi$, thyristor conduction angle $\theta = \pi - \alpha$, and voltage adjusting range is $0 \sim U_i$. With the increase of α , the power factor of the circuit is also reduce. Load impedance angle $\varphi = \tanh^{-1}(\omega L/R)$, when the load is the inductive load. The current in the circuit is discontinuous when $\varphi \ll \alpha \ll \pi$. The effective value of inductive load voltage is:

$$U_o = \sqrt{\frac{1}{\pi} \int_{\alpha}^{\alpha+\pi} (\sqrt{2}U_i \sin \omega t)^2 d\omega t} = U_i \sqrt{\frac{\theta}{\pi} + \frac{\sin 2\alpha - \sin(2\alpha + 2\theta)}{2\pi}} \tag{3}$$

2.2 Three-Phase AC-Voltage Regulator Method for Thyristor

The key point of the thyristor three-phase power-supply regulator is the trigger angle is calculated in accordance with the inter-phase regular of the three-phase power supply to trigger the corresponding six thyristors reliably and effectively, on the basis of the power-supply voltage or current synchronization signal is collected accurately. Double pulse synchronous trigger mode by power-supply voltage is used in this system. These trigger mode requirements control the trigger thyristor turn-on and turn-off in turn, which to the synchronization signal as a reference, according to the pre-designed sequence, even if this outside is disturbed does not change the sequence [3–5].

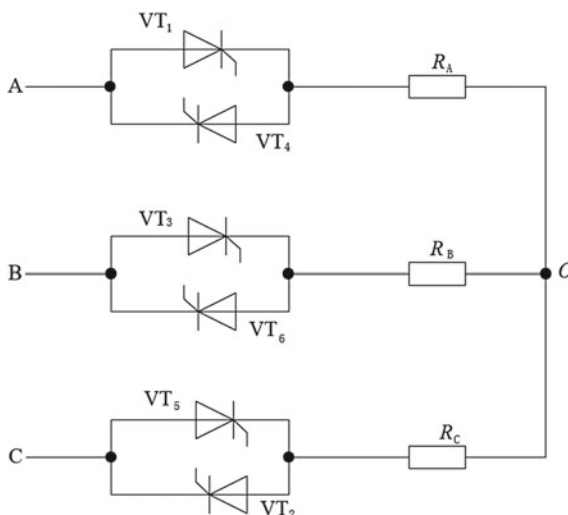


Fig. 2. Three-phase and three-wire circuit

So as to achieve high reliability, high precision of the three-phase voltage regulation must satisfy the following conditions: accurate synchronizing signal acquisition; High precision and strong anti-interference ability of the thyristor trigger pulse and isolation of output. In order to achieve the above objectives on the basis of references to the other researches, proposed using the FPGA of EP4CE115F29C7 combining specific synthetic trigger circuit of main circuit. Take three-phase and three-wire star type circuit as an example, as shown in Fig. 2. The output voltage is shown in Fig. 3, when the trigger pulse when $\alpha = 60^\circ$ that is shown in Fig. 4,

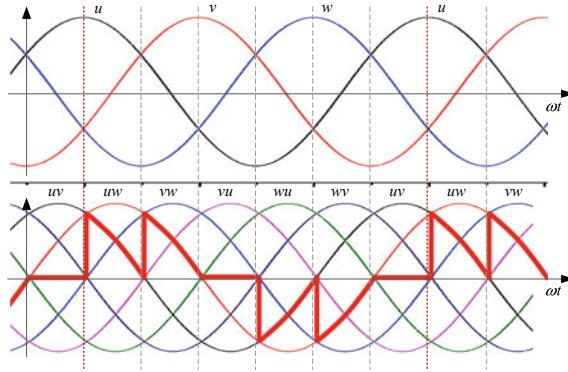


Fig. 3. The output of three-phase AC power

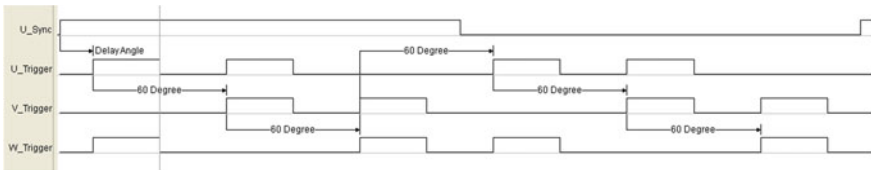


Fig. 4. Thyristor double pulse trigger signal

2.3 System Model and Simulation

The model of the three-phase power regulator system was established in MATLAB/Simulink with 10Ω resistance in series of $1\mu\text{H}$ inductors as load. The three-phase power output voltage is shown in Fig. 5, when α is 30° , 60° , 90° , and 120° .

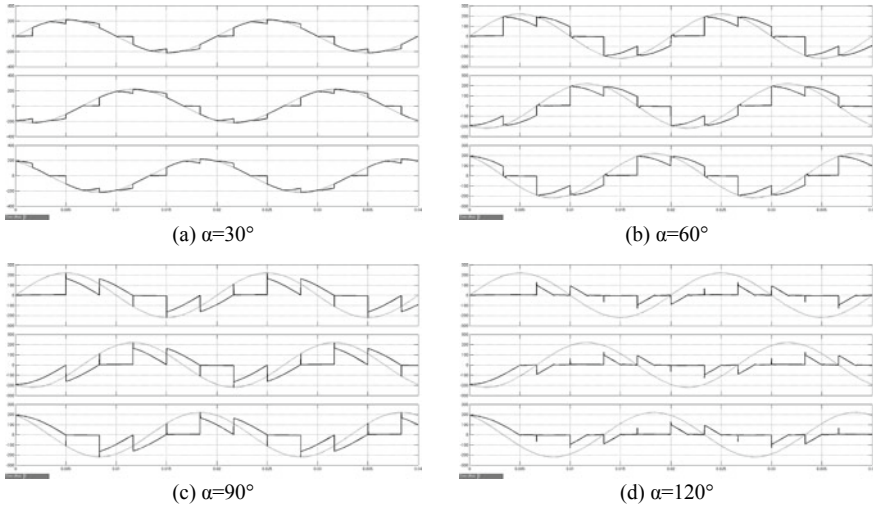


Fig. 5. The output of three-phase AC power supply at different phase angles

2.4 The Principle of Control System

According to inputs deviation values, the PID control carries out multiplication and addition operation through the proportion, integral, and differential rules. Finally, the operational results for the output are taken. In the analog system, the PID algorithm equation is

$$p(t) = K_p[e(t) + \frac{1}{T_I} \int e(t)dt + T_D \frac{de(t)}{dt}] \tag{4}$$

where $P(t)$ is output of controller, $e(t)$ is deviation, K_p is proportion factor, T_I is integral time constant, and T_D is differential time constant. As for FPGA, it can only calculate the controlled quantity by time sampling deviation. So it is necessary to discretize (4). The discrete equation of PID is

$$P(k) = K_p\{e(k) + \frac{T}{T_I} \sum_{j=0}^k e(j)T + \frac{T_D}{T}[e(k) - e(k - 1)]\} \tag{5}$$

where k is sampling sequence, T is sampling period, $e(k)$ is the deviation at the time of the k -th sample, $e(k - 1)$ is the deviation at the time of the $(k - 1)$ -th sample, and $P(k)$ is the output of controller at the time of the k -th [6].

3 System Implementation

3.1 Hardware Design

The system of three-phase AC-voltage regulator energy-saving is composed of the output phase to phase voltage detection circuit, zero-crossing synchronization signal detection circuit, phase sequence detection circuit, display and data output circuits, the core circuit of FPGA, trigger pulse generation and the output circuit, photoelectric isolating and driving circuit of thyristor, and voltage regulating circuit of Thyristor, as shown in Fig. 6. The system can calculate the power factor angle according to the zero synchronization signal and the phase difference of current and voltage of load are detected, in addition it can regulate the output voltage according to adjust triggering angle of output trigger pulse sequence, so as to the power factor is improved greatly. The load voltage is collected by the mutual inductor. The current detection sensor uses Hall sensor. Its maladjustment current is less than 30 mA and the respond speed is less than 10 s. The analog–digital conversion is implemented with AD9235 of which the conversion accuracy is 12-bit and 65MSPS data rates.

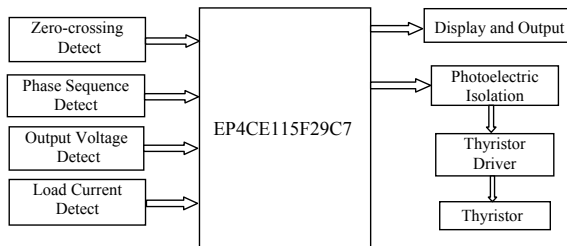


Fig. 6. The block diagram of system structure

3.2 Software Design

The design of software is carried on TOP-DOWN method based on Quartus II. The whole design's Analysis, Synthesis, Simulation and configuration are completed by Quartus II. Qsys is an object-oriented development tool for SOPC, which was introduced by Altera. The Nios II processor core and peripheral interfaces for this project can be built through the Qsys tool. The design of SOPC platform is shown in Fig. 7 where AD_9235CTRL is A/D converter driver modules. The starting point of the trigger pulse signal is calculated by Nios II with as shown in Fig. 8. According to the power angle φ , the voltage of need to downregulated, and the U-phase voltage synchronization signal as reference signal [7].

3.3 System Debug

The beam pumping unit with 30 KW three-phase asynchronous motor is used as test object in this system. The load changes from time to time depending on the running

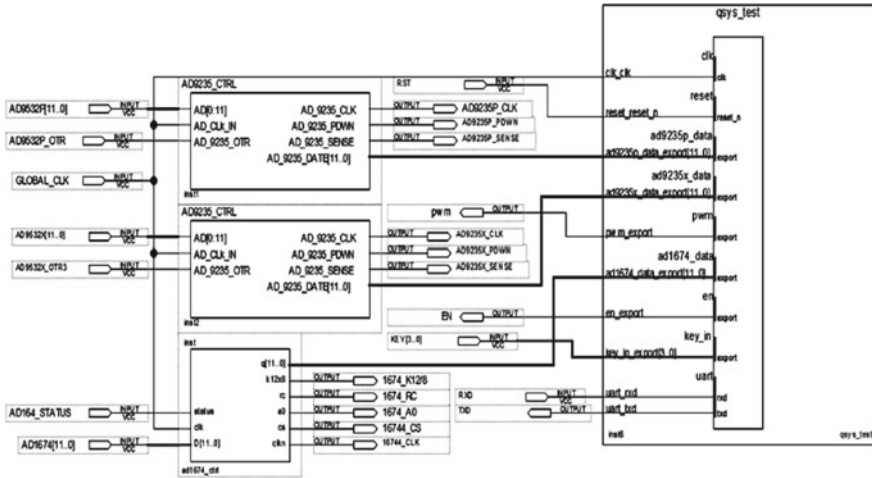


Fig. 7. The SOPC platform

state in this system. The three-phase output voltage with a phase angle of 30° and 60° is shown in Fig. 9. As shown in these figure, the test results are consistent with the simulation. By experiment, the scope of regulating voltage from 150 to 220 V range is adjusted, but after more than 90° increase in the number of higher harmonic will lead to motor stalling [8].

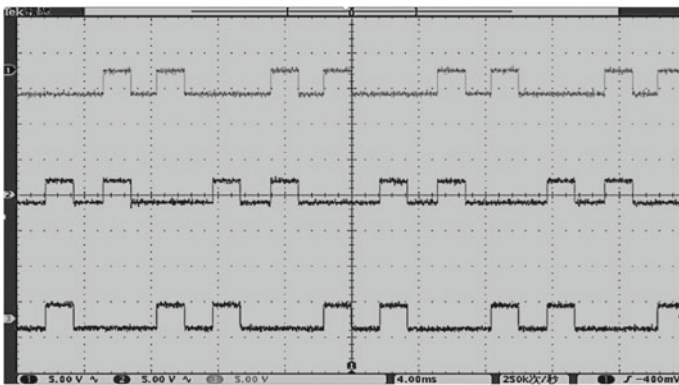


Fig. 8. Trigger pulse signal of the FPGA running

The load current and output voltage are detected and sent to the computer terminal through serial port for analysis. The result of the test is shown in Fig. 10. According to the comparison of load current and output voltage, the system has good load tracking characteristics and good energy-saving effect.

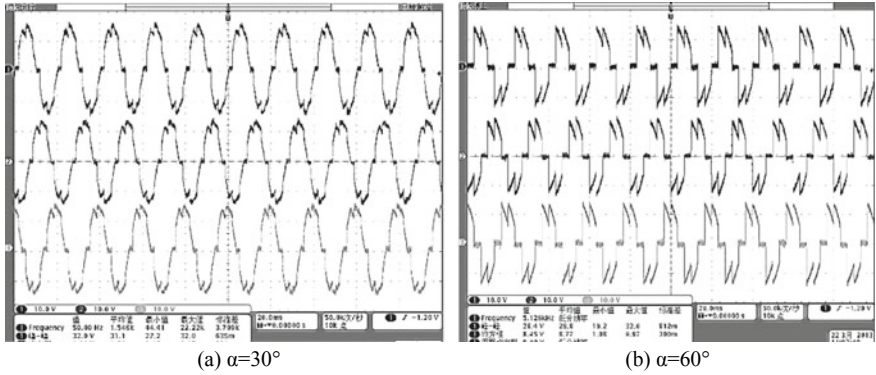


Fig. 9. The wave of regulator output at 30° and 60° with three-phase asynchronous motor

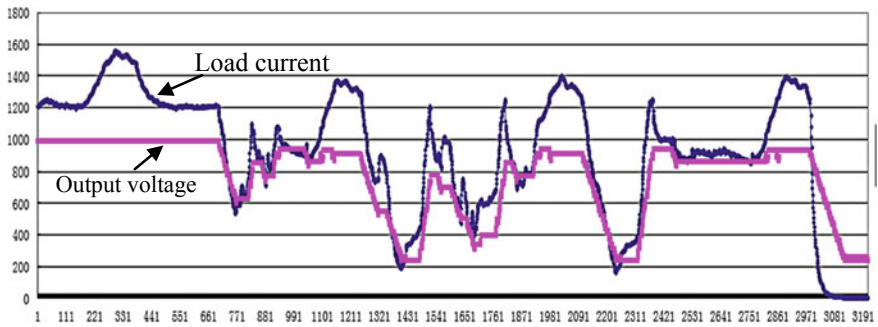


Fig. 10. Load tracking waveform

4 Conclusion

The system using FPGA with the advantages of high speed, work stability, and easily modifying have realized three-phase AC-voltage regulator energy-saving system and corresponding with the external circuit, the whole system performance is greatly improved. The design of the control algorithms with Quartus II and the Nios II embedded processor is customized with Qsys, which have the advantages of a great deal of flexibility, augmentability, and short design development cycle [9, 10]. This energy-saving system can be applied to the ship electric propulsion, dynamic positioning and pumping units and other energy-saving equipment, which is of further reference significance for the realization of energy conservation.

References

1. Zhou, G.S., Zhang, Y., Shen, H., et al.: Design of SCR three-phase AC-voltage regulator based on STM32 MCU. *Electron. Des. Eng.* (2013)
2. Solet, D., Béchenec, J.L., Briday, M., Faucou, S., Pillement, S.: Hardware runtime verification of embedded software in SOPC. *Ind. Embed. Syst.* 1–6 (2016)
3. Zinoviev, G.S., Sidorov, A.V., Kharitonov, S.A.: Three-phase AC voltage regulator as part of an autonomous system. In: *International Conference on Actual Problems of Electronics Instrument Engineering*, pp. 762–765. IEEE (2015)
4. Zhou, G., Tu, Y., Zhao, S.: A three-phase AC-voltage regulator system. *Telkomnika Indones. J. Electr. Eng.* **12**(5), 3501–3508 (2014)
5. Li, G., Liu, H., Gu, J., et al.: A novel three-phase series AC voltage regulator. In: *Proceedings IPEMC 2000. The Third International Power Electronics and Motion Control Conference, 2000*, vol. 3, pp. 1108–1110. IEEE (2002)
6. Sun, H., Wang, X., Lin, Q., Wang, X., Su, S.: The design of the DC servo motor controller based on fuzzy immune PID algorithm. In: *Chinese Control Conference*, pp. 4724–4729 (2017)
7. Jiang, M.C., Pan, P.E., Liu, T.C., et al.: A soft-switching modular three-phase AC voltage regulator. In: *Telecommunications Energy Conference*. IEEE (2016)
8. Rozhdenko, D.A., Udovichenko, A.V.: Digital control system with system of automatic control for three-phase two-zone thyristor AC voltage regulator. In: *International Conference of Young Specialists on Micro/nanotechnologies and Electron Devices*, pp. 374–376. IEEE (2013)
9. Orallo, C.M., Carugati, I., Funes, M., et al.: Thyristor gate control implementation on FPGA for particle accelerator facilities. In: *Micro-Nanoelectronics, Technology and Applications*, pp. 48–53. IEEE (2014)
10. Lati, A., Irki, Z., Sakhi, S., Nemra, A., Hamerlain, M.: Implementation of features detection and matching algorithms on FPGA using Nios II. In: *International Conference on Modelling*, pp. 55–60 (2017)



A Flexible Broadband Single RF Architecture Based on Time-Modulated Array

Xiaofeng Ling and Nan Wang^(✉)

School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China
{xflying,wangnan}@ecust.edu.cn

Abstract. With a single RF front-end to equivalently realize RF transformation of multiple antenna elements, single RF array is an important way to reduce the manufacturing cost of massive MIMO systems. Time-division-multiplexing-based single RF array is suitable for uplink receiver array. However, there are some bottlenecks in broadband applications, such as very fast switching speed requirement of RF switch, wide bandwidth requirement of analog front-end and low signal fidelity, which stop the wide use of this technology. To overcome the above bottlenecks, a time-modulated-array-based broadband single RF receiver array architecture is proposed in this paper, and the theory of time-modulated array along with the idea of signal reconstruction is also introduced to broadband the single RF array. With these methods, the performance requirements of RF switching speed and analog front-end are expected to be reduced significantly while the signal fidelity degradation can be safely ignored. The experimental results show that the proposed architecture can improve the bandwidth of the system with traditional RF switching elements.

Keywords: Single RF array · Time-modulated array · Massive MIMO · Signal reconstruction · Time-division multiplexing

1 Introduction

Large-scale antenna array is one of the key technologies in modern mobile communication and radar systems, and it offers a variety of interesting applications for civilian and military organizations, including direction of arrival estimation, multiple-input-multiple-output (MIMO) operation, digital beamforming and interference suppression [1]. However, these conventional array systems typically require a separate RF front-end and digitizer for each of the array antenna elements, thereby increasing the size, weight, power consumption, and cost of the systems. Therefore, single RF array technology, which uses a single RF front-end

equivalent to realize multiple elements of RF conversion, is proposed to effectively overcome the economic and complex problems of large-scale antenna array.

The idea of single RF was first proposed by Fredrick et al. [2] in 2002, and their work enables a drastic reduction in hardware requirements for smart antenna system through time-division multiplexing; however, this technique also has several short-comings, such as: signal distortion, xxxcan only be applied the prototype with narrow band pass and a limited number of antenna elements due to the high need for RF switching speed and audio-to-digital converter (ADC) performance. Therefore, different alternatives to realizing orthogonal transmission systems that result in a single RF front-end MIMO system [3], and several types of code-division-multiplexing-based single RF front-end have been recently [4, 5] to accommodate various multi-antenna schemes including spatial multiplexing, spatial diversity, and beamforming; however, the noises from the receivers are also enlarged, resulting in a significant gain loss. To improve the quality of output signal, Henault et al. [6] designed a single RF MIMO receiver that adopts a beam-switching antenna instead of a conventional array antenna, and this design architecture can be deployed in a very small physical space while achieving a full spatial multiplexing gain; however, this technique is only suitable for the low frequency waves in time-division scenarios.

In this paper, the theory of time-modulated array and method of signal reconstruction to broadband single RF array are demonstrated, and this method significantly reduces the performance requirements of RF switch speed and analog front-end on the premise of signal fidelity remaining unchanged. To better describe this proposed method, we first build the theory model of the array system, and then demonstrate the reconstruction condition of antenna element signal, and finally present the specific signal reconstruction method. The experimental results show the effectiveness of the proposed method in reducing the requirements of RF switching speed when trying to maintain the system performance: to a single RF array with 15 receiver elements and 10 MHz receiving bandwidth, the switching speed requirement could be reduced by 70% at a penalty of 2 dB array gain loss, and therefore, significantly improves the system bandwidth.

2 Preliminaries of Single RF Digital Beamforming System

The modern standard digital beamforming system provides significant performance improvement over analog beamforming system [1]. The antenna system needs the array signals of each array element, and the usual way to obtain these signals is to connect each array with a microwave digital receiver. Figure 1a presents the conventional structure of this system. However, this architecture incurs significant hardware cost because we need to deploy M RF receivers for M signal arrays. The main idea to put down the hardware cost is to reduce the number of RF receivers from M to 1, and the related research works mainly fall into two categories: The first structure is called time-division multiplexing as showed in Fig. 1b, in which a microwave switch selects one channel from all M

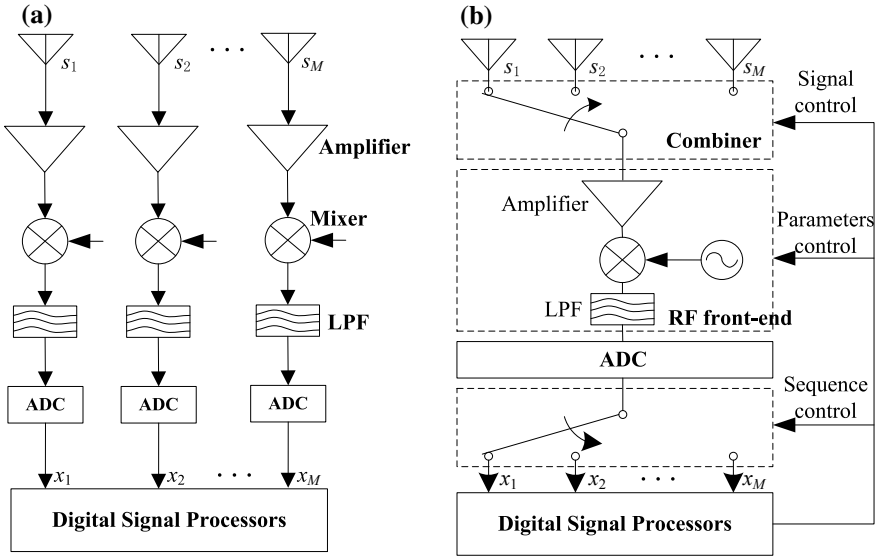


Fig. 1. Digital beamforming systems. **a** The standard digital beamforming system. **b** Time-division-multiplexing-based single RF digital beamforming system.

antenna channels and record the array signal sequentially in one sampling cycle just like the time-division multiplexing signal selection technology [2]. The second type is the code-division multiplexing, where the input signals from antenna elements share a single RF path by code modulation and demodulation [4, 5].

Compared with code-division multiplexing, time-division multiplexing method employs RF switches which have very simple structure to replace the complex code modulation and demodulation circuit, and thus, obviously reduces the complexity of the system and avoids the problem of big gain loss of code-division multiplexing modulator. However, the present time-division multiplexing single RF structure requires that the switching speed must be at least $2 * M$ times of the array bandwidth (M is the number of antenna elements), and therefore, the switching speed of RF switch becomes a bottleneck in realizing the broadband single RF arrays.

In this paper, a broadband single RF receiver array based on time-modulated-array theory is developed, which is demonstrated in Fig. 2a. The time modulation circuit and real-time signal reconstruction processing algorithm are also introduced to overcome the bottlenecks of the current single RF structures.

3 The Proposed Architecture

3.1 System Composition

In this subsection, the system of our proposed broadband single RF receiver array is demonstrated, which is presented in Fig. 2a. This proposed architecture can be divided into the following four parts.

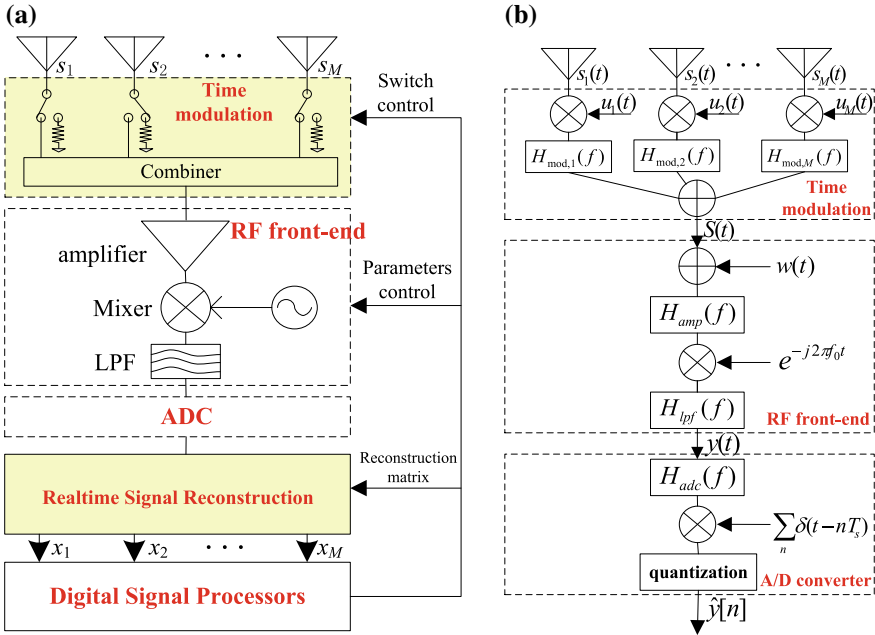


Fig. 2. a System architecture of the proposed single RF receiver array. b Theoretical model of the proposed single RF receiver array.

1. **Antenna array:** It consists of M antenna elements, and the received signal of each element is denoted as s_m , ($m = 1, 2, \dots, M$). We consider all the antenna arrays to be the same, and the differences between each array can be neglected. Instead, the signal reconstruction becomes one of the key considerations when deploying the antenna arrays.
2. **Time modulation:** It employs M independent single-pole double-throw RF switches to replace the single-pole M -throw RF switches of the time-division multiplexing structure. The switches in our system must have high switching speed, and this enlarges frequency response variation between switch circuit and combiner, which must be counted.
3. **RF front-end and ADC:** It has the classic superheterodyne digital receiver structure. In order to support the broadband characteristics of the array, the digital receivers should have very large instantaneous bandwidth (no less than M times of array bandwidth) to support the broadband characteristics of the array, and the noises of amplifier, mixer and ADC must be counted to maintain a high quality of signal processing result. In addition, the bandwidth of the digital receiver is extremely large, and the nonideal characteristics of each element also need to be dealt with.
4. **Real-time signal reconstruction:** It uses the high-speed digital signal obtained by the ADC to estimate the input signal of each antenna element, and the FPGA which has a strong parallel processing capability is recommended as the processing platform to achieve the real-time performance.

3.2 The System Modeling Based on the Time-Based Modulation Array Theory

The signal flow of each component can be established following the previous theoretical modulations, and the signal flow of the whole single RF array can be built as shown in Fig. 2b. $u_m(t)$ is the switch function of the m -th antenna array at time t . This is a periodical function with the period to be τ , which can be described as follows:

$$u_m(t) = \begin{cases} 1, & \mu_{1,m}\tau \leq t \leq \mu_{2,m}\tau; \\ 0, & \text{others;} \end{cases} \quad (1)$$

$H_{mod,m}(f)$ is the frequency response function of the m -th array signal caused by switch and combiner. $w(t)$ and $e^{-j2\pi f_0 t}$ represent the hot noise and the local oscillator of RF front-end, respectively. $\sum_n \delta(t - nT_s)$ is the sampling impact string function whose period is T_s . $H_{amp}(f)$, $H_{lpf}(f)$ and $H_{adc}(f)$ are the transmission functions of amplifier, filter and AD converter, respectively.

Then, the mapping between the input signal sets $\{s_m(t)|m = 1, 2, \dots, M\}$ ($\{s_m(t)\}$ for short) and the ADC output $\hat{y}[n]$ is established. Each switch function $u_m(t)$ can be designed according to the requirement, and then, the output response function before sampling can be calculated after selecting the switch function set $\{u_m(t)|m = 1, 2, \dots, M\}$ ($\{u_m(t)\}$ for short). Finally, the completed output response function of the array model $y(t)$ and its discrete form $\hat{y}(t)$ can be derived as follows:

$$y(t) = F(\{s_m(t)\})|_{u_m(t)} \quad (2)$$

$$\hat{y}(t) = \hat{F}(\{s_m([n])\})|_{u_m(t)} \quad (3)$$

According to the theory of time-modulated array, the energy of signal $s(t)$ after time modulation is spread to the each harmonic whose frequency is $f_{sw} = 1/\tau$, and its frequency domain is expressed as follows:

$$S(f) = \sum_{m=1}^M \sum_{k=-K}^K G_m^{(k)} \cdot S_m(f - k \cdot f_{sw}) \quad (4)$$

where $S_m(f)$ is the frequency domain of the input signal from the m -th array element, $G_m^{(k)}$ is the coefficient of the Fourier series with k order, and $H_{mod,m}(f)$ is the RF switch frequency response. Therefore, the frequency domain of signal $\hat{y}(t)$, which is denoted as $\hat{Y}(f)$, can be expressed as follows.

$$\begin{aligned} \hat{Y}(f) = & H_{adc}(f)H_{lpf}(f) \left[\sum_{m=1}^M \sum_{k=-K}^K G_m^{(k)} S_m(f - f_0 - k f_{sw}) \right. \\ & \left. + W(f - f_0) \right] H_{amp}(f - f_0) \end{aligned} \quad (5)$$

where $W(f)$ is the frequency domain of the noise signal $w(t)$.

The time domain of $\hat{y}(t)$ can be calculated by transforming $\hat{Y}(f)$ to time domain, and discrete form of $\hat{y}[n]$ is described as $\hat{y}[n]$. If Nyquist sampling condition is satisfied, $\hat{y}[n]$ contains all the information of $\hat{y}(t)$, and the input signal can also be described by its discrete form $s_m(n)$.

3.3 Array Signal Reconstruction

By introducing the time-modulated-array theory, the single RF receiver array can be designed with less limitations, and the time to switch on and off the switches in time modulation $\{u_m(t)\}$ can be set freely; however, not all possible $\{u_m(t)\}$ ensure the holographic characteristics of the whole single RF receiver array. Therefore, we must first determine the conditions that the input single set $\{s_m(t)\}$ can be derived from the ADC output $\hat{y}[n]$, and then, the inverse function of $\hat{F}(\cdot)$ (Eq. 3) must be derived.

The signal reconstruction process can be divided into the following three stages:

1. **Step 1:** The frequency domain of $\hat{y}(t)$, which is reconstructed from $\hat{y}[n]$, is described as $\hat{Y}(f)$, and $\hat{y}[n]$ must satisfy Nyquist sampling condition.
2. **Step 2:** Estimate the frequency response of $s(t)$ by $\hat{Y}(f)$, and it is completed by two steps of equalization and demodulation. The estimated value of $S(f)$ in the Passband range is denoted as $S'(f)$.
3. **Step 3:** The set of discrete forms of signals from array elements is denoted as $\{s_m(n)\}$, and it is derived from $S'(f)$. In this work, we assume that the bandwidth of $\hat{Y}(t)$ is M times of the bandwidth of $s_m(t)$, and $S'(f)$ can be described as follows:

$$S'(f) = \sum_{m=1}^M \sum_{k=-(M-1)/2}^{(M-1)/2} G_m^{(k)} \cdot S_m(f - k \cdot f_{sw}) \tag{6}$$

The time domain of the above equation is described as follows.

$$S'(t) = \sum_{m=1}^M \sum_{k=-(M-1)/2}^{(M-1)/2} G_m^{(k)} e^{-j2\pi k f_{sw} t} s_m(t) \tag{7}$$

where $s'(t)$ is the time domain representation of $S'(f)$. Considering that $f_s = M \cot f_{sw}$, the discrete form of $S'(t)$ is described as follows:

$$[s'[M_n + 1], s'[M_n + 1], \dots, s'[M_n + 1]]^T = A \cdot [s_1[n], s_1[n], \dots, s_1[n]]^T \tag{8}$$

where A is a matrix, whose element is denoted as $a_{pm} \sum_{k=-K}^K G_m^{(k)} \cdot e^{-\frac{j2\pi k p}{M}}$.

Then, the $\{s_m(n)\}$ can be calculated from Eq. (8), with a precondition that the matrix A is reversible.

4 Experimental Results

To demonstrate the effectiveness of our proposed theory, we simulated the time-modulated-array-based single RF receiver array architecture, and compare the simulated results against the time-division-multiplexing-based single RF array [6]. The number of antenna elements in this experiment is set to be $M = 15$, RF

switching period is set to be $\tau = 50$ ns, and the instantaneous bandwidth of the array is set to be 10 MHz.

To simplify the experiments, we assumed that the RF switching time modulation circuit has an ideal cut-off characteristic (which can be described by Eq. 1). In addition, the simulated transfer functions of RF front-end amplifier, RF front-end filter and the ADC are assumed to satisfy the flat bandpass characteristics, which means that the frequency responses of $H_{amp}(f)$, $H_{ipf}(f)$, and $H_{adc}(f)$ in bandpass remain as constants. Furthermore, the switching control signal of our time modulation circuit is set the same as that of the time-division-multiplexing-based single RF array. The pulse interval delay of adjacent array element is $\frac{\tau}{M} \approx 3.33$ ns, and the duty cycle of pulse is $\mu_{2,m} - \mu_{1,m}$, which can be set by the values among $[\frac{1}{M}, \frac{1}{2}]$.

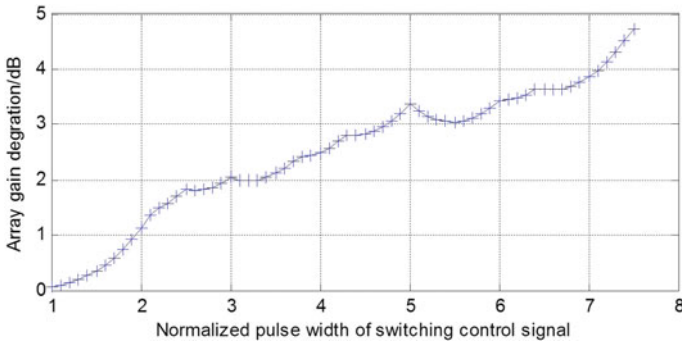


Fig. 3. The array gain loss under different duty cycles of pulse.

The simulation results of the array gain loss under different duty cycles are presented in Fig. 3. The X-axis gives the normalized pulse width (normalized by $\frac{\tau}{M}$, which is the pulse width of time-division-multiplexing), and the Y-axis gives the array gain loss. The experimental results show that our proposed architecture has a better flexibility in obtaining the array gain by modifying the switching pulse width. Alternatively, the required RF switching speed, which is the multiplicative inverse of RF control pulse width, can be greatly reduced with a penalty of small array gain loss. The experimental result show that the switching speed requirement could be reduced by about 30% if we allowed $2dB$ array gain loss, and the array bandwidth can be three times of that in the time-division-multiplexing-based single RF array if the RF switches used are the same, and thus, significantly improves the performance of system bandwidth.

5 Conclusions

Time-division-multiplexing based single RF array is suitable for up-link receiver array, however, there are some bottlenecks in broadband applications, such as very fast switching speed requirement of RF switch, wide bandwidth requirement

of analog front-end and low signal fidelity. Therefore, a time-modulated-array-based single RF architecture is proposed in this paper to solve the above problems, by reducing the performance requirements of RF switch and analog front-end. The experimental results demonstrate that this proposed method can significantly increase the RF switch flexibility with a small penalty of array gain loss, and therefore, improve the performance of system bandwidth.

Acknowledgements. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant 61604054, and Shanghai Aerospace Science and Technology Innovation Fund (SAST2016097, SAST2017112).

References

1. Barton, P.: Digital beam forming for radar. *IEE F-Commun. Radar Signal Process.* **127**(4), 266–277 (1980)
2. Fredrick, J.D., Yuanxun, W., Itoh, T.: A smart antenna receiver array using a single RF channel and digital beamforming. *IEEE Trans. Microwave Theory Tech.* **50**(12), 3052–3058 (2002)
3. Mohammadi, A., Ghannouchi, F.M.: Single RF front-end MIMO transceivers. *IEEE Commun. Mag.* **49**(12), 104–109 (2011)
4. Tzeng, F., Jahanian, A., Pi, D., Heydari, P.: A CMOS code-modulated path-sharing multi-antenna receiver front-end. *IEEE J. Solid-State Circuits* **44**(5), 1321–1335 (2009)
5. Zhang, J., Wu, W., Fang, D.G.: Single RF channel digital beamforming multibeam antenna array based on time sequence phase weighting. *IEEE Antennas Wirel. Propag. Lett.* **10**, 514–516 (2011)
6. Henault, S., Jackson, B.R., Antar, Y.M.: Compensation of time-division multiplexing distortion in switched antenna arrays with a single RF front-end and digitizer. *IEEE Trans. Antennas Propag.* **61**(8), 4383–4388 (2013)
7. Kalis, A., Kanatas, A.G., Papadias, C.B.: A novel approach to MIMO transmission using a single RF front end. *IEEE J. Sel. Areas Commun.* **26**(6), 972–980 (2008)



An Open Integrated Electronic System Software Architecture Design for Launch Vehicle

Feng Zhang^(✉), Guo-wei Yao, Qian Wang, and Yun Xia

Research and Development Center, Academy of Launch Vehicle Technology,
China, No.1 Nandahongmen Road, FengTai District, Beijing, China
zhangfeng_3399@163.com

Abstract. In view of advantages and disadvantages analysis of software communication architecture (SCA), general open architecture (GOA) and avionics software system structure (ARINC653), a new type of integrated system software architecture which is suitable for launch vehicle was designed in this paper. According to the characteristics of the launch vehicle software, the new software system architecture including four layers, such as resource service layer, operating system layer, middleware layer, and application layer. Each layer is independent to each other, and all layers can integrate the functions of the traditional launch vehicle software system. The new software system architecture can support flexible functional expansion and tailoring, which enables it to quickly adapt to the requirements of different types of launch vehicles, and greatly reduce the repetitive design work.

Keywords: Launch vehicle · Integrated electronic system · Software architecture

1 Introduction

The discrete software structure was adopted in traditional launch vehicle's designing. According to specific task requirements, the control system, telemetry system, exterior measuring, and safety control system are constructed in parallel from top to bottom. This discrete structure has the advantages of simple interface and easy to analyse the cause of failure. But every subsystem was designed independently, the resources cannot be shared, which leads to problems such as redundancy of hardware resources, various kinds and large quantity of equipment. Especially in the design of different models of launch vehicles, it is necessary to redesign the hardware and software if the function is extended or cut.

In recent years, with the increasing of space missions, high-density launch and simultaneous development of multiple models launch vehicles have become the normal state of China's aerospace industry. In addition, emergency rescue in orbit, natural disasters and other emergency tasks, also need the launch vehicle to have the ability to respond quickly [1]. Although the space launch capability has been greatly improved in recent years, it is far from enough to meet the needs of high-density launch. Therefore, it is urgent to design an open integrated electronic system architecture [2].

The United States implemented the JTRS program to address the problem of frequent updating and high cost of radio stations, the software communication architecture [3] specification is proposed and continuously revised [4]. However, the architecture implementation is complex and the hierarchy from hardware to software is not clear enough. Under the background of increasing requirement of performance and function upgrade of avionics system, in 1996, the United States published the SAEAS4893, generic open architecture (GOA) framework, which stipulates the hierarchical division of the system hardware and software, as well as the type of interface. It is beneficial to reduce the cost, ensure the upgrades and extensibility of the system [5], but the reliability of the application function is not considered. To realize physical and functional integrate, the integrated modular avionics system is proposed, it not only faced with illegal access problems brought by the highly shared resources, but also must ensure the application software maintainability and portability [6]. To solve the above problems, ARINC653 standard is proposed, which defines the standard interface of application software from the perspective of single module. This standard uses the idea of spatial and temporal partition isolation to ensure the security of the system. However, all application functions operate independently in the partition and are directly provided service by the core operating system, which is not applicable in systems with high real-time requirements. Because these structures are not designed for high-speed, highly reliable spacecraft like the launch vehicle, they cannot be used directly in the launch vehicle.

An open integrated electronic system architecture was designed in this paper, which adopts the open hierarchical structure, the partition real-time operating system and middleware technology.

2 Traditional Software Architecture Analysis

2.1 Software Communication Architecture (SCA) Analysis

The Software Communication Architecture (SCA) [3] is a standard set of specifications for all software radio stations developed by the Joint Program Office (JPO) in the United States. The core idea of the SCA standard is that the system functionality is defined by software. Based on the same idea, China has also issued software communication architecture specification. The overall architecture of SCA is shown below [7] (Fig. 1).

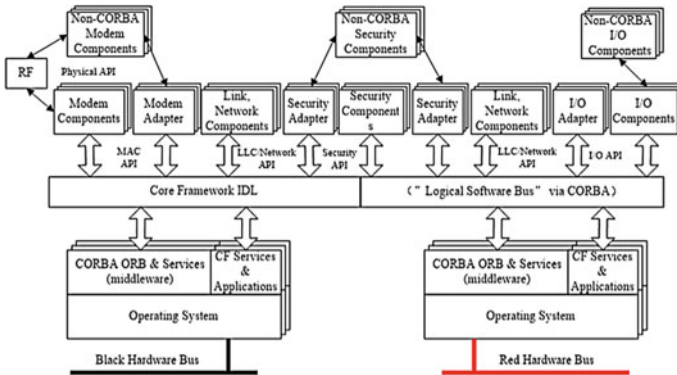


Fig. 1. The overall architecture of the SCA

The CORBA middleware and core framework, which are designed above the operating system, are the two main features of SCA software architecture. The interface between them and the operating system and the service interfaces they provided for the upper layer application are shown below [7] (Fig. 2).

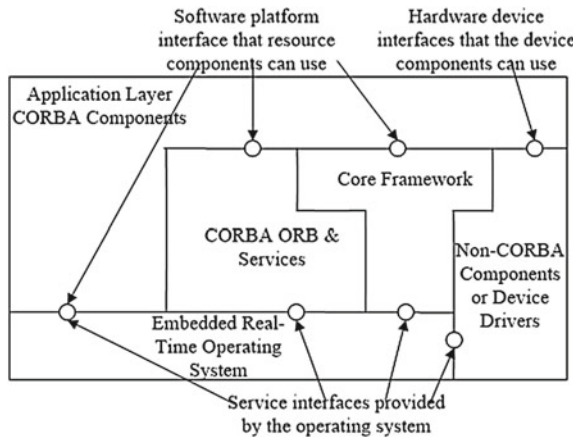


Fig. 2. Interface between CORBA components and underlying support software

As can be seen from the figures

1. The real-time operating system provides basic services for CORBA components, core framework, NON-CORBA components, or device drivers.
2. The upper layer waveform application software can obtain the basic services by calling the API provided by the embedded real-time operating system.
3. The upper layer waveform application software can realize messaging by calling CORBA components.

4. The core framework can provide the upper layer waveform application software with abstraction of hardware and low-level software, the file access services, and management services, such as configuration, installation and unloading.
5. For NON-CORBA components, the upper layer waveform application software can access the API provided by the NON-CORBA component adapter.
6. The upper layer waveform application software can directly call the API provided by the device driver.

The advantage of the SCA architecture is that the waveform is defined by software to achieve different radio functions, so the software associated with waveform generation is included in the application layer. Accordingly, the SCA software architecture hierarchical division is not in accordance with the processing order of sending or receiving a waveform, i.e., the order of RF transceiver, RF front end processing, signal processing, and information processing. This leads to the fact that SCA architecture is not clearly divided from the bottom to the top in terms of hardware to software. Because of this, the upper layer waveform application software can invoke the API provided by the device driver and interact with the hardware driver directly. In addition, the embedded real-time operating system, CORBA components, core framework and the NON-CORBA component or device drivers can provide API for upper layer waveform application software, the interface is scattered and not conducive to standardization.

2.2 General Open Architecture (GOA) Analysis

SAE AS4893 is an important open system structure standard. The main features of GOA are hierarchical division and interface classification. The architecture diagram is shown below [8] (Fig. 3).

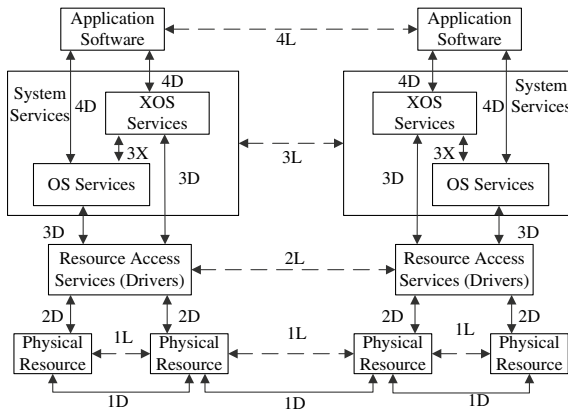


Fig. 3. GOA architecture diagram

The physical resource layer, resource access layer, system service layer and application software layer are included in the GOA architecture. According to the description of each layer in the standard, the hierarchy of architecture is divided according to the order of physical resources, underlying software, operating system and top application software. The hierarchy is relatively clear. It is conducive to the isolation from hardware to software, enabling the system to be more open and easy to be upgraded and expanded.

The extended operating system in the system service layer is other components outside the operating system, which allows the abstraction of the public needs of different applications. It can be understood that when some software components are invoked by multiple applications, they can be extracted as part of the extended operating system. In fact, compared to the ASAAC standard, the general system management (GSM) in the ASAAC is equivalent to the XOS in the GOA [9], so the system's parameter configuration function can be used as a XOS service. This is beneficial to the standardization and modular design of the software and avoids the repeated development of the same processing function. At the same time, the extracted software components can be repeatedly verified, and the reliability is improved.

However, the applications, which located above system service layer, includes multiple functions, they are running on the same platform and sharing system resources, resource competition or illegal access may occur in the working process. Therefore, it is not suitable for important functions that require high real-time performance or systems require high operational reliability.

2.3 ARINC653 Software System Architecture Analysis

ARINC653 software architecture is an application interface standard for the new generation of aircraft data integration requirements. Its main advantage is to use space and time partition isolation to achieve high reliability and security of software [10].

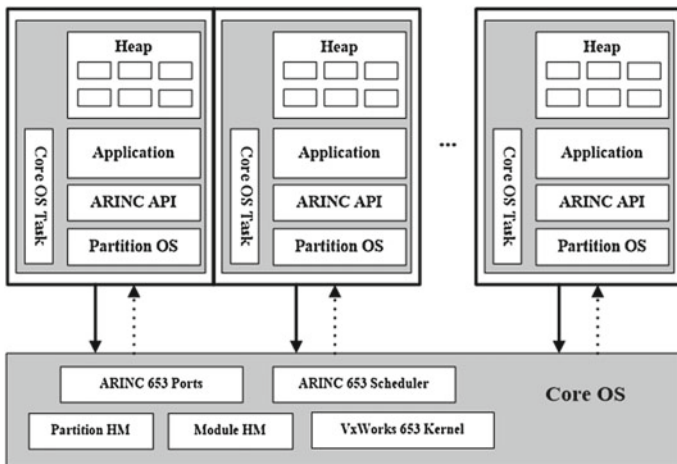


Fig. 4. ARINC653 software architecture

A generic APEX interface is defined between the core operating system layer and the application layer (Fig. 4).

The application software can be decoupled by means of partition isolation, so that each partition has no influence on each other. When one partition fails, other partition applications on the processor can still run normally, which improves the reliability of software operation. Due to partition isolation, each partition program can only be run in the allocated time, so as to avoid the resource competition question, and partition program cannot access other partitions address, which improved the security of the software.

However, in the ARINC653 architecture, all functional applications are placed on top of the core processing software, partitioned, and run-on schedule. This will reduce the real-time performance of application software for handling interrupts, and for shareable software components in a partition, when multiple applications need to call them, they need to communicate across partitions, which reduces processing efficiency. Therefore, it is necessary to balance security and real time in software system design. If partition isolation is needed, note that processing with high real-time requirements is not recommended as an application in a partition.

3 Requirements and Principles for Overall Architecture Design of Integrated Electronic System Software on Launch Vehicle

The main characteristics of electronic system on launch vehicle can be summarized as follows:

1. The electronic system on launch vehicle is short in working time and has high real-time requirement.
2. High reliability requirement for single mission.
3. The working parameters of wireless subsystems such as space-based telemetry, ground-based telemetry, satellite navigation, and safety control are relatively fixed, and there is no need to redefine them in the task phase.
4. For launch vehicles, the number of launches is limited, the mission time is short and the amount of test data obtained is small. Therefore, the high reliability of the system should be ensured in design stage.

The middleware and core framework used in SCA software architecture build a middle layer on top of the operating system. The middle layer provides messaging service, service interfaces, and control over the top layer applications. This idea can solve the communication and management problems when multiple systems are integrated.

The hierarchy in GOA architecture is very clear, which can separate hardware from software, applications from operating systems. The architecture is good for system upgrade and expansion. The extended operating system proposed in GOA architecture is also worthy of reference. The idea is to extract shareable software components from the application layer, make components that belong to management or underlying services part of the XOS, which can simplify software development.

The idea of partition isolation in ARINC653 software architecture allows partitions to be independent of each other. After multiple systems are integrated, it can be used to ensure the security and reliability of software.

Through the above analysis, it can be concluded that the design of integrated electronic system on launch vehicle should be based on the following principles:

1. The architecture is layered, the layers are isolated downward, and the interfaces between layers should be standardized.
2. The time and space partition isolation technology are used to improve the security and reliability of application at run time.
3. The middleware layer is used to isolate the operating system from the application software to improve the software portability.
4. The principles of middleware layer design are: it is convenient for application layer software development and management of application layer software, and the shareable service software components in the application layer can be classified into the middleware layer.

4 Software Architecture Design of Integrated Electronic System on Launch Vehicle

4.1 Overall Architecture Design

The designed software architecture of integrated electronic system on launch vehicle is shown below. The architecture consists of four layers, from top to bottom: application layer, middleware layer, operating system layer, and resource service layer (Fig. 5).

4.2 Resource Service Layer Design

Resource service layer interact directly with hardware. It uses hardware driver to convert the operating system calls into direct access or control of the hardware, provides hardware device abstraction interfaces to the operating system layer.

This layer contains core drivers and architecture support drivers. The core drivers complete the initial configuration of CPU and related minimum system hardware, and complete the configuration, initialization, boot, and operation of the operating system as well. The architecture support drivers are primarily composed of other hardware drivers, including drivers for serial port, Flash, developing and debugging devices, and interface software for adapting the operating system to different CPUs.

4.3 Operating System Layer Design

In this paper, an embedded real-time operating system with partition support is adopted. Its functions mainly include two aspects, one is partition management, the other is to provide common services.

Partition management mainly include the following functions: manage the partition time and space, ensure that the partition runs in the allocated time and address space;

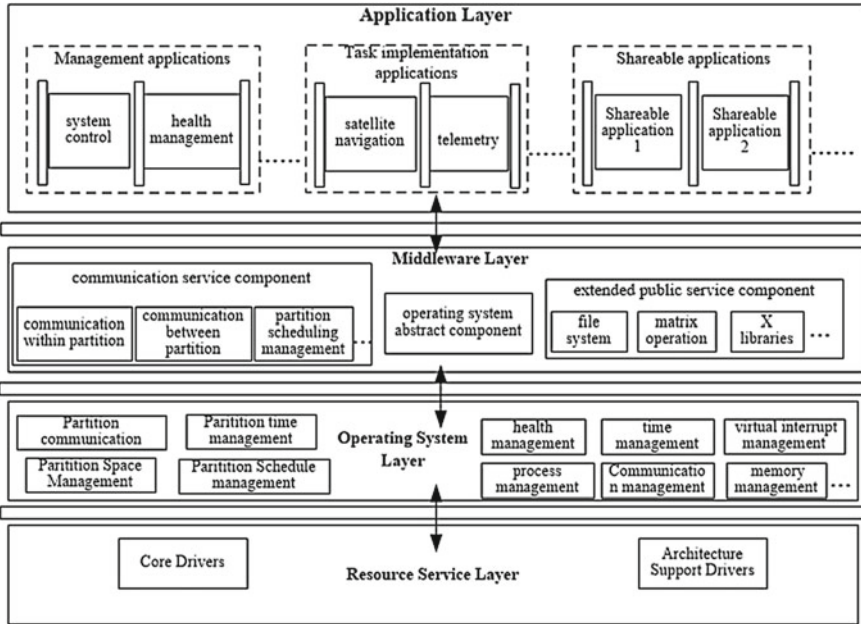


Fig. 5. Software architecture of integrated electronic system on launch vehicle

provide communication services within and between partitions; manage partition schedule, support schedule definition, and schedule switching, etc.

Common services mainly include the following functions: process management, time management, virtual interrupt management, memory management, health management, etc.

4.4 Middleware Layer Design

The middleware layer is the key in software design. It includes three components: operating system abstract component, communication service component, and extended public service component.

1. Operating system abstract component.

The operating system abstract component calls the operating system layer APIs, encapsulates them, and then provides service to the application layer with the encapsulated APIs. The operating system abstraction component isolates the application layer from the operating system layer.

2. Communication service component.

The communication service component provides a variety of ways for communication within each partition of the application layer, such as blackboard, signal volume, etc. And it provides multiple ways for communication between partitions, such as shared memory, message queues, etc. Finally, it also provides partition scheduling

management service, using schedule to plan the running order and time of each partition, and when a partition fails, the schedule can be switched.

3. Extended public service component.

Extended public service component provides the application layer with interfaces that the operating system does not have and can simplify the application layer software development. It includes such as file system service, matrix operation service, etc. This component can be selected and configured according to specific requirements.

4.5 Application Layer Design

Time and space are isolated between partitions in the application layer. This layer consists of three types of applications: management applications, task implementation applications, and shareable applications. Each application is located on a partition of the application layer, and the key application software can be backed up in other partition of the same hardware module or different hardware module.

1. Task implementation applications.

This kind of application is a series of digital signal processing software, which can be used to realize the integrated functions. For the integrated electronic system on launch vehicle, it includes flight control, telemetry, satellite navigation, safety control and other functions. The digital signal processing software includes framing, navigation message processing, filter processing, guidance rate calculation, etc.

2. Shareable applications.

This kind of applications refers to software components that can be shared in task implementation applications. It can be obtained by combing and clustering the processes of functions such as telemetry, satellite navigation and flight control etc.

3. Management applications.

Management application is used to complete the system-level management function of integrated electronic system on launch vehicle. Its functions include system control, health management etc. Enable the system to have the system-level health status monitoring capability, and the graceful degradation capability in fault state.

5 Advantage Analysis of the Software Architecture Design

- Based on the idea of hierarchical design, the software architecture of launch vehicle electronic system is designed into four layers. It can isolate hardware from upper software, and isolate applications from operating system. The isolation between software layers facilitates software upgrade, maintenance, and migration.
- Based on heterogeneous platform universal middleware design idea, set up middleware layer between operating system layer and application layer. It can separate the top layer application from the underlying software and hardware, as well as

provide messaging and service interfaces for the top layer application. This can be a good solution to the multisystem integrated management problems.

- The extended service component provided by the middleware layer can be composed of two parts: the business component; the software components in the application layer that are invoked by multiple functions and have certain real-time requirements can be extracted as middleware components. Such components can improve software processing efficiency.
- With the embedded operating system that supports partition isolation, the partition configuration, operation, management and scheduling of different applications are realized. Meanwhile, without additional hardware, the application software can be multiple redundant backup in different partitions. It satisfies the demand of high reliable operation of electronic system on launch vehicle.

6 Conclusions

In this paper, the software architecture of integrated electronic system is designed based on the open and hierarchical principle. On the basis of realizing the integrate of electronic system, it satisfies the demand of universalization and high reliability of electronic system on launch vehicle. The software architecture can effectively reduce the weight, volume, and power consumption of the electronic system and improve the overall performance of China's future spacecraft.

References

1. Chen, H.P., Luo, H., Liu, Y., et al.: Analysis and enlightenment on development trend of launch vehicle abroad. *J. Aerosp. Ind. Manag.* **11**, 34–37 (2015)
2. Song, Z.Y.: Research on avionics system configuration of next generation launch vehicle. *J. Manned Spacefl.* **22**(3), 317–322 (2016)
3. Wang, Y.: Design and Development of Frequency-Hopping Communication Waveform Based of SCA, D. Hangzhou Dianzi University, Hangzhou (2010)
4. Cai Z., Zhang X.Q.: Overview of SCA 4.0 specification, *J. Commun. Technol.* **46**(7), 126–128 (2013)
5. Huang, Y.K., Xue, Q.H., Li, W.M.: A study on open systems architecture standards. *J. Avion. Technol.* **36**(1), 34–41 (2005)
6. Cui, X.N., Hu, L.P., Ye, H., et al.: Research on software interfaces of integrated avionics system. *J. Comput. Sci.* **38**(2), 122–126 (2011)
7. GJB: Software communications architecture for tactical radio systems, 5028–2004
8. SAE AS4893: Generic open architecture (GOA) framework. USA (1996)
9. Huang, Y.K., Wu, J.M., Gu, T., et al.: Analysis and research on SAE AS4893 general open architecture (GOA) framework. *J. Avion. Technol.* **38**(1), 40–46 (2007)
10. Pu, X.-B.: Modern Avionics System and Integration, 1st edn. Aviation Industry Press, Beijing (2013)



Read and Write Performance Research and Optimization for eMMC Device Driver

Yanlin Chen, Songyan Liu^(✉), Yifei Niu, Huan Liu,
and Xiaowen Wang

Electronic Engineering College, Heilongjiang University, 150080 Harbin, China
390607441@qq.com, liusongyan@hlju.edu.cn, {niu_yifei,
liuhuan920913, xiaowengoing}@163.com

Abstract. The eMMC (embedded multimedia card) device was originally applied for high storage devices on mobile devices by a driver provided a good internal programming interface to the hardware. This article analyzes the read and write characteristics of eMMC devices in high-speed mode. It is proposed to improve the read and write performance of eMMC devices by using frequency replacement algorithm and double buffer as the basic improvement scheme.

Keywords: Device driver · eMMC · Operating system · Interface protocol

1 Introduction

The eMMC device integrates a NAND flash controller, a NAND Flash array and an eMMC interface on a single chip [1], which ensures high-performance communication between the host controller and the eMMC memory card. And multiple eMMC storage cards can be combined into large-scale storage arrays to facilitate the storage of large-capacity data [2]. This article proposes the use of scheduling algorithms, DMA technology and replacement of FIFO strategies in the eMMC high-speed mode to solve the problems of storage volatile, storage management and high bit error rate, providing practical significance for the development of high-speed storage device driver.

2 eMMC Device Driver Read and Write Mode

The eMMC device supports multiple read/write modes, such as high-speed SDR, HS200, and HS400 modes. The HS400 mode has a maximum transfer speed of 400 MB/s. When the eMMC device reaches the transfer mode, the block size (represented by *size_of_block*) of each write is set by command, and the number of blocks of the current transfer packet is calculated based on the start address (represented by *add_start*) and the end address (represented by *add_end*) as shown in Eq. (1).

$$num_of_block = \frac{add_end - add_start}{size_of_block} \quad (1)$$

Then reset the FIFO queue and interrupt status, configure registers by the calculated *num_of_block*, place current task status in the data transfer state. If there is a DMA request [3], the DMA descriptor is started by acquiring the physical address of the data map. Through the number of blocks calculated, the single block or multi-block operation is decided. For multiple block read operations, the host first sends a read operation command, after receiving the response correctly, it sends the data block with 16 detection bits through the data line [4]. If it is incorrect, re-read or enter the interrupt. If correct, the slave sends a stop command after waiting for two clock cycles. Multiple block write operations are similar to multiple block read operations, the multi-block read and write operation timing is as shown in Fig. 1.

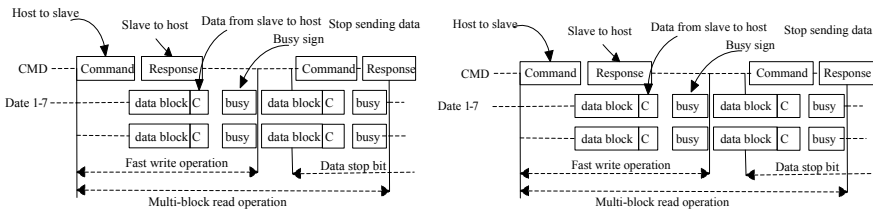


Fig. 1 Schematic diagram of sending data in multi-block read mode

3 The Optimization for Device Driver Performance

As the transmission speed increases, data transmission requires more efficient cache calls and storage strategies in the eMMC device. Double buffer and frequency-based page replacement algorithms are proposed to improve the reliability.

3.1 The Design of Double Buffers

The operating system usually reads and writes in units of 512 bytes, and NandFlash has a sector size of 512 bytes. Therefore, a basic unit of I/O transfer between the eMMC device driver and the device can allocate a buffer composed of a plurality of sectors to summarize the write request of small data blocks and then write it to the device.

In a single buffered scheme, the transmitted data is input to the system buffer, and the process will do a pre-read. Since the writing of data is logically continuous, it is expected that this data block will eventually be used, and only an unnecessary block will be read at the end of the processing sequence. When the data is ready to be sent to the device, coping it from the user space to the system buffer, and the requesting process can freely continue to execute or swap out if necessary. This method increases the storage speed of the device compared to the case without buffering, and the user program can process data block that has been read or written while the next block is being read.

In high-speed storage devices, data block write speed is fast and read speed is slow, which is easy to cause write data to be overwritten, so two system buffers can be allocated to improve this problem. That is, while a process is transferring data to and

from a buffer (the data is fetched from this buffer), the operating system is emptying (or filling) another buffer. We can roughly estimate the execution time as $\max [C, T]$, where C is the computation time between input requests and T is the time required to input a block of data. Therefore, if $C \leq T$, the block-oriented device can be run at full speed. Device buffering scheme is as shown in Fig. 2.

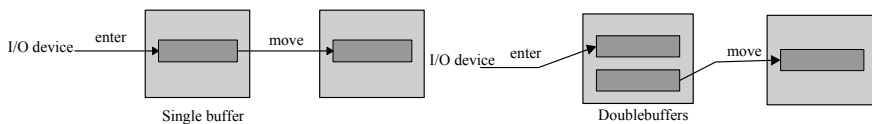


Fig. 2 Device Buffering Scheme

3.2 Improvement Based on Frequency Substitution Algorithm

When a new block is read into the NandFlash storage array, an existing block must be replaced for constrained cache space. Therefore, it is necessary to design a high-performance page replacement algorithm in the eMMC devices.

The LFU algorithm wastes the buffer space due to locality. So a frequency-based permutation algorithm is proposed in [ROBI90]. As shown in Fig. 3, the blocks are logically grouped into a stack with a new area and an old area. When a cache hits, the block being accessed is moved to the top of the stack. If the block is already in the new zone, its time accessing counter does not increase, otherwise the counter increases by 1 once as a hit, the block with the lowest access counter and no access to the new zone is swapped out. The method is possible to replace a block that is relatively frequently accessed because there is not a long enough interval to establish the number of accesses.

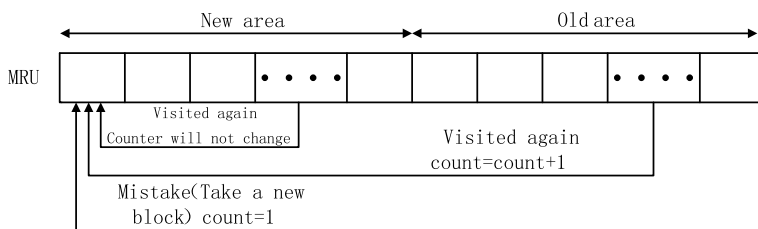


Fig. 3 Frequency replacement algorithm

Therefore, a further improvement is to divide the stack into three zones: a new area, a middle area, and an old area. Assuming that there is a large enough intermediate area, the blocks that are accessed relatively frequently are reduced in time for the data block to be read back into memory due to a miss. Since only the blocks in the old zone meet the replacement conditions, there is an opportunity to increase their access counters before they become eligible for the replacement. Because of the higher hit rate, less time for allocating a new buffer, the performance of write is correspondingly improved.

4 Conclusion

This paper analyzes the read and write operations in high-speed mode, and proposes dual buffers between the operating system and the driver to meet the requirement of the read and write performance for large-capacity storage devices. By improving the frequency-based permutation algorithm between the driver and the NandFlash storage array, the problem of mismatch between write and read speeds is effectively improved, and the read and write performance of the eMMC device is improved overall.

References

1. Park, S.H., et al.: A pattern adaptive NAND flash memory storage structure. *IEEE Trans. Comput.* **61**(1), 134–138 (2011)
2. Kim, C., Lee, C.: Design of eMMC controller with multiple channels. In: *SOC Design Conference IEEE*, pp. 317–318 (2016)
3. Aljumah, A., Altaf, M.: Amba based advanced DMA controller for SoC. *Int. J. Adv. Comput. Sci. Appl.* **7**(3) (2016)
4. Mielke, N., et al.: Bit error rate in NAND flash memories. *Reliability Physics Symposium, 2008. IRPS 2008. IEEE International IEEE*, pp. 9–19 (2008)



Power Analysis Method Based on DC Bus Voltage Waveform

Yicheng Wang^(✉), Zhiyong Huang, Xiaolong Luo,
and Yinguo Huang

State Key Laboratory of Precision Measuring Technology and Instruments,
Tianjin University, Tianjin 300072, China
17695798266@163.com, hyg@tju.edu.cn

Abstract. This essay analyzes the measurement of the power factor in the drive motor system of the frequency converter. Due to the fact that it is inconvenient to utilize the power analyzer to acquire the motor power factor directly. A method is proposed. We can collect the DC bus signal waveform. Then we could reconstruct motor stator three-phase voltage and current to test indirectly the motor power factor. In addition, there are some simulations and data collected from experimental device in order to verifying accuracy. Congruously, the experimental results show that within a certain error range, the motor power factor can be measured by voltage and current reconstruction.

Keywords: DC bus voltage · Phase voltage reconstruction · Phase current reconstruction · Power analysis · Power factor

1 Introduction

Power factor is an important parameter of the power system. The traditional power factor measurement is usually the use of AC voltage and current sensors on the stator side. Then we could measure the three-phase voltage and current on the stator side. We could obtain power factor by power analyzer. However, the price of the power analyzer and AC sensors is high. This method inevitably increases the test cost. Moreover, in many cases, the accuracy requirements are not very strict and environment is complex. Consequently, the use of power analyzers is not an appropriate for power factor measurement.

This essay will introduce a new method which could indirectly measure the motor power factor. It can be considered by collecting voltage and current signals from DC bus, and then using the data analysis reconstruction algorithm to reconstruct the voltage and current of the stator phase. And we could calculate motor power factor. The main benefits of this method are that it can dramatically decline measurement costs and data analysis can be easily achieved in the upper computer, improving the convenience of measurement.

2 Principle

The typical voltage source PWM inverters is widely used in the frequency converters. Its basic structure is shown in Fig. 1. The basic structure consists of three pairs of power devices. The basic working principle of this system is as follows: The state of the switch tubes on the three-phase upper arm is defined as S_a , S_b and S_c . Furthermore, state is defined as “1” when it is turned on, and it is “0” when it is turned off [1]. When the state of the switch is 100, the state of the upper three power tube is on, off and off. The current of the DC bus is equal to the A-phase of the output. When the state of the switch is 010, the B-phase current of the output is equal to the bus current. When the upper arms are all on or off, the output vector is 000 or 111. In this case, Bus current does not indicate any phase current.

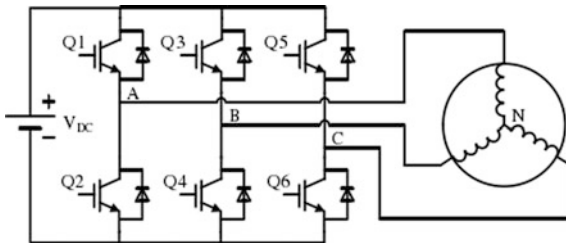


Fig. 1. Three-phase voltage source inverter with load

The on-off state of the upper and lower power devices is opposite. This means that there are eight kinds of working states which can form 8 spatial voltage vectors. Six of them are nonzero space voltage vectors that including $U_0(100)$, $U_{60}(110)$, $U_{120}(010)$, $U_{180}(011)$, $U_{240}(001)$, $U_{300}(101)$. The remaining two zero vectors are $U_{000}(000)$ and $U_{111}(111)$ respectively. The space vector graphic is divided into six sectors by these eight working conditions [2].

The on-off state of power tube is usually controlled by PWM wave. SPWM is used more often in daily utilization. The width of each pulse and the distance between them are determined by the intersection of sine wave and triangle wave [3].

However, SPWM output voltage has some harmonic component, leading to large switching loss and dead zone issues. Therefore, the SVPWM technology is developed to alleviate this problem. Compared with SPWM, SVPWM has more advantages. When the voltage vector is output, the flux path of the motor is close to the circle and it is easy to achieve [4].

In the practical application of SVPWM, the inner tangent hexagonal pattern ($n = 6$) is usually used. As shown in Fig. 2, each side is still composed of vectors. Each sector is modulated by two nonzero vectors and zero vectors. Symmetric modulation mode is often used. In one cycle, the operating time of each voltage component is divided into two sections. The corresponding phase currents of the DC bus at different switching times can be obtained from Table 1. Through observing Fig. 3, in the second sector, the data of the i_b and i_c phase currents can be obtained by sampling and reconstruction.

And i_a can be calculated according to formula (4). In the same way, this conclusion can be acquired in other sectors.

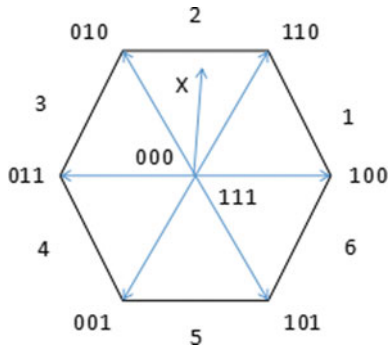


Fig. 2. Inverter output space vector

Table 1. DC bus current corresponding phase current at different switching times

Section	Voltage vector	Corresponding phase current
1	$V_1(100)$	$i = i_a$
	$V_2(110)$	$i = -i_c$
2	$V_2(110)$	$i = -i_c$
	$V_3(010)$	$i = i_b$
3	$V_3(010)$	$i = i_b$
	$V_4(011)$	$i = -i_a$
4	$V_4(011)$	$i = -i_a$
	$V_5(001)$	$i = i_c$
5	$V_5(001)$	$i = i_c$
	$V_6(101)$	$i = -i_b$
6	$V_6(101)$	$i = -i_b$
	$V_1(100)$	$i = i_a$

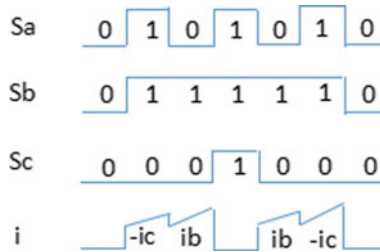


Fig. 3. SVPWM waveforms in sector 2

However, in the actual system, it is difficult to collect accurate data of phase current. The minimum operating time T_{min} of a nonzero voltage space vector must satisfy formula (1).

$$T_{min} = T_{dead} + T_{conv} + T_{sett} \tag{1}$$

T_{dead} , T_{conv} , and T_{sett} , respectively indicates inverter dead time, current sampling A/D conversion time and current setting time. T_{sett} can be also explained as the stable settling time of the DC bus current [5].

After Fig. 4 shows that the switch state changes, there is a current setting time, and there are current sampling time and an inverter dead zone. For sampling time is too short, lead to cannot complete reconstruction. We can through the translation pulse method and vector insertion method to solve the problem.

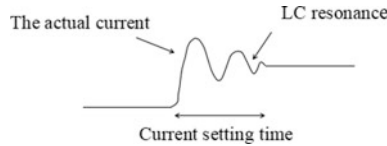


Fig. 4. Schematic diagram of bus current after switching state changes

Set r_s for motor each phase of the stator resistance. The inductance is L . The counter electromotive force is e . And the motor midpoint is n . The median points of three pairs of power devices a , b , and c respectively. The DC capacitance midpoint for o , then you can get the formula (2) [6].

$$\begin{cases} U_{an} = U_{ao} - U_{no} = i_a r_s + L_s \frac{di_a}{dt} + e_a \\ U_{bn} = U_{bo} - U_{no} = i_b r_s + L_s \frac{di_b}{dt} + e_b \\ U_{cn} = U_{co} - U_{no} = i_c r_s + L_s \frac{di_c}{dt} + e_c \end{cases} \tag{2}$$

Add the two sides of formula (2) separately.

$$U_{ao} + U_{bo} + U_{co} - 3U_{no} = 3(i_a + i_b + i_c)r_s + L_s \frac{d(i_a + i_b + i_c)}{dt} + e_a + e_b + e_c \tag{3}$$

Under the condition that the three-phase system has no center line and the three-phase balance of the reverse electromotive force. We can be obtained formulas (4) and (5) [7].

$$i_a + i_b + i_c = 0 \tag{4}$$

$$e_a + e_b + e_c = 0 \tag{5}$$

So, we can get the formula (6).

$$\begin{cases} U_{an} = \frac{2}{3}U_{ao} - \frac{1}{3}U_{bo} - \frac{1}{3}U_{co} \\ U_{bn} = \frac{2}{3}U_{bo} - \frac{1}{3}U_{co} - \frac{1}{3}U_{ao} \\ U_{cn} = \frac{2}{3}U_{co} - \frac{1}{3}U_{ao} - \frac{1}{3}U_{bo} \end{cases} \quad (6)$$

We can find that phase voltage U_{an} , U_{bn} , U_{cn} could be expressed by U_{ao} , U_{bo} , U_{co} . The U_{ao} , U_{bo} and U_{co} are expressed as bus voltage and three switch tube function. We use T_1 , T_2 , and T_3 to represent the instantaneous phase voltage duty ratio given by the three power tubes of the upper half bridge within a period T_s . So here we can get the formula (7).

$$\begin{cases} U_{an} = V_d \left[\frac{2}{3}T_1 - \frac{1}{3}T_2 - \frac{1}{3}T_3 \right] \\ U_{bn} = V_d \left[\frac{2}{3}T_2 - \frac{1}{3}T_1 - \frac{1}{3}T_3 \right] \\ U_{cn} = V_d \left[\frac{2}{3}T_3 - \frac{1}{3}T_1 - \frac{1}{3}T_2 \right] \end{cases} \quad (7)$$

According to formula (7), it can be found that for each switching cycle, the output three-phase voltage can be obtained as long as the duty ratio of each switching circuit is obtained.

The proposed method reconstructs the output three-phase voltage by using the detected DC-link voltage and the output three-phase current reference [8].

When the power system is running, the power grid needs to provide two parts of electricity: active power for one-way conversion to other energy and reactive power for electromagnetic energy exchange [9]. The measurement of both electric power and energy, especially in no sinusoidal conditions, is still an open research problem in the metrology community [10]. Many researches have been made on the definition and calculation of reactive power at home and abroad, and they have been put forward corresponding theories. The traditional power definition method is based on sinusoidal A-C circuit. In a single-phase sinusoidal circuit.

Apparent power

$$S = UI \quad (10)$$

Active power

$$P = UI \cos \varphi \quad (11)$$

Reactive power

$$Q = \sqrt{S^2 - P^2} = UI \sin \varphi \quad (12)$$

The φ is the impedance angle. Reactive power is considered to be the energy exchange between power and load [11]. The $\cos \varphi$ is the value calculated in the article. In the motor system, the motor is equivalent to a large inductor. There only exists inductive reactive power.

The accurately and fast estimation of Phase Difference (PD) is required between the voltage and current of an AC electrical power system to calculate the Power Factor

(PF) for defining how effectively the electrical energy is converted into the useful form [12].

The voltage and current signal reconstructed from the DC bus. The voltage and current phase difference are calculated by the upper computer. In the motor system, the current hysteresis voltage signal is a phase difference φ . And then we can get the power factor.

3 Simulation and Experiment

In matlab/simulink, the input and output of the inverter are simulated. We assume the drive signal is SVPWM wave, and the current waveform of the bus in Fig. 6 and the corresponding driving signal waveform at the corresponding time can be obtained. As can be seen from Fig. 5, the bus current does contain all the information of output three-phase current [13].

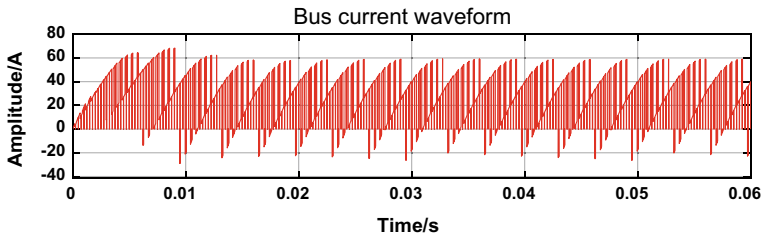


Fig. 5. Bus current waveform

To observe the waveform of Fig. 6, we can find that the figure of SVPWM waveform in five sectors and six sectors. So, we need sampling the signal in the three-vector corresponding to the DC bus waveform that we can get three-phase current waveform. We can get the three-phase current and voltage waveform shown in Figs. 7 and 8.

By analyzing Figs. 7 and 8, we can calculate the power factor of the stator end of the motor model directly at 0.809 and the error is 1.3%, which meets the test requirements.

The experiment waveform data is shown in Figs. 9 and 10.

Blue waveform is DC bus current waveform, red waveform is DC bus voltage signal. There are a great quantity of laboratory electromagnetic interference, so there is a big wave distortion. From the waveform, it can be found that the signal of each phase waveform is contained in the bus current (Figs. 11, 12 and 13)

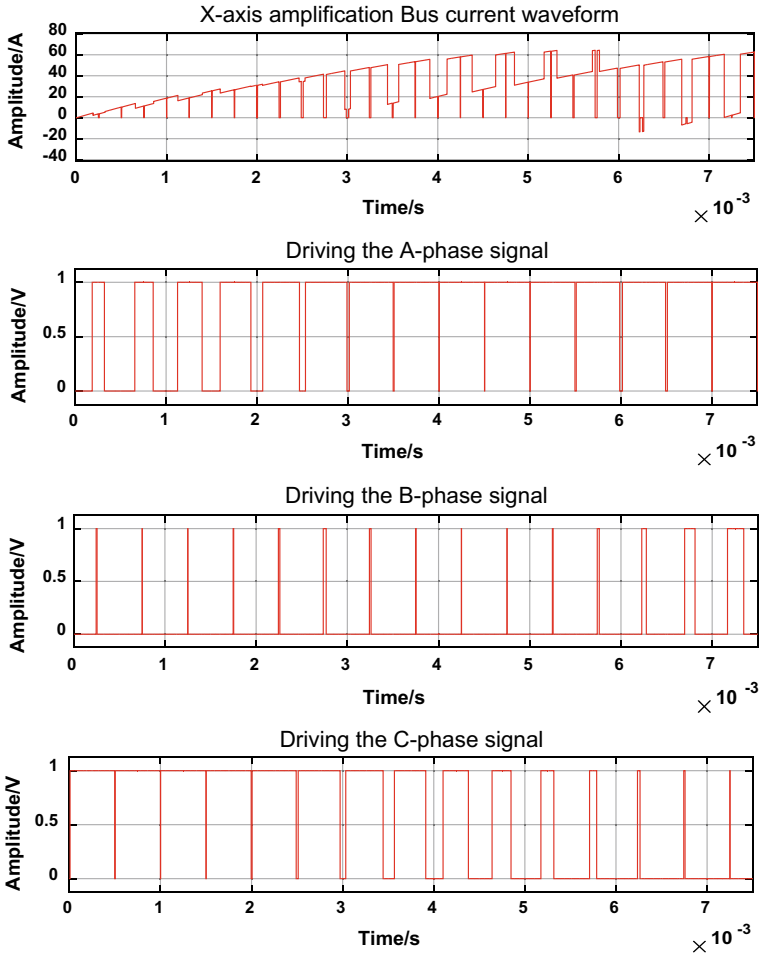


Fig. 6. Bus current corresponds to the drive waveform

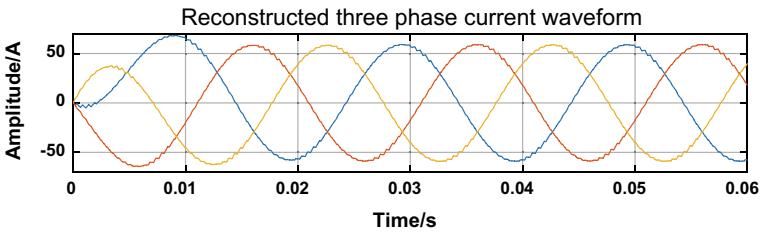


Fig. 7. Reconstructed three-phase current waveform

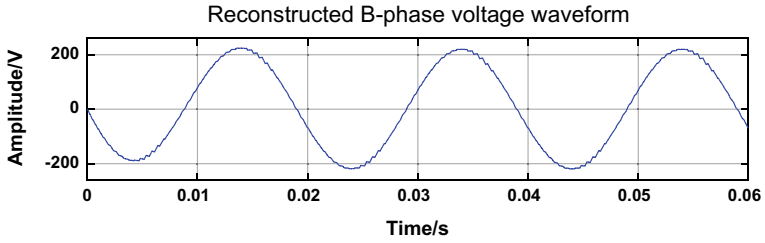


Fig. 8. Reconstructed B-phase voltage waveform

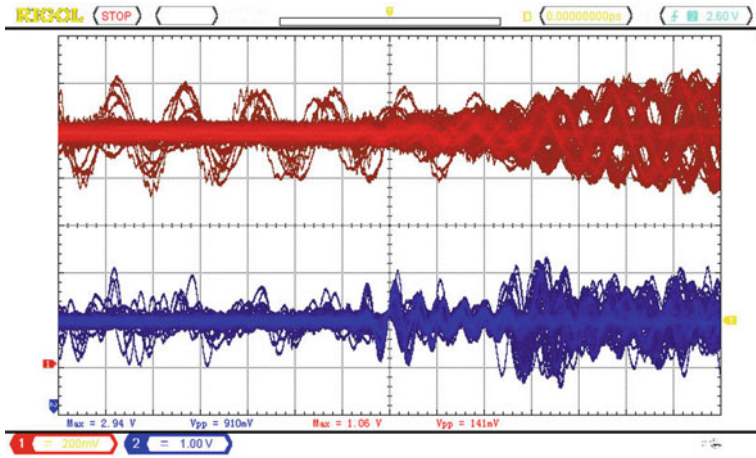


Fig. 9. Bus voltage and current afterglow

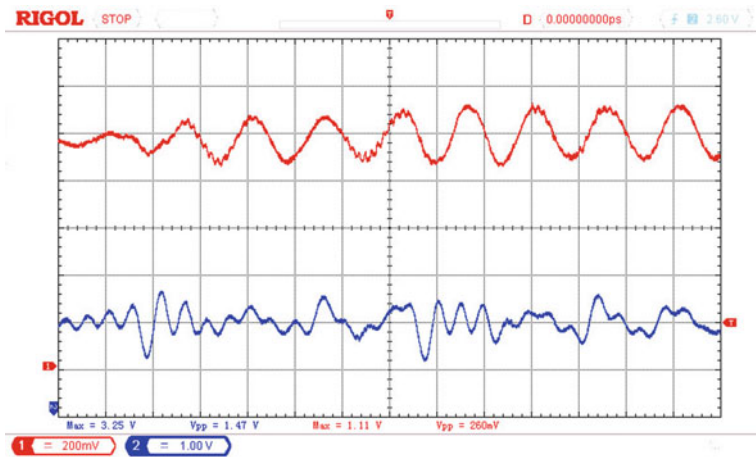


Fig. 10. Bus voltage and current waveform

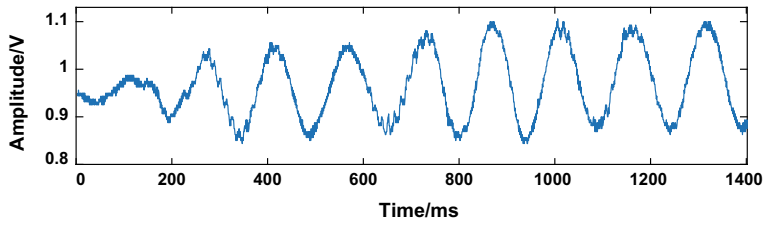


Fig. 11. Voltage waveform of bus at 500 rpm

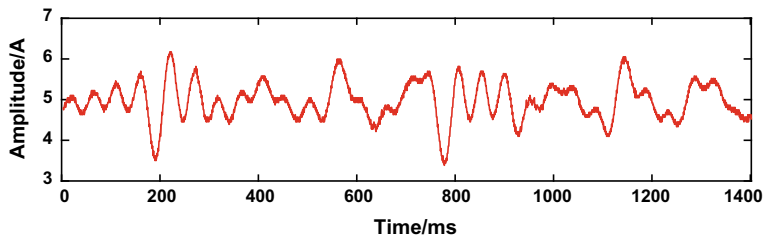


Fig. 12. Current waveform of bus at 500 rpm

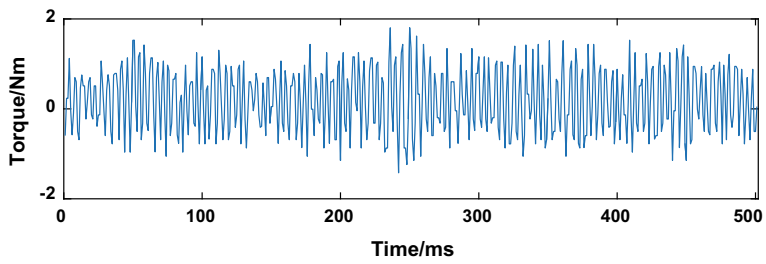


Fig. 13. Torque variation waveform at 500 rpm

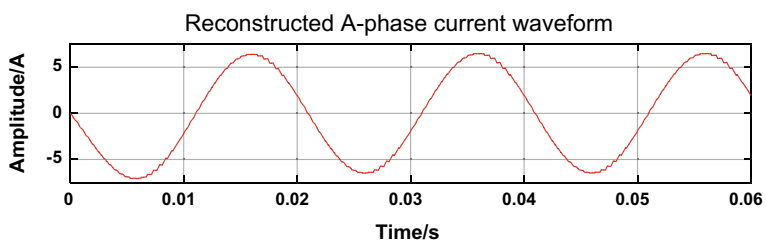


Fig. 14. Reconstructed A-phase current waveform at 500 rpm

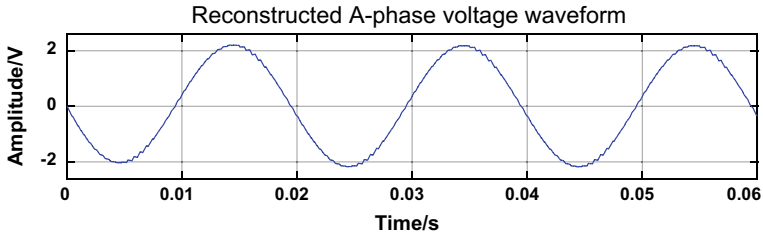


Fig. 15. Reconstructed A-phase voltage waveform at 500 rpm

After theoretical analysis, we can get the following reconstructed waveform.

From Figs. 14 and 15, we can get the power factor is 0.88, and the theoretical power factor is 0.89. After calculation, the error is 1.1%, which meets the test requirements (Figs. 16, 17, 18 and 19).

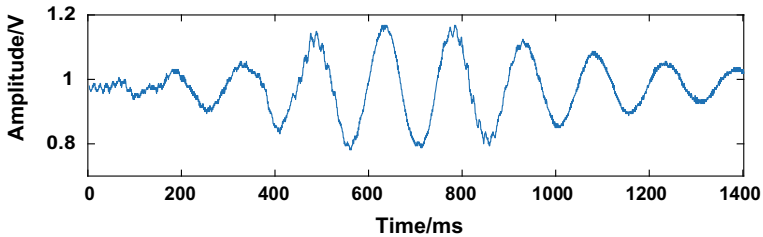


Fig. 16. Voltage waveform of bus at 1000 rpm

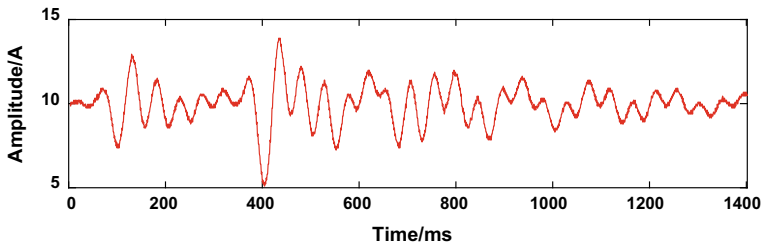


Fig. 17. Current waveform of bus at 1000 rpm

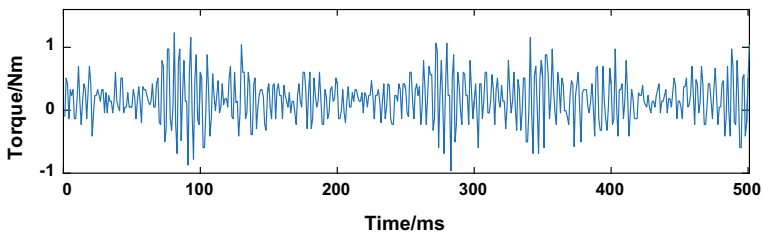


Fig. 18. Torque variation waveform at 1000 rpm

After theoretical analysis, we can get the following reconstructed waveform.

From Figs. 19 and 20, we can get the power factor is 0.906, and the theoretical power factor is 0.902. After calculation, the error is 0.5%, which meets the test requirements.

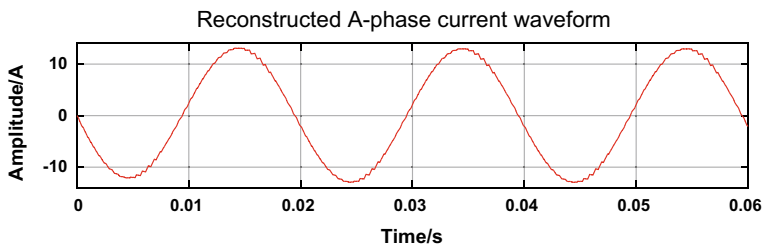


Fig. 19. Reconstructed A-phase current waveform at 1000 rpm

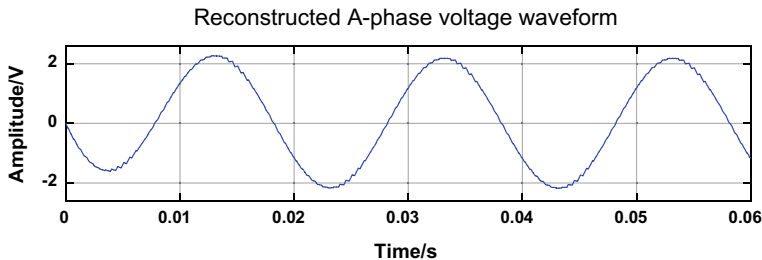


Fig. 20. Reconstructed A-phase voltage waveform at 1000 rpm

4 Conclusion

In order to detect the power factor of the motor, the power factor of the stator end of the motor should be considered. DC bus signal and corresponding SVPWM waveform signal are collected, and then the bus signal is sampled at the corresponding switching time. Simulation and experiment show that this method can reconstruct the voltage and current signal of stator terminal. By calculating the phase difference of voltage and current, we can get the power factor. The value are close to the actual value. However, in the process of reconstruction, there are waveform distortion, which is caused by sector switching and power tube switching. All of these can be solved by other algorithms, such as translation pulse method, vector insertion method, etc., which will be further studied in future experiments.

References

1. Cheng, M., Jiang, Y., Wang, Q.: PWM inverter phase current reconstruction strategy for line voltage Modulation. *J. Mot. Control* **22**(02), 9–16 (2018)
2. Chu, J., Hu, W., Huang, W., Yang, F.: A phase current sampling reconstruction technique for inverter. *Trans China Electrotech. Soc.* **25**(01), 111–117 (2010)
3. Zhang, Y.: *Application Technology of SPWM Frequency Conversion Speed Regulation*. China Machine Press, p. 12 (2011)
4. Zeng, Y.: *Principle, Algorithm and Application of SVPWM Technology of Variable Frequency Speed Regulation*. China Machine Press, p. 10 (2010)
5. Gu, Y., Ni, F., Yang, D., Dang, J., Liu, H.: Phase current reconstruction method based on bus current sensor. *J. Mot. Control* **13**(06), 811–816 (2009)
6. Zhang, J., Zhang, X., Chen, T., Peng, Y., Sun, Y., Tian, H.: Phase voltage reconstruction of synchronous motor and observation of magnetic chain. *Electr. Driv.* **43**(08), 29–32 (2013)
7. Sun, X., Shang, Y., Zhong, Y., Zhang, W.: Research on output voltage reconstruction technology of three-phase voltage source inverter. *J. Xi'an Univers. Technol.* **02**, 122–124 (2005)
8. Haga, H., Hara, Y., Kondo, S.: Method for Reconstructing Three-phase Voltage of Current Source Inverter via Detection of DC-link Voltage. Nagaoka University of Technology (2012)
9. Wang, W.: *Research on High Precision Reactive Power Measurement*. Southwest JiaoTong University (2007)
10. Fiorucci, E.: The measurement of actual apparent power and actual reactive power from the instantaneous power signals in single-phase and three-phase systems. *Electr. Power Syst. Res.* (2014)
11. Shi, Y.: *Calculation and Analysis of Reactive Power Under Power Electronic Load*. SCUT (South China University of Technology) (2011)
12. Sarbulut, L.: A simple power factor calculation for electrical power systems. *Electr. Power Energy Syst.* 06 (2011)
13. Li, Y.: *Research on Phase Current Reconstruction Technology of Permanent Magnet Synchronous Motor*. Harbin Institute of Technology (2014)



Automatic Rapid Electrokinetic Pattern System

Yanwei Wang^{1(✉)} and Jiaqi Zhen²

¹ College of Mechanical Engineering, Harbin Institute of Petroleum, Harbin, China

xianxinyue@163.com

² College of Electronic Engineering, Heilongjiang University, Harbin 150080, China

Abstract. Rapid electrokinetic patterning (REP) is an optoelectronic technique to manipulate particles. The number of particles can be adjusted by varying REP parameters such as frequency and voltage. An experimenter using REP may want more or fewer particles than are collected in a certain experiment. We use evolution of life curve to analysis REP system, and use nine screens to analysis system. Currently an experimenter may try to vary the experiment parameters to acquire the desired number of particles. This approach is cumbersome. Our goal is to set up an automatic REP system. It can help us to acquire the particles in an automated way. Therefore, we used subfield analysis to set up the automatic rapid electrokinetic pattern system.

Keywords: Particle image velocity · Rapid electrokinetic patterning · Manipulation particles

1 Background

In order to acquire the velocity of micro- or nanofluidic, the need for these measurement techniques has continued to grow since the development of the first micro flow characterization technique, microparticle image velocity (μ PIV), in 1998 [1]. Since then the number and variety of methods of manipulating micro and nanofluidic flows has exploded. It is now common to find microfluidic devices with several different forces acting on the flow that are not typically significant at macroscales. Surface tension gradients, electrophoretic effects, and electrothermal forces are all common drivers of microfluidic flows. Recently, there are many methods to evaluate microparticle image velocity. And there are many practice problems related with microfluid are found and some different kinds of microfluid models are designed to investigate its property. Thus, with the rapid development, μ PIV could be used in more wide flow field [2–4].

2 Micro-PIV Mathematical Model

Micro-PIV (particle image velocity) is a measurement technique which allows for capturing velocity information of whole flow fields in fractions of a second.

Figure 1 shows the setup for PIV recording. Small tracer particles are added to the flow. It is assumed that the tracer particles move with the local flow velocity between the two illuminations. The light scattered by the tracer particles is recorded via a high quality lens. With modern Charge Coupled Device (CCD) cameras (1000×1000 sensor elements and more), it is possible to capture more than 100 PIV recordings per minute. By comparing these particle images, we can extract the displacement information from a PIV recording.

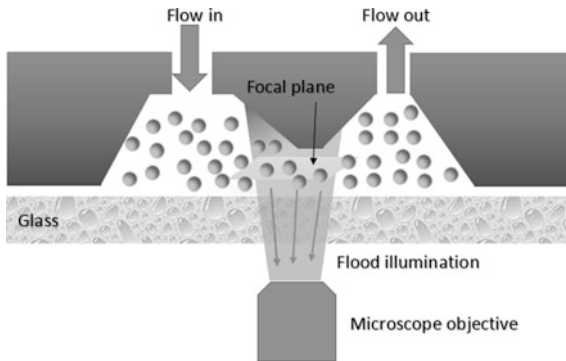


Fig. 1. Micro-PIV schematic

Our goal is to set up an automatic REP system. It can help us to acquire the particles in an automated way. The principles are the following, first, the CCD images the particles; second, the computer identifies the appropriate feedback signal; the computer controls the signal generator and amplifier. There are two images from different times. Comparing these two images, we can find the particle velocity. Cross correlation function k of two image intensity functions $f_k(i, j)$ and $g_k(i, j)$ computed for a correlation region measuring p by q pixels.

$$\varphi_k(m, n) = \sum_{j=1}^q \sum_{i=1}^p f_k(i, j) g_k(i + m, j + n) \quad (2.1)$$

3 Design Procedure

3.1 Description of Requirements

Figure 2 gives us the Causal Analysis. It is analysis the reason that we cannot control particle following the adjustment.

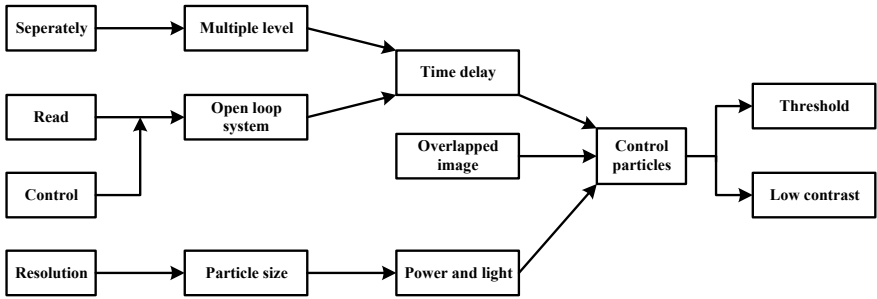


Fig. 2. Casual analysis

From Fig. 2 we can know that threshold and low contrast, both of them influence the result. From the top line, the main reason about control particle on time is separate unit, and the open loop system is another factor.

3.2 REP Equipment

In order to set up an automatic REP system, we build the Substance field model. It can help us to acquire the particles in an automated way. Thus we analysis our system and construct the Substance field model. Figure 3 shows the construction.

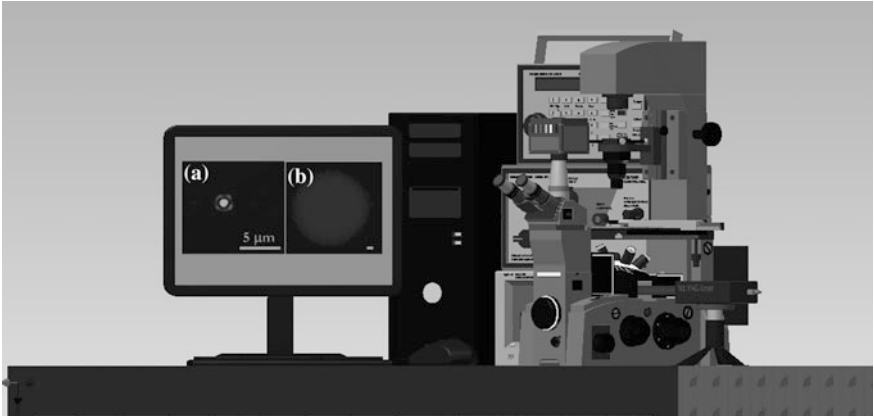


Fig. 3. Set up of REP equipment

With the addition of the electricity field, we use signal generator and ampler to adjust the power of electric. At the same time, the time delay is increased, on the other hand when we see the picture, we also need time to adjust the equipment, it also wastes

time. In order to minutes the time, we use computer to connection the particle and the electric equipment.

Therefore, we used subfield analysis to set up the Automatic Rapid Electrokinetic Pattern System in Fig. 4.

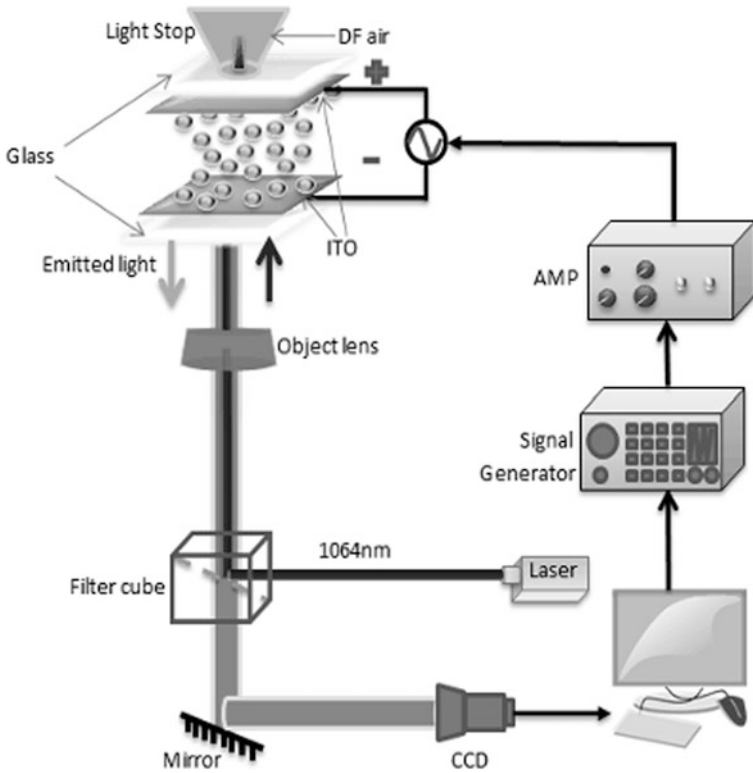


Fig. 4. Automatic rapid electrokinetic pattern system

3.3 Inventive Principles and the Matrix of Contradictions

Sometimes, the ratio is very low, it is difficult to distinct the background and the particle, Fig. 5 shows the example.

We use the neighbor fixed instead of the whole picture, and use center point to calculate the location and record each center points. The result is shown in Fig. 6.



Fig. 5. Particle image

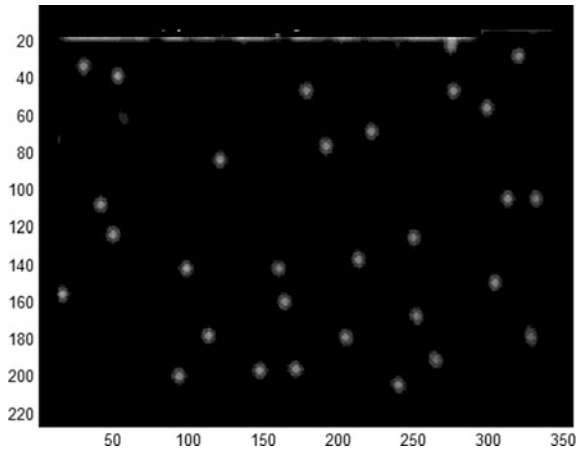


Fig. 6. Processed image

4 Conclusions

Automatic rapid electrokinetic pattern system is set up, and the laser and electric field is used to control particle motivation. Moreover, the particle is closed to the neighbor particles, and the contrast ratio between the particle and the background are similar, it is difficult to identify and counting particles. In this circumstance, we set up matrix of contradictions, and find out the particle and background contradiction factors, and use 40 innovation principles to build a band-pass filter which uses the highest and lowest gray pixels. It is valid for simulation particles and experiment particles. Especially some particles image with low contrast ration, noise, and short distances.

Acknowledgements. This work was supported by the National Natural Science Foundation of China under Grant No. 61501176, Heilongjiang Province Natural Science Foundation (F2018025), University Nursing Program for Young Scholars with Creative Talents in Heilongjiang Province (UNPYSCT-2016017), The postdoctoral scientific research developmental fund of Heilongjiang Province in 2017 (LBH-Q17149).

References

1. Santiago, J.G., Wereley, S.T., Meinhart, C.D., et al.: A particle image velocimetry system for microfluidics. *Exp. Fluids* **25**, 316–319 (1998)
2. Wang, Y.W., Craig, S., Pan, X.D., et al.: Micro/nano flow characterization techniques. *Encycl. Nanotechnol.* **8**, 1–11 (2015)
3. Williams, S.J., Park, C., Wereley, S.T.: Advances and applications on microfluidic velocimetry techniques. *Microfluid. Nanofluid.* **8**, 709–726 (2010)
4. Kwon, J.S., Wereley, S.T.: Light-actuated electrothermal microfluidic motion: investigation and physical interpretation. *Microfluid. Nanofluid.* **5**, 609–620 (2005)

Wireless System



A Secondary Surveillance Radar Data Analysis Technique Based on Geometrical Method

Jing Gao^(✉), Jie Zou, and Ning Guo

The Second Research Institute of CAAC, Chengdu, China
124775272@qq.com

Abstract. Due to the influence of special topography and meteorology conditions, the southwest ATC secondary surveillance radar has a unique characteristic which consists of a complex interference condition and a difference in data character compare with other areas. In order to analyze the blind area of the radar coverage and the surveillance of unreasonable target movement in this area, a radar target data analysis technique based on geometrical method is proposed in this paper. The data analysis technique consists of a horizontal position discrimination region method and a vertical altitude discrimination interval method. The technique can be implemented by the air traffic control department with a rational analysis of critical data components in the radar data. Therefore, by using such technique to achieve the purpose of improving the capability of critical data monitoring and maintenance, providing reasonable issue detection, and to ensure the safety of civil aviation operation.

Keywords: Secondary surveillance radar · Geometrical method · Data analysis · Elevation coverage

1 Introduction

The secondary surveillance radar (SSR) has been utilized by ATC as an initiative and efficient tool to maintain traffic control operations [1]. Based on the influences of topography and meteorology conditions in southwest zone, the data analysis process becomes more particular regarding the data characters and interference parameters. Thus, the evaluation of SSR data analyze can be carried out from two aspects, to evaluate the quality of connectivity and signal stability at physical layer, or to evaluate the conformity from data formation point of view [2]. Critical data such as altitude, speed will only go through limitation determination process since there is no deeper data quality analysis can be done by daily system operation and maintenance [3].

Therefore, based on the traditional radar data analysis method, it is an essential process to consider a better technique to analyze the data quality of SSR regarding the complex terrain factors and the target movement characteristics.

2 Data Reception and Processing

2.1 Nonstandard Radar Data Processing

In general, one frame of data contains only one service information or target information as shown in Fig. 1, the standard ASTERIX format can be used for analysis [4].

01 03 02 00 0b f0 03 20 02 68 28 20 a9

Fig. 1. Under normal circumstances of a frame of service information

A radar located in southwest sending the assembly data is more special, as shown in Fig. 2. If the standard format is used, the data of the underlined part will be treated as discarded and will result in a systematic loss of critical information.

```

01 03 02 00 0b f0 03 20 02 68 28 20 a9 01 00 db ff d4 03 20 a0 01 46 42 51
68 e0 11 d1 e4 09 08 79 72 11 23 46 04 8c 20 aa b9 40 ff d4 03 20 a0 00 19
49 1c 69 d0 12 ed e0 ba 08 70 72 97 23 51 05 2c 20 ad ba 40 ff d4 03 20 b0
01 0b 13 0f 6a 78 04 cf f7 c8 05 ff f7 8b 29 4d 02 1c 20 ae c7 40 ff d4 03
20 a0 00 4c 68 2a 6b 0c 19 96 d2 a4 08 53 6d 10 2f 0f 05 cc 20 af ae 40 ff
d4 03 20 b0 01 a1 0d 56 68 40 03 a6 fa 79 05 98 ee b8 2e a1 01 a8 20 a9 cb
40 ff d4 03 20 a0 00 07 57 1e 69 b0 16 a9 da cd 07 ef df dd 20 77 05 54 20
ac b6 40 ff d4 03 20 a0 01 14 59 2f 6f 1c 11 fa d7 32 07 64 b0 5b 28 13 04
b4 20 b7 b8 40 ff d4 03 20 a0 01 01 37 42 6d 08 0c 65 e7 50 04 59 3e 49 22
02 02 00 20 b3 b3 40
    
```

Fig. 2. Southwest a radar a frame of raw data

According to the data characteristics, SIC and SAC are used for nonstandard one frame data segmentation, and then reverse check the SIC character data (radar format type CAT characters, data length characters, and FS characters) data rationality, as above conditions are in line with the standard ASTERIX format for analysis. The multiple standard data format included in the segmentation is shown in Fig. 3.

```

01 03 02 00 0b f0 03 20 02 68 28 20 a9
01 00 db ff d4 03 20 a0 01 46 42 51 68 e0 11 d1 e4 09 08 79 72 11 23 46 04 8c 20 aa b9 40
ff d4 03 20 a0 00 19 49 1c 69 d0 12 ed e0 ba 08 70 72 97 23 51 05 2c 20 ad ba 40
ff d4 03 20 b0 01 0b 13 0f 6a 78 04 cf f7 c8 05 ff f7 8b 29 4d 02 1c 20 ae c7 40
ff d4 03 20 a0 00 4c 68 2a 6b 0c 19 96 d2 a4 08 53 6d 10 2f 0f 05 cc 20 af ae 40
ff d4 03 20 b0 01 a1 0d 56 68 40 03 a6 fa 79 05 98 ee b8 2e a1 01 a8 20 a9 cb 40
ff d4 03 20 a0 00 07 57 1e 69 b0 16 a9 da cd 07 ef df dd 20 77 05 54 20 ac b6 40
ff d4 03 20 a0 01 14 59 2f 6f 1c 11 fa d7 32 07 64 b0 5b 28 13 04 b4 20 b7 b8 40
ff d4 03 20 a0 01 01 37 42 6d 08 0c 65 e7 50 04 59 3e 49 22 02 02 00 20 b3 b3 40
    
```

Fig. 3. Adjusted data

2.2 Quality Analysis Based on Radar Data

Traditional air traffic control secondary surveillance radar quality analysis usually includes two methods—communication level and data level. The quality analysis on

the communication layer mainly includes whether the signal arrives, the standard signal pin, the transmission baud rate, the CRC check, the framing and the number of frames, the delay, and the signal stability. The data layer on the main includes the format of the grammar check, if the rotation cycle is normal, whether the angle and distance of measurement are in deviation, the target (position, altitude, speed) loss/jump, coverage [5].

3 Radar Data Quality Analysis Based on Geometric Region

Since the aircraft is a moving target, according to the parameters of current location, speed, direction, integrated aircraft performance, radar wave error, and radar coverage, we can determine the aircraft forecast data (position, direction) of the relevant range at the next radar cycle. If there is a radar detection situation outside the forecast delineation, a clear question and a source of origin may be given. Target positioning becomes one of the most important analysis. This paper presents a radar target position detection discrimination method based on the geometric region (referred to as the target position discrimination method TPDM). This method includes the horizontal position discrimination region and the vertical altitude discrimination interval.

In TPDM, some part of aircraft performance parameter values are required, the current performance of all aircraft to analyze the parameters to get the minimum taxi distance turning radius: $R_{\min} = 31.14$ m, Maximum cruising speed: $V_{\max} = 900$ km/h, Maximum rate of increase: $C_{\text{up}} = 17$ m/s, Maximum rate of decline: $C_{\text{down}} = 32$ m/s. These important parameters are the fundamental values of the target location discrimination algorithm.

3.1 Horizontal Position Discrimination Area

Based on the flight direction, the maximum flight speed and the radar scan period of the last radar detection target position O_s , the next detection target position C_r of the radar can be determined and rounded at C_r and the radius is R_{\min} , and O_s for two outer lines l_1 and l_2 , where C_{\max} is the limit boundary range of the radar detection target, and the geometric principle is shown in Fig. 4.

In order to facilitate the calculation of the coordinates of each point, so that the current O_s point as the origin of the coordinates to the east as the X axis, north as the Y axis to establish the Cartesian coordinate system, as shown in Fig. 5. Let O_s be the coordinates (x_0, y_0) , C_r (x_r, y_r) , C_{\max} (x_{\max}, y_{\max}) , A (x_A, y_A) , B (x_B, y_B) .

The straight line distance from the current point to the point C_r is $S = \frac{3}{2}(v_0 - \frac{1}{3}v')t$, where v' is the historical track point velocity at the previous cycle. According to the track direction θ , the coordinates (x_r, y_r) of the center of the circle can be calculated, i.e.,

$$\begin{cases} x_r = x_0 + S \cos(\theta) \\ y_r = y_0 + S \sin(\theta) \end{cases} \quad (1)$$

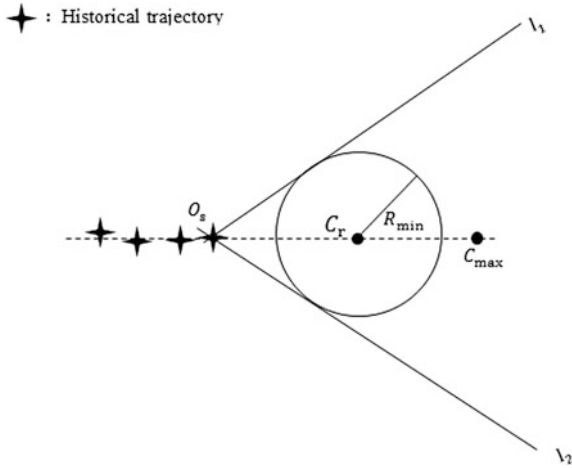


Fig. 4. Predicted radar scans the target position area

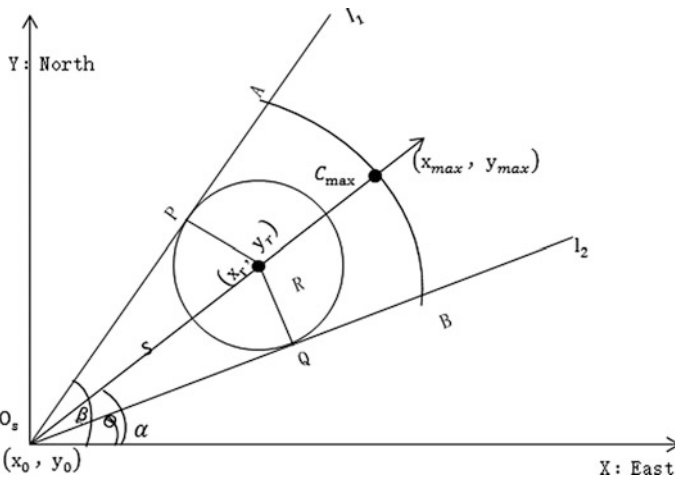


Fig. 5. Radar data detection geometric area (horizontal area) schematic

Now take (x_r, y_r) as the center of the circle, draw a circle with a radius of R_{max} , and made O_s point for the ray l_1, l_2 round tangent, tangent points are P and Q, can be obtained l_1 and l_2 the linear equation:

$$l_1 : y = y_0 + \tan \beta(x - x_0); \tag{2}$$

$$l_2 : y = y_0 + \tan \alpha(x - x_0); \tag{3}$$

where the angle α and β between the two lines and the east X axis is given by the following formula:

$$\beta = \theta + \arcsin \frac{R_{min}}{v_0 t}; \quad \alpha = \theta - \arcsin \frac{R_{min}}{v_0 t} \tag{4}$$

When the speed is the maximum cruising speed, there are $S_{max} = V_{max}t$, The coordinates (x_{max}, y_{max}) of the furthest point C_{max} at the next scanning period can be calculated, i.e.,

$$\begin{cases} x_{max} = x_0 + S_{max} \cos(\theta) \\ y_{max} = y_0 + S_{max} \sin(\theta) \end{cases} \tag{5}$$

Now S_{max} is the radius, O_s for the circle to draw the arc AB intersects two lines for the A and B, there $|AO_s| = |BO_s| = S_{max}$, so A, B coordinates were

$$\begin{cases} x_A = x_0 + S_{max} \cos \beta \\ y_A = y_0 + S_{max} \sin \beta \end{cases} \tag{6}$$

$$\begin{cases} x_B = x_0 + S_{max} \cos \alpha \\ y_B = y_0 + S_{max} \sin \alpha \end{cases} \tag{7}$$

At this point, O_sAB three points of the coordinates have been calculated, the three points surrounded by the sector-shaped area is the radar data detection geometric area (horizontal area).

3.2 Vertical Altitude Discrimination Interval

The geometrical principle is shown in Fig. 6. The altitude interval falls between the maximum climb rate and the maximum decline rate. The h_0 is the altitude of the current track point, and the altitude may be between $[h_B, h_A]$ after the next scan period t , h_A is the altitude of the highest point A, h_B is the altitude of the lowest point B. Relationship there are:

$$h_A = h_0 + tC_{upMax} \tag{8}$$

$$h_B = h_0 - tC_{downMax} \tag{9}$$

$[h_B, h_A]$ determine the vertical area of the geometric region of the radar data detection. Where C_{upMax} and $C_{downMax}$ are the maximum climb rate and the minimum decline rate.

4 Analysis of ATC SSR Data

In this paper, the data is derived from the detection data of a radar station in the southwest. The scanning period is 4 s. The nonstandard radar data processing method is used to process the SSR data of the air traffic control and analyzed according to the

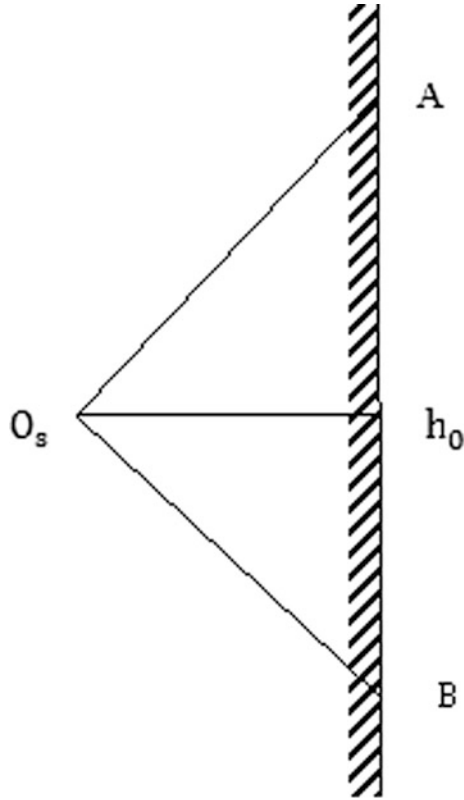


Fig. 6. Radar data detection geometric area (vertical area) schematic

standard ASTERIX. The specific method is based on the basic data quality analysis mentioned in the above 2.2.

4.1 Judgment Method

To determine whether the coordinates of a_3 fall in sector O_sAB as follows:

- (1) To determine whether the direction of a_3 fall between $[\alpha, \beta]$;
- (2) To determine the angle ϑ (that is, the direction of a_3 angle) of the ray of the $O_s a_3$ point and the X axis is in the $[\alpha, \beta]$ between; calculation formula:
 $\vartheta = \arctan\left(\frac{y_3 - y_0}{x_3 - x_0}\right)$, if $\vartheta \notin (\alpha, \beta)$, then the result is a_3 data is incorrect; If $\vartheta \in (\alpha, \beta)$, proceed to method (3);
- (3) To determine whether a_3 is within the sector O_sAB ; calculate the distance $|O_s a_3|$, which is given by the following formula:

$$|O_s a_3| = \sqrt{(x_3 - x_0)^2 + (y_3 + y_0)^2} \tag{10}$$

If $|O_s a_3| < S_{max}$, get a_3 in the sector, otherwise, The result of the judgment is that the a_3 data is incorrect.

According to the radar high-level coverage and vertical altitude interval to re-judge whether a_3 can be scanned by radar, as shown in Fig. 7, where h_0 is the elevation of a_2 , $h_1 \sim h_4$ are radar coverage in the direction of a_2 visible The elevation of the mountain. Whether a_3 can be scanned is as follows:

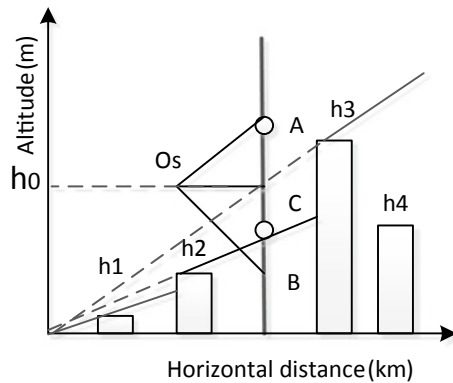


Fig. 7. Vertical radar coverage

- (1) According to the elevation and direction, the first contour line can be found, which is greater than or equal to a_2 elevation, as shown in Fig. 7.
- (2) Calculating the value h_C at intersection point C between line AB and line h_2h_3 , then the vertical interval modified to $[h_C, h_A]$;
- (3) Determining the altitude H_3 of a_3 . If it is found that a_3 in the vertical altitude interval, a_3 data is correct, otherwise a_3 data is not correct.

4.2 Data Analysis

One of the radar scans position data $a_1(-23802.2, 6677.26)$, $a_2(-23930.1, 7078.89)$, $a_3(-24115.6, 7411.31)$. Altitude $h_1 = 2900$, $h_2 = 2900$, $h_3 = 2900$. Flight direction were 103.5984, 120.04, 123.766. Where a_1 , a_2 is the validated data, and the data to be verified a_3 is analyzed. According to the target position discrimination method, a_2 is the target track point data of the current radar scanning. The speed is 510 km/h, the direction is 124.04, the altitude is 2900 m, and the speed of a_1 is 500 km/h. The three coordinate points of the sector O_sAB at this time are $O_s(-23930.1, 7078.89)$, $A(-24478, 7915.7)$, $B(-24382, 7970.8)$. The vertical altitude interval $[2772, 2968]$, $\alpha = 116.887$, $\beta = 123.1209$. According to the judgment method of 4.1, it can be seen that a_3 falls in the sector-level O_sAB horizontal position area. As shown in Fig. 8, a_3 position data is reasonable, the flight direction is reasonable, the speed value is

reasonable. According to the method of 4.1 to calculate $h_c = 2810$ m, corrected vertical altitude interval, and a_3 altitude $h_3 = 2900$ m, we can see that the altitude data to determine reasonable.

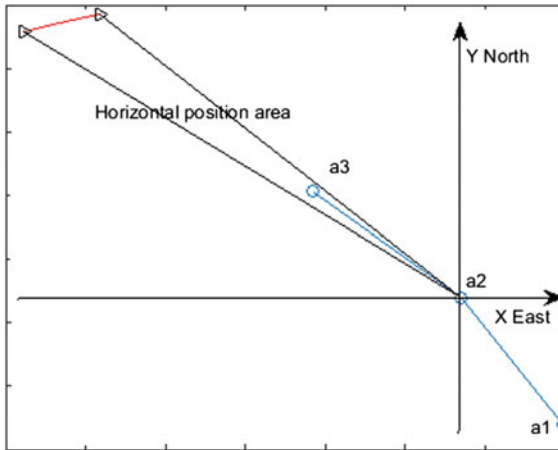


Fig. 8. To be analyzed in the horizontal area outside the situation

5 Conclusions

This paper illustrated an additional method of ATC secondary surveillance radar data analysis without considering the communication level and the standard ASTERIX data format level. Using the GIS characteristics to obtain 36° radar full terrain coverage. Based on geometric location area and the proposed radar detection target location identification method which including the horizontal position discrimination method and vertical altitude discrimination interval method, to achieve the ATC secondary surveillance radar data positioning quality and data rationality analysis.

This method combines the elevation and motion characteristics to solve the problem of rationality of the critical data determination, in order to make a more comprehensive secondary surveillance radar data analysis. The actual data are used to solve the problem of data quality reliability which overcomes the issue of multi-radar data fusion. By using the method as such, we can ensure the safety of critical data of ATC automation, and maintain the safety of civil aviation operation and air traffic control.

References

1. Weeda, D.J.A., Ligthart, L.P., Nieuwkerk, L.R., Nieuwkerk, L.R., van der Klein, D.C.M.: Quantitative estimation of secondary surveillance radar information. *J. Navig.* **1**, 26–35 (1992)
2. Kim, E., Sivits, K.: Blended secondary surveillance radar solutions to improve air traffic surveillance. *Aerop. Sci. Technol.* 203–208 (2015)

3. Wyndham, B.A.: Secondary surveillance radar. *Aeronaut. J.* (919), 375–375 (1988)
4. Senoguchi, A., Koga, T.: Analysis of downlink aircraft parameters monitored by SSR mode S in ENRI. In: *Digital Avionics Systems Conference* (2009)
5. Guo-hai, C.A.O., Yun-feng, W.A.N.G.: Anomalies analysis of data fusion for ATC primary and secondary radar. *Telecommun. Eng.* **54**(4), 452–456 (2014)



Spectral Efficiency Analysis of Hybrid Precoding in Millimeter Wave MIMO Systems

Jing Li¹, Dianwu Yue^{1(✉)}, and Meng Wang^{1,2}

¹ College of Information Science and Technology, Dalian Maritime University,
Dalian 116026, China
dwue@dmlu.edu.cn

² Department of Electronic Engineering, Dalian Neusoft University of Information,
Dalian 116026, China

Abstract. The utilization of millimeter wave (mmWave) frequency has shown great potential for future wireless cellular systems. Compared with conventional fully digital precoding, hybrid precoding achieves a high performance with less radio frequency (RF) chains and lower hardware complexity. In this paper, we investigate the hybrid precoding for multiuser communication in a downlink mmWave massive MIMO system. First, we present a hybrid precoding scheme designed based on zero-forcing (ZF) precoding. Then, we derive the spectral efficiency of this hybrid precoding scheme and get its closed-form expression which can be represented by the digamma function. Finally, simulation results demonstrate the validity of the analytical result.

Keywords: Hybrid precoding · Millimeter wave · Zero forcing · Spectral efficiency

1 Introduction

As a promising technology for upcoming fifth generation (5G) systems, mmWave communications have attracted tremendous attention due to large bandwidths and rich spectral resource [1]. Massive MIMO technology can be well combined with mmWave as the decrease in wavelength of mmWave makes it possible to place a very large number of antennas in a much smaller dimension. Moreover, the large antenna arrays can provide sufficient multiplexing gain by precoding and combining to overcome the free-space path loss of mmWave channel.

It is not practical to apply conventional fully digital precoding structure in mmWave systems because the implementation of fully digital precoding requires one dedicated RF chain per antenna element, which leads to unacceptable power consumption and cost at mmWave systems. To address the RF hardware con-

straint, hybrid precoding is proposed by achieving near-optimal performance with much less RF chains.

To date, most researches on hybrid precoding for mmWave MIMO systems have been focused on the improvement of hybrid precoder structure and the design of optimization algorithm. There are several typical algorithms depending on different precoder structures. The orthogonal matching pursuit (OMP) algorithm utilizes the spatially sparse characteristic of the mmWave channel for fully connected structure [2], the success interference cancellation (SIC) algorithm leverages the block diagonalization property of RF precoding matrix for partially connected structure [3], and the matrix factorization algorithm designs the precoding matrix based on matrix polar decomposition for hybridly connected structure [4]. More recently, the researches have been shifted to the analysis of spectral efficiency and energy efficiency with different hybrid precoder structures [3, 5]. However, most literature just list the spectral efficiency expression and compare the spectral efficiency of different precoding schemes via numerical simulation, which motivates us to derive a closed-form expression of spectral efficiency for hybrid precoding based on a specified algorithm and structure.

In this paper, we consider a multiuser mmWave massive MIMO system, present a hybrid precoding scheme designed based on ZF precoding, and then get the closed-form expression of spectral efficiency which can be represented by the digamma function. Finally, numerical results demonstrate the validity of the analytical expression.

2 System Model

Consider a narrowband downlink multiuser mmWave MIMO system as shown in Fig. 1. A base station (BS) with N_t transmit antennas and N_t^{RF} RF chains is assumed to communicate with K users. Each user has N_r antennas. To transmit N_s data streams, the RF chains number of the BS must satisfy $N_s \leq N_t^{RF} \leq N_t$. We focus on the multiuser beamforming case in which the BS communicates with every user via only one stream. Therefore, the total number of streams $N_S = K$. The BS applies an $N_t \times N_t^{RF}$ analog precoder $\mathbf{F} = [\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_{N_t^{RF}}]$ and a $N_t^{RF} \times N_s$ digital precoder $\mathbf{B} = [\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_{N_s}]$. The digital precoder modifies both phase and amplitude while only phase changes are made by the RF precoder with a variable analog phase shifters network. Thus, each entry of the RF precoder is of constant modulus, which is normalized to satisfy $\mathbf{F}_{i,j} = \frac{1}{\sqrt{N_t}} e^{j\varphi_{i,j}}$, where $\varphi_{i,j}$ is the angle of phase shifter and $\mathbf{F}_{i,j}$ is the (i, j) th element of \mathbf{F} . To meet the total transmit power constraint, the entries of \mathbf{B} are normalized to satisfy $\|\mathbf{FB}\|_F^2 = N_s$, where $\|\cdot\|_F$ is the Frobenius norm. The transmitted signal is given by

$$\mathbf{x} = \mathbf{FB}\mathbf{s} \quad (1)$$

where $\mathbf{s} = [s_1, s_2, \dots, s_K]^T$ is the $K \times 1$ vector of transmitted symbols, satisfying $\mathbb{E}\{\mathbf{s}\mathbf{s}^H\} = \frac{P}{K}\mathbf{I}_K$, where P is the total transmitted power. $(\cdot)^T$ is the transpose

of a matrix and $(\cdot)^H$ means the conjugate transpose of a matrix. We assume equal power allocation for different users' streams, and after the processing of the RF combiner at the k th user, the signal is represented as

$$y_k = \mathbf{w}_k^H \mathbf{H}_k \sum_{k=1}^K \mathbf{F} \mathbf{B}_k s_k + \mathbf{w}_k^H \mathbf{n}_k \tag{2}$$

where \mathbf{w}_k is the RF combiner of the k th user, \mathbf{H}_k is the $N_r \times N_t$ vector which represents the mmWave channel between the BS and the k th user, and $\mathbf{n}_k \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})$ is the additive white Gaussian noise vector. Throughout this paper, we assume that the channel state information is known to both the transmitter and receiver. The received signal-to-interference-plus-noise ratio (SINR) of the k th user can be expressed as

$$\text{SINR}_k = \frac{\frac{P}{K} |\mathbf{w}_k^H \mathbf{H}_k \mathbf{F} \mathbf{B}_k|^2}{\sum_{k \neq u} \frac{P}{K} |\mathbf{w}_k^H \mathbf{H}_k \mathbf{F} \mathbf{B}_u|^2 + \sigma^2} \tag{3}$$

If Gaussian inputs are used, the system can achieve a long-term average (over the fading distribution) spectral efficiency, which is given as

$$R = \sum_{k=1}^K \mathbb{E} \{ \log_2(1 + \text{SINR}_k) \} \tag{4}$$

Due to high free-space path loss, mmWave channels are expected to have limited scattering [2], and the propagation environment is well characterized by the extended Saleh–Valenzuela model [6]. Under this model, the mmWave channel between the k th user and the BS can be expressed as

$$\mathbf{H}_k = \sqrt{\frac{N_t N_r}{L}} \sum_{\ell=1}^L \alpha_{k,\ell} \mathbf{a}_r(\theta_{k,\ell}) \mathbf{a}_t^H(\phi_{k,\ell}) \tag{5}$$

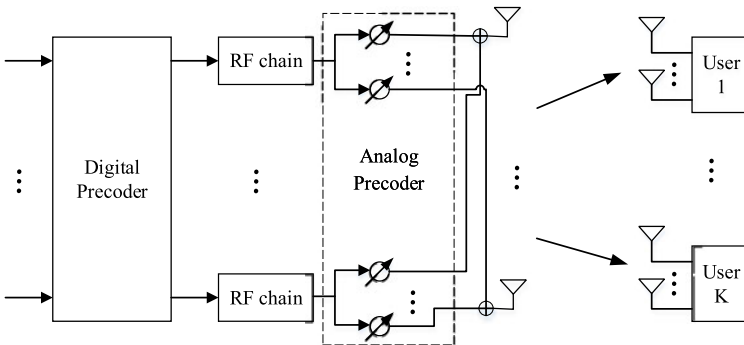


Fig. 1. System model of the mmWave MIMO system with hybrid precoding structure

where $\alpha_{k,\ell}$ denotes the complex gain for the ℓ th path including the path loss. We assume that α_k are i.i.d. random variables following the complex Gaussian distribution, i.e., $\alpha_{k,\ell} \sim \mathcal{CN}(0, 1)$. The variables $\theta_{k,\ell}, \phi_{k,\ell} \in [0, 2\pi]$ are the ℓ th path's angle of arrival and departure (AoA/AoDs), respectively. Finally, $\mathbf{a}_r(\theta_{k,\ell})$ and $\mathbf{a}_t(\phi_{k,\ell})$ are the antenna array response vectors of the BS and the k th user, respectively. If a uniform linear array (ULA) is assumed, the transmit array response vector $\mathbf{a}_t(\phi_{k,\ell})$ is defined as

$$\mathbf{a}_t(\phi_{k,\ell}) = \frac{1}{\sqrt{N_t}} \left[1, e^{j\frac{2\pi}{\lambda}d \sin(\phi_{k,\ell})}, \dots, e^{j\frac{2\pi}{\lambda}(N_t-1)d \sin(\phi_{k,\ell})} \right]^T \quad (6)$$

where d is the distance between antenna elements and λ is the signal wavelength. The user's receive array response vector $\mathbf{a}_r(\theta_{k,\ell})$ can be written in the similar fashion.

3 Analog and Digital Precoder Design

In conventional massive MIMO systems, ZF precoding is generally acknowledged as optimal for fully digital precoding. To realize the fully digital beamforming performance, it is sufficient that the number of RF chains in fully connected architecture is greater than or equal to twice the number of data streams [7]. For easy analysis, we choose $N_t^{RF} = 2N_s$. Denote the fully digital ZF precoder as \mathbf{V} , and $\mathbf{V}_{i,j} = v_{i,j} e^{j\gamma_{i,j}}$ is the (i, j) th element of \mathbf{V} . Choose the k th column of the digital precoder $\mathbf{B}_k = [0, \dots, \sqrt{N_t}b_{2k-1}, \sqrt{N_t}b_{2k}, \dots, 0]^T$. Then, satisfying $\mathbf{FB} = \mathbf{V}$ is equivalent to

$$b_{2k-1} e^{j\varphi_{i,2k-1}} + b_{2k} e^{j\varphi_{i,2k}} = v_{i,k} e^{j\gamma_{i,k}} \quad (7)$$

for $i = 1, \dots, N_t$ and $k = 1, \dots, N_s$. The equation has multiple solutions. From the triangle inequality, the norm of the composite complex number satisfies

$$\left| |b_{2k-1}| - |b_{2k}| \right| \leq \left| b_{2k-1} e^{j\varphi_{i,2k-1}} + b_{2k} e^{j\varphi_{i,2k}} \right| = v_{i,k} \leq |b_{2k-1}| + |b_{2k}| \quad (8)$$

Along this line, we further choose one solution $b_{2k-1} = b_{2k} = v_{\max}^k$, where $v_{\max}^k = \max\{v_{i,k}\}$ for $i = 1, \dots, N_t$. It can be verified by the extended Euler formula $e^{j\vartheta} + e^{-j\vartheta} = 2 \cos(\vartheta)$ that the following is a solution for the above equation.

$$\varphi_{i,2k-1} = \gamma_{i,k} - \arccos\left(\frac{v_{i,k}}{2v_{\max}^k}\right) \quad (9)$$

$$\varphi_{i,2k} = \gamma_{i,k} + \arccos\left(\frac{v_{i,k}}{2v_{\max}^k}\right) \quad (10)$$

With the previous analysis, we adopt the algorithm for hybrid precoder design in the case that $N_t^{RF} = 2N_s$. Multi-stream baseband precoding can be applied to the mmWave channel matrix $\mathbf{H} = [\mathbf{H}_1^H, \mathbf{H}_2^H, \dots, \mathbf{H}_K^H]^H$, where the fully digital ZF precoder can be performed as

$$\mathbf{V} = \mathbf{H}^H (\mathbf{H}\mathbf{H}^H)^{-1} \mathbf{A} \quad (11)$$

where \mathbf{A} is a diagonal matrix, introduced for column power normalization of the hybrid precoder. Considering the special structure of the column vector \mathbf{B}_k in the algorithm, the digital precoder \mathbf{B} is a block diagonal matrix. As the matrix \mathbf{V} is not block diagonal and $\mathbf{FB} = \mathbf{V}$, the analog precoder \mathbf{F} is also not block diagonal, which means the hybrid precoder we presented is a fully connected structure as shown in Fig.1.

4 Spectral Efficiency Analysis

In the following, we quantitatively analyze spectral efficiency achieved by the designed scheme. For the k th single-antenna user, the RF combiner and the receive array response vector $\mathbf{a}_r(\theta_{k,\ell})$ become constants. Then, the mmWave channel between the k th user and the BS can be simplified as

$$\mathbf{H}_k = \mathbf{D}\mathbf{A}^H \quad (12)$$

where $\mathbf{D} = \sqrt{\frac{N_t}{L}} [\alpha_{k,1}, \alpha_{k,2}, \dots, \alpha_{k,L}]$ and $\mathbf{A} = [\mathbf{a}_t(\phi_{k,1}), \mathbf{a}_t(\phi_{k,2}), \dots, \mathbf{a}_t(\phi_{k,L})]$. In the spectral efficiency analysis, we introduce the asymptotic orthogonality property of mmWave MIMO channel assuming the number of transmit antennas approaches infinity [8]. Then, for a ULA system with large scale antennas, we assume that the transmit antenna vector is asymptotically orthogonal. In combination with the asymptotic orthogonality property, substitute (11) and (12) into (3), we can obtain

$$\begin{aligned} \text{SINR}_k &= \frac{P}{K\sigma^2} |\mathbf{w}_k^H \mathbf{H}_k \mathbf{F} \mathbf{B}_k|^2 \\ &= \frac{P}{K\sigma^2} |\Lambda_{k,k}|^2 \\ &= \frac{PN_t}{KL\sigma^2} \sum_{\ell=1}^L |\alpha_{k,\ell}|^2 \\ &= \beta \sum_{\ell=1}^L |\alpha_{k,\ell}|^2 \end{aligned} \quad (13)$$

where $\beta \triangleq \frac{PN_t}{KL\sigma^2}$. As the assumption $\alpha_{k,\ell} \sim \mathcal{CN}(0, 1)$, it can be concluded that $|\alpha_{k,\ell}|^2$ follows an exponential distribution with parameter 1, i.e., $|\alpha_{k,\ell}|^2 \sim \text{Exp}(1)$. When the number of transmit antennas approaches infinity, different paths of the users are independent. Denote $x \triangleq \sum_{\ell=1}^L |\alpha_{k,\ell}|^2$. The entry x , which expresses the summation of L independent variables that follow the same exponential distribution, follows Gamma distribution with a shape parameter L and an inverse scale parameter 1, i.e., $x \sim \Gamma(L, 1)$. The probability density function of x is given by

$$f(x) = \begin{cases} \frac{x^{L-1} e^{-x}}{\Gamma(L)}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

where $\Gamma(L) = (L-1)!$ is the Gamma function. Substitute (13) and (14) into (4), the spectral efficiency achieved by the adopted scheme can be written as

$$\begin{aligned}
R &= \sum_{k=1}^K \mathbb{E} \{ \log_2 (1 + \beta x) \} \\
&= \sum_{k=1}^K \int_0^{\infty} \log_2 (1 + \beta x) \frac{x^{L-1} e^{-x}}{\Gamma(L)} dx \\
&\stackrel{(a)}{\approx} \log_2 e \sum_{k=1}^K \int_0^{\infty} \ln(\beta x) \frac{x^{L-1} e^{-x}}{\Gamma(L)} dx \\
&= \log_2 e \sum_{k=1}^K \left[\int_0^{\infty} (\ln \beta) \frac{x^{L-1} e^{-x}}{\Gamma(L)} dx + \int_0^{\infty} (\ln x) \frac{x^{L-1} e^{-x}}{\Gamma(L)} dx \right] \\
&\stackrel{(b)}{=} \sum_{k=1}^K \left[\log_2 \beta + \log_2 e \frac{\Gamma'(L)}{\Gamma(L)} \right] \\
&= K \log_2 e [\ln \beta + \psi(L)]
\end{aligned} \tag{15}$$

where (a) is derived from $\ln(1 + \beta x) \approx \ln(\beta x)$ as N_t goes to infinity. The left item of (b) results from $\int_0^{\infty} \frac{x^{L-1} e^{-x}}{\Gamma(L)} dx = 1$ which can be computed from the probability density function $f(x)$ in the Eq. (14), while the right item of (b) follows from the equation $\int_0^{\infty} (\ln x) x^{L-1} e^{-x} dx = \frac{d\Gamma(L)}{dx} = \Gamma'(L)$. The entry $\psi(L)$ is the value of the digamma function $\psi(z)$ which is defined as $\psi(z) = \frac{\Gamma'(z)}{\Gamma(z)}$ at the point $z = L$. When z is greater than zero, the value of $\psi(z)$ increases monotonically as L increases.

As the signal-to-noise ratio (SNR) can be denoted as $\frac{P}{\sigma^2}$, according to the Eq. (15), the spectral efficiency monotonically increases as SNR and the number of transmit antennas increase when $\beta \geq 1$. The effect of channel paths on the spectral efficiency depends on the magnitude of $\psi(L)$ and $\ln L$.

5 Simulation Results

In this section, simulation results are presented to show the performance of the designed algorithm and the validity of the derived expression. Assume that $K = 4$ data streams are sent from the BS to four single-antenna users while the number of BS RF chains is 8. Considering the sparsity of the mmWave MIMO channel, we choose the path number $L = 10$. Assume the noise variance is 1. For each simulation, the spectral efficiency is plotted over 1000 channel realizations. For comparison, we also simulate the low-complexity phase zero precoding (PZF) scheme proposed in [9] and beamspace MIMO (B-MIMO) scheme proposed in [10].

Figure 2 illustrates the spectral efficiency of different hybrid precoding schemes for $N_t = 256$. As expected, the analytical spectral efficiency of the proposed scheme is very accurate in high SNR regime. Compared with PZF and B-MIMO schemes, the designed algorithm has a better performance. The

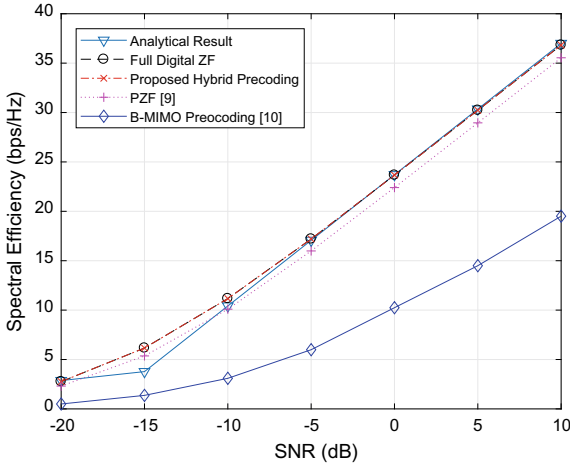


Fig. 2. Spectral efficiency versus SNR

analytical spectral efficiency in the Eq. (15) is derived when the value N_t goes to infinity. On the condition that N_t is a relatively small number, SNR is low and $\beta < 1$ which leads to $\log_2(\beta)$ decreases as SNR increases, and the spectral efficiency also decreases. When the SNR increases and reaches a specified value, $\beta \geq 1$ leads to the result that $\log_2(\beta)$ increases as SNR increases. This explains the reason why SNR = -15 dB and -20 dB have a close value.

Next, we analyze the spectral efficiency versus the number of transmit antennas as shown in Fig. 3. In order to express the asymptotic property clearly, we

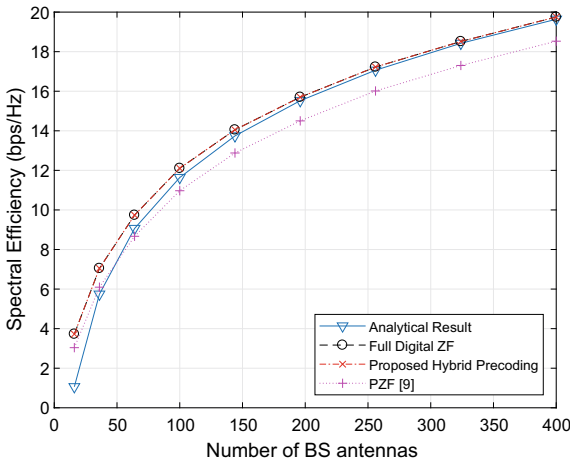


Fig. 3. Spectral efficiency versus transmit antennas

choose $\text{SNR} = -5$ dB as the gaps of spectral efficiency between different precoding schemes are relatively clear at this point. In Fig. 3, it is observed that, as the BS transmit antennas N_t increases, the analytical result gradually approaches the practical result and the performance gap between the analytical result and the practical result decreases. It indicates the validity of our analytical result. Also, the spectral efficiency increases as the transmit antennas increases, which is also consistent with our theory analysis.

6 Conclusion

In this paper, we presented a high-performance hybrid precoding algorithm for downlink multiuser mmWave systems leveraging the large number of deployed antennas. The performance of the presented algorithm was analyzed for single-antenna users when the system dimensions were very large. Moreover, the closed-form expression of spectral efficiency for the adopted scheme was derived. Simulation results demonstrated the validity of the analytical result.

References

1. Roh, W., et al.: Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results. *IEEE Commun. Mag.* **52**(2), 106–113 (2014)
2. El Ayach, O., et al.: Sparsely sparse precoding in millimeter wave MIMO systems. *IEEE Trans. Wirel. Commun.* **13**(3), 1499–1513 (2014)
3. Gao, X., Dai, L., Han, S., et al.: Energy-efficient hybrid analog and digital precoding for mm-wave MIMO systems with large antenna arrays. *IEEE J. Sel. Areas Commun.* **34**(4), 998–1009 (2015)
4. Zhang, D., Wang, Y., Li, X., et al.: Hybridly-connected structure for hybrid beamforming in mm-wave massive MIMO systems. *IEEE Trans. Commun.* **66**(2), 662–674 (2017)
5. Yu, X., Shen, J.C., Zhang, J., et al.: Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems. *IEEE J. Sel. Top. Signal Process.* **10**(3), 485–500 (2016)
6. Saleh, A.A.M., Valenzuela, R.: A statistical model for indoor multipath propagation. *IEEE J. Sel. Areas Commun.* **5**(2), 3679–3684 (2013)
7. Sohrabi, F., Yu, W.: Hybrid digital and analog beamforming design for large-scale antenna arrays. *IEEE J. Sel. Top. Signal Process.* **10**(3), 501–513 (2016)
8. El Ayach, O., et al.: The capacity optimality of beam steering in large millimeter wave MIMO systems. In: *IEEE Workshop on Signal Processing Advances in Wireless Communication*, pp. 100–104 (2012)
9. Liang, L., Xu, W., Dong, X.: Low-complexity hybrid precoding in massive multiuser MIMO systems. *IEEE Wirel. Commun. Lett.* **3**(6), 653–656 (2014)
10. Sayeed, A., Brady, J.: Beamspace MIMO for high-dimensional multiuser communication at millimeter-wave frequencies. In: *Global Communications Conference (GLOBECOM)*, pp. 3679–3684 (2013)



An FPGA-Based Balanced and High-Efficiency Two-Dimensional Data Access Technology for Real-Time Spaceborne SAR

Tianyuan Sun^{1,2}, Bingyi Li^{1,2}, Xiaoning Liu^{1,2},
and Yizhuang Xie^{1,2}(✉)

¹ Radar Research Lab, School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China

xyz551_bit@bit.edu.cn

² Beijing Key Laboratory of Embedded Real-Time Information Processing Technology, Beijing 100081, China

Abstract. With the development of satellite load and very large scale integrated (VLSI) circuit technology, spaceborne real-time synthetic aperture radar (SAR) imaging systems have become a solution for rapid response to hazards. Through analyzing the algorithm pipeline flow as well as introducing the storage-computation model, a balanced and high-efficiency 2-D data access technology based on cross-mapping data storage method has been achieved to suit the large point processing for real-time spaceborne SAR system. A prototype based on NetFPGA-SUME board with Xilinx XC7VX690T is given to verify the performance of the proposed design. Taking Stripmap SAR imaging of 16384 * 16384 granularity raw data (5 m resolution, 25 km width) as an example, the imaging based on chirp scaling algorithm takes 6.63 s, which is better than some other real-time processing methods.

Keywords: SAR · Real-time processing · FPGA parallel accelerating component

1 Introduction

Spaceborne synthetic aperture radar (SAR) is a remote sensor operating in the microwave band which has the features of all-weather and all-day. As one of the most important methods of space-to-earth observation, SAR has been mainly used for maritime surveillance and national security, disaster management and ecosystem monitoring [1]. Following the development of the yield of wide open missions, the demand of real-time performance for SAR data processing raised rapidly. Different kinds of chips such as field programmable gate array (FPGA), application-specific integrated circuit (ASIC), digital signal processors (DSP), and graphic processing units (GPU) are chosen to build a high-performance spaceborne SAR real-time processing platform. DSP has been proposed to accelerate the image formation processing [2]. However, due to its drawbacks such as fixed architecture, low data throughput capacity, and complexity of high-speed design, DSP is not feasible to satisfy the demands of SAR imaging systems. In recent years, GPU implementations have been proposed

which are able to process SAR images at very high parallelism in real time [3]. However, these GPU-based SAR systems have the disadvantage of high power consumption of per device. FPGA-based architectures enable for massive parallel processing and flexible use of hardware resources while only low power is needed. It solves the low ratio between computational power and power consumption problem of DSP and GPU [1].

The proposed design introduces the CS algorithm for SAR image processing [4]. The [5] represents that four times all-pixels FFT are the most computation-hungry operations of CS implementation, and the efficiency burden mainly occurs in the data access after corner turning (matrix transposition). Several optimized methods have been introduced for this problem. The [6] uses the window access mode to accelerate the matrix transposition and the [7] uses ping-pong buffers for DDR SDRAM in multi-DSP system to solve this problem, respectively. In this paper, the cross-mapping method [8] has been introduced in our 2-D SAR data access strategy generation. Through analyzing and modeling the algorithm pipeline flow, a balanced and high-efficiency 2-D data operation technology has been achieved for real-time spaceborne SAR system. A NetFPGA-SUME-based imaging prototype system according to the proposed method is implemented, and the Stripmap SAR raw data with the parameters of Chinese Gaofen3 satellite [9] has been adopted to test the efficiency and real-time performance of the system. The results indicate that our design can be potentially used for spaceborne on-board real-time processing.

2 CS Algorithm and Storage Strategy

In this paper, we only discuss the SAR imaging processing after A/D conversion. The static mapping of CS image processing can be described by the end-to-end processing chain as shown in Fig. 1. N_R and N_A represent the sample numbers of the range direction and azimuthal direction. As the left flowchart in Fig. 1 shows, the standard CS algorithm can be partitioned into two parallel branches. The right branch which is denoted by the solid-line rectangular boxes represent phase function estimation (PFE) and quantify operations. The left branch is denoted by the gray, dashed rectangular boxes, which represent four times FFT and three matrix transpose operations. For each stage, two branches are mutually independent, and thus they can work synchronously. This paper focuses on the left branch optimization for FPGA implementation.

In general, the SAR raw data are stored range-first. Thus, three matrix transpose operations are required during the total processing procedure. Generally, DDR SDRAM always be introduced as external bulk-storage memory for raw data, which means the FPGA should establish a high-speed data access link between the inner processing module and external DDR SDRAM. The DDR-based matrix transpose is achieved using the submatrix cross-mapping method. In this way, matrix transpose operation can be performed simultaneously with the data reading before each FFT stage, and the access efficiency of the range and the azimuth can be balanced. Figure 1 also shows the main two steps of the method. First, divide the $N_A \times N_R$ raw data matrix into $M \times N$ submatrixes of $N_a \times N_r$ order; second, map each submatrix into the three-

dimensional storage array of the DDR. Each submatrix corresponds to a row of DDR, and every two rows of data in a submatrix are cross-mapped to this DDR row. Moreover, each of the two adjacent submatrixes in two dimension is distributed in different banks of DDR.

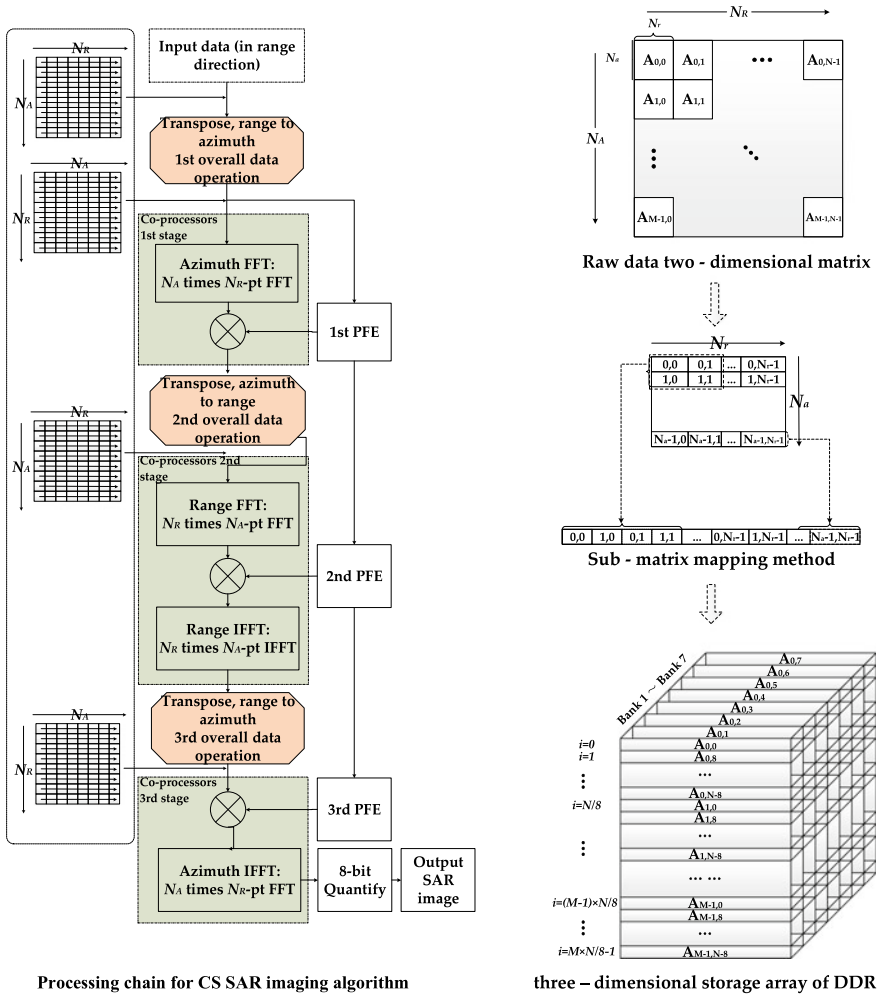


Fig. 1. Flowchart of CS algorithm and storage strategy.

3 Computation-Storage Balanced Modeling Analysis

In this paper, a generic model is given to analyze the bandwidth matching problem between the proposed-like universal systems first. It is assumed that an independent processing flow in the system operation requires i storage operations $m_1 \dots m_i$ and j computing operations $p_1 \dots p_j$. The flow diagram is shown in Fig. 2.

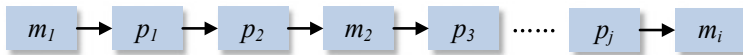


Fig. 2. Schematic diagram of sequential pipelining system.

The bandwidth of each storage operation unit and arithmetic unit is represented by $B_{m1} \dots B_{mi}$, and $B_{p1} \dots B_{pj}$, respectively. So the storage and computation capability of system can be represented as follows:

$$B_{memory} = [x_1 B_{m1}, x_2 B_{m2}, \dots, x_i B_{mi}]_{\min} \tag{1}$$

$$B_{processor} = [y_1 B_{p1}, y_2 B_{p2}, \dots, y_j B_{pj}]_{\min} \tag{2}$$

where x_i as well as y_i means the parallelism of memory or arithmetic unit, which depends on the number of sub-memory (such as RAM and FIFO) and computing core. Thus, the bandwidth matching problem is divided into following two conditions:

$$\begin{aligned} & \max [x_1 B_{m1}, x_2 B_{m2}, \dots, x_i B_{mi}]_{\min} \\ & \text{s.t. } [x_1 B_{m1}, x_2 B_{m2}, \dots, x_i B_{mi}]_{\min} \leq [y_1 B_{p1}, y_2 B_{p2}, \dots, y_j B_{pj}]_{\min} \end{aligned} \tag{3}$$

where m_i and m_j represent the i -th storage and j -th computing operation, respectively. Assumed that system is designed under the limited B_p to achieve maximum utilization of computing bandwidth, the equivalent optimization problem will be changed as shown in (4), which is called calculation first solution.

$$\begin{aligned} & \max [y_1 B_{p1}, y_2 B_{p2}, \dots, y_j B_{pj}]_{\min} \\ & \text{s.t. } [x_1 B_{m1}, x_2 B_{m2}, \dots, x_i B_{mi}]_{\min} \geq [y_1 B_{p1}, y_2 B_{p2}, \dots, y_j B_{pj}]_{\min} \end{aligned} \tag{4}$$

Furthermore, a typical flow of the CS algorithm implementation can be represented by Fig. 3, which is treated as a single sequential pipelining system with storage and computing units. The whole procedure can be divided into four separated pipelines, two for azimuth and two for range. Each pipeline contains a computing process which includes FFT/IFFT operation, phase function estimation (PFE) operation, and a number of storage processes. For large point SAR imaging, the PFE can be approximately considered as a real-time operation, thus FFT/IFFT rate determines the bandwidth of each computing operation. The two azimuths as well as range pipelining processes can be seen as the same pipeline. So, we get $B_{m4} = B_{m13}$, $B_{m3} = B_{m12}$, $B_{m2} = B_{m11}$, $x_4 = x_{13}$,

$x_3 = x_{12}$, $x_2 = x_{11}$, $B_{p1} = B_{p4}$, $y_1 = y_4$ in azimuth and $B_{m7} = B_{m10}$, $B_{m6} = B_{m9}$, $B_{m5} = B_{m8}$, $x_7 = x_{10}$, $x_6 = x_9$, $x_5 = x_8$, $B_{p2} = B_{p3}$, $y_2 = y_3$ in range.

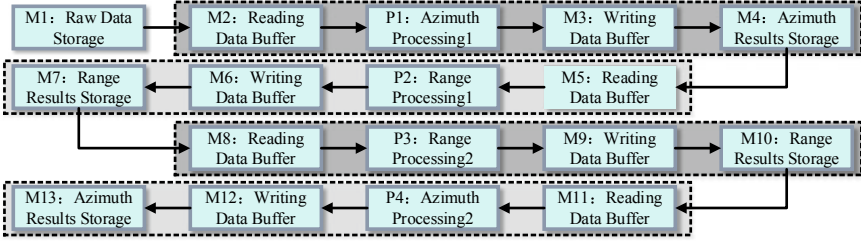


Fig. 3. Flowchart of CS algorithm hardware implementation.

In order to balance two-dimensional data access by cross-mapping, the same implementation method should be designed for range and azimuth. By taking azimuth processing as an example, we get the storage first solution and its calculation is as follows:

$$\begin{aligned} & \max [x_2 B_{m2}, x_3 B_{m3}, x_4 B_{m4}]_{\min} \\ & \text{s.t. } [x_2 B_{m2}, x_3 B_{m3}, x_4 B_{m4}]_{\min} \leq [y_1 B_{p1}]_{\min} \end{aligned} \quad (5)$$

$$\begin{aligned} & \max [y_1 B_{p1}]_{\min} \\ & \text{s.t. } [x_2 B_{m2}, x_3 B_{m3}, x_4 B_{m4}]_{\min} \geq [y_1 B_{p1}]_{\min} \end{aligned} \quad (6)$$

These solutions can be used as the simple method to design the key structure of CS algorithm under different needs.

On-chip memory such as RAM and FIFO can be handled as reading and writing buffers (RAM or FIFO). The optimization problem above has the following relationship:

$$\begin{cases} B_{m2} = F_{m2} W_{m2} \\ B_{m3} = F_{m3} W_{m3} \\ B_{m4} = B_{ddr} = \rho \eta B_{ddr_peak} \\ B_{p1} = B_{ffi} = F_{ffi} W \end{cases} \quad (7)$$

where F_{m2} , F_{m3} represent the clock frequency of the buffer used in storage operation B_{m2} and B_{m3} . W_{m2} , W_{m3} are the word length of B_{m2} and B_{m3} separately. B_{ddr_peak} is the peak bandwidth of DDR. And B_{ddr} , the actual bandwidth of DDR weighted by η and ρ , which are both determined by the data access mode (range/azimuth) of the DDR. η represents the actual DDR bandwidth loss ratio and ρ is the utilization of burst length. B_{ffi} , equals to B_{p1} , can be calculated by F_{p1} and W . F_{p1} represents the clock frequency of FFT processing module and W is the word length of SAR data.

4 Experiments and Results

To validate the effectiveness of the proposed design methodology, we establish a prototype using a single NETFPGA-SUME, which consists of one Xilinx XC7VX690T FPGA, a pair of 4 GB DDR3 and 36 MB QDRII, etc. A sub-board with zynq7020 SoC is connected to the NetFPGA-SUME via the FMC interface. The SoC sub-board is only used as the data I/O of the entire system by its Gigabit Ethernet, and it can be used for some software expansion functions. The photo of the hardware system is shown in Fig. 4. In this paper, entire design test and CS imaging processing are implemented by the NetFPGA-SUME.

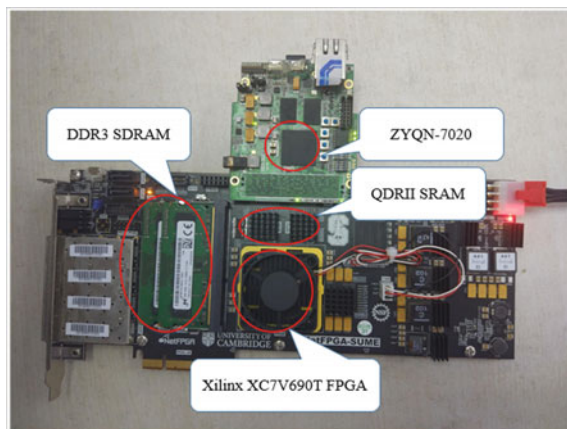


Fig. 4. Photo of hardware system.

Therefore, the system model includes three kinds of memory bandwidth, DDR3, QDRII, and internal RAM. QDRII is used as the writing buffer and RAM is used as the reading buffer. As the source of DDR is limited, we design the system based on the storage first method. The storage first solution (5) is changed as follows:

$$\begin{aligned}
 & \max [B_{ddr}, mB_{qdr}, kB_{ram}]_{\min} \\
 & \text{s.t. } [B_{ddr}, mB_{qdr}, kB_{ram}]_{\min} \leq [nB_{fft}]_{\min}
 \end{aligned} \tag{8}$$

Both range and azimuth processing can use (8) to analyze. The prototype design with the highest storage bandwidth utilization can be achieved by confirming the appropriate m , k , n . Table 1 shows several important clock frequencies in the system. F_{ddr_io} and F_{qdr_io} represent the clock frequency of the DDR I/O and QDR I/O. The word length of both RAM and FFT is 64 bit which equals to the single precision floating point complex form of SAR raw data.

Table 2 shows the summary of parameters for parallel processing design where the data outside the brackets is the constraint result obtained from (8) and the determined values of parameters are inside the brackets. Following our design, the operation of

Table 1. Basic clock frequencies of the system implementation.

Parameter	Parameter value
FPGA main frequency	200 MHz
F_{ddr-io}	400 MHz
F_{ram}	200 MHz
F_{qdr-io}	500 MHz
F_{fft}	150 MHz

Table 2. Summary of the parameters for parallel processing design.

Parameter	Parameter value	
	Range operation	Azimuth operation
B_{ddr}	4 GB/s (4 GB/s)	4 GB/s (4 GB/s)
B_{qdr}	4.5 GB/s (4.5 GB/s)	4.5 GB/s (4.5 GB/s)
B_{ram}	1.6 GB/s (3.2 GB/s)	1.6 GB/s (1.6 GB/s)
B_{fft}	1.2 GB/s (1.2 GB/s)	1.2 GB/s (1.2 GB/s)
m	$\geq 1(1)$	$\geq 1(1)$
k	$\geq 3(2)$	$\geq 3(4)$
n	$\geq 3(4)$	$\geq 3(4)$

range and azimuth has the same results. One thing to be noted is that Bram of range operation is set to 3.2 GB/s indeed. That is because the data in once burst eight of DDR contain four azimuth lines and two range lines, and then we increase the word length of RAM in range operation. The number of this memory is halved as well.

In order to pipeline the computing operation, reading and writing buffers use ping-pong structure which has double consumption of storage resources. Four lines of data as a group are in ping or pong [5]. FFT operations can be operated near pipeline with four parallelism, thus the system efficiency is mainly determined by DDR bandwidth.

The FPGA resource occupation is shown in Table 3. According to the above timing analysis, it takes 6.63 s for the system hardware to process Chinese GaoFen3 25 km width, 5 m resolution Stripmap SAR raw data with $16384 * 16384$ data granularity. Table 4 shows a comparison between the proposed parallel system and some previous works. Compared with [1, 5, 10, 11], and taking into account the data granularity

Table 3. Resource occupation of FPGA.

Parameter	Value
Number of slice registers	136730 (15.78%)
Number of LUTs	107298 (24.77%)
Number of RAMB36s/FIFOs	660 (44.90%)
Number of DSP48s	378 (10.50%)

Table 4 Comparison with previous works

Works	Year	Schemes	Data granularity	Time (s)
Proposed	2018	FPGA	16384 * 16384	6.63
[10]	2017	FPGA	16384 * 16384	10.6
[5]	2017	FPGA + ASIC	16384 * 16384	12.1
[1]	2016	FPGA + Microprocessor	6472 * 3328	8
[11]	2008	Multi-DSP	4096 * 4096	13

processed, the proposed system undoubtedly shows advantages in both processing time and power consumption. Figure 5 shows the final imaging result with reliable quality of the proposed system.

**Fig. 5.** Imaging result of the system hardware.

5 Conclusion

In this paper, in order to improve the real-time performance of SAR imaging process, a balanced and high-efficiency 2-D data access technology is proposed. With cross-mapping methods and CS hardware implementation model, the feasible and efficient prototype is implemented. Four FFT cores are instantiated and operated in parallel. The platform requires 6.63 s to focus 25 km width, 5 m resolution Stripmap SAR raw data with a granularity of 16384 * 16384. The proposed method not only can be designed and implemented quickly, but it can also achieve high processing efficiency.

Admittedly, this technology is suited for other CS-based modes and other granularity imaging. With the development of anti-radiation reinforcement technology and system fault-tolerant technology, the proposed framework is undoubtedly expandable and feasible for potential spaceborne real-time SAR imaging processing.

Acknowledgements. This work was supported by the National Natural Science Foundation of China under Grant 31727901, and the Hundred Leading Talent Project of Beijing Science and Technology under Grant Z141101001514005.

References

1. Lou, Y., Clark, D., Marks, P., Muellerschoen, R.J., Wang, C.C.: Onboard radar processor development for rapid response to natural hazards. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **9**, 2770–2776 (2016)
2. Langemeyer, S., et al.: A compact and flexible multi-DSP system for real-time SAR applications. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium. IGARSS*, vols. I–VII, pp. 1657–1659. IEEE (2003)
3. Zhang, F., Li, G., Li, W., et al.: Accelerating spaceborne SAR imaging using multiple CPU/GPU deep collaborative computing. *Sens.* **16**(4), 494 (2016)
4. Cumming, I.G., Wongm, H.C.: Digital processing of synthetic aperture radar data: algorithms and implementation. *International Radar Conference*, vol. 1, pp. 168–175
5. Yang, C., Li, B., et al.: A spaceborne synthetic aperture radar partial fixed-point imaging system using a field-programmable gate array—application-specific integrated circuit hybrid heterogeneous parallel acceleration technique. *Sens.* **17**, June 2017
6. Zhou, J.: Fine-grained algorithm and architecture for data processing in SAR applications. National University of Defense Technology, Oct 2010
7. Gu, C., Chang, W., Li, X., et al.: Multi-core DSP based parallel architecture for FMCW SAR real-time imaging. *Radio Eng.* **2015**, 1084–1090 (2015)
8. Liu, X., Xie, Y., Zhao, B., Chen, H., Yang, C.: A SAR matrix transpose method of efficiency balanced two dimensional access. *Acta Electron. Sin.* **44**(1), 33–38 (2016)
9. Zhang, Q.: System design and key technologies of the GF-3 satellite. *Acta Geod. Cartogr. Sin.* (2017)
10. Li, B., Shi, H., Chen, L., et al.: Real-time spaceborne synthetic aperture radar float-point imaging system using optimized mapping methodology and a multi-node parallel accelerating technique. *Sens.* **18**(3), 725 (2018)
11. Desai, N.M., Kumar, B.S., Sharma, R.K., Kunal, A., Gameti, R.B., Gujratty, V.R.: Near real time SAR processors for ISRO’s multi-mode RISAT-I and DMSAR. In: *Proceedings of 7th European Conference on Synthetic Aperture Radar*, pp. 1–4, 2–5 June 2008



Research on the Realization of Excitation Signal SPWM of Radar Synchronous Motor Based on DSP

Liang Zhao^(✉) and ShiHui Zheng

Civil Aviation Flight University of China GuangHan, Sub-College, Guanghan,
China

30375838@qq.com

Abstract. To generate the excitation signal of radar synchronous motor which can obtain the azimuth and elevation of radar antenna, high-speed DSP was used to generate SPWM by the arithmetic of asymmetry rule sampling. The advantage, theory, and mathematic model of SPWM were deeply analyzed. Based on TMS320LF2407A, the process of obtaining SPWM was given. Finally, the test waves were analyzed. The result of experiment shows that it has advantages of high precision, real time, and simple structure to use DSP on the generation of SPWM.

Keywords: Synchronous motor · DSP · SPWM · Asymmetry rule sampling algorithmic

1 Introduction

There are many traditional methods for measuring angle of radar antenna. According to different sensors, it can be divided into potentiometers, brush contact encoders, photoelectric encoder, AC micromotor, and so on. With the development of microelectronic technology, it is more and more popular to adopt inductive AC micromotor, synchronous motor, and all-electronic solid-state synchronous motor. This method not only has high reliability, but also has strong anti-interference performance. Moreover, it has the advantages of high precision, real time, simple structure, convenient, long service life, maintainability, etc [1]. Synchronous motor requires an AC excitation voltage signal, which usually uses frequency at 50, 60, and 400 Hz [2]. SPWM can greatly reduce the harmonic components which will get a satisfactory driving effect. SPWM is the abbreviation form of sinusoidal pulse width modulation, a method to change the pulse wave of equal width to the pulse wave of gradually changing width according to the regulation of sine.

At present, the methods of producing SPWM can be divided into two categories: hardware and software methods. It is difficult to improve the circuit by using the hardware method, such as generation SPWM waves by analog and digital mixing circuits. Using software method may obtain the lowest cost of circuit, which is

generated by real-time calculation. It requires a very high speed of the controller, and DSP is the most cost-effective controller to meet this requirement. TMS320LF2407A is a DSP chip especially used for industrial control in TI Company. By using the DSP, not only the excitation signal of the self-aligning machine can be generated in real time, but also the precise control of the relevant motion system can be realized.

2 SPWM Principle

The principle of generating positive SPWM by software method is using a set of isosceles triangle wave compared with sine wave, and wave of the setting isosceles triangle is called carriers, while sine wave is called modulated wave. As shown in Fig. 1, the time when the carrier and modulation waves intersect acts as the “on” or “off” moments of the switch. DSP calculates the corresponding value of the comparator in real time, a pulse width is obtained according to the sine law of change of pulse modulation waveform.

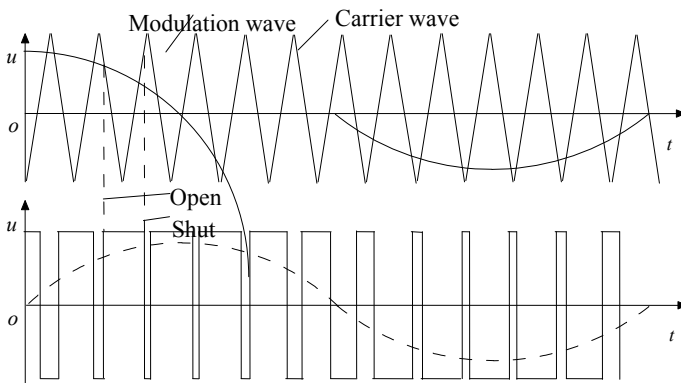


Fig. 1. SPWM wave generation method

3 Mathematical Model

The real-time calculation of SPWM signals needs a mathematical model. There are various ways to build mathematical model, such as harmonic elimination method, equal area method, sampling SPWM, and series methods derived from them. Sampling SPWM method varies from natural sampling, symmetric regular sampling to asymmetric regular sampling [7]. The waveform generated by the asymmetric rule sampling method is closer to sine wave than other methods. This paper focuses on asymmetric regular sampling SPWM method, generating SPWM wave by asymmetric rule sampling method as shown in Fig. 2.

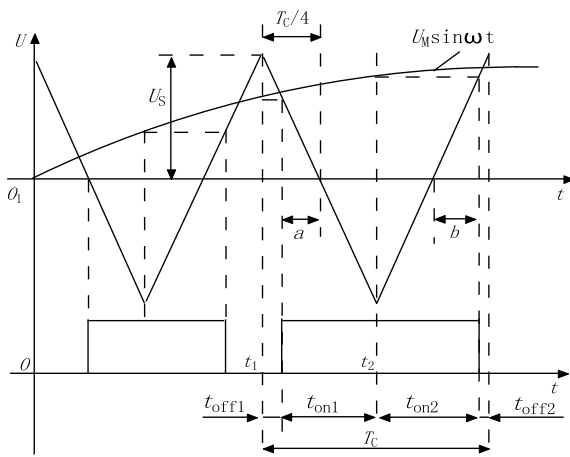


Fig. 2. Generating SPWM wave by asymmetric rule sampling method

As shown in Fig. 2, U_S is the peak value of the triangular carrier, U_M is the amplitude of the modulated wave, and T_C is the period of the triangular carrier. Sine and triangle waves create two intersections at the waist. In triangular wave, sampling at the vertex and bottom of symmetry axis, further improves the similarity between step wave and sine wave. And, t_1, t_2 are the two sampling times, determine the “open”, “close” time of the SPWM wave, namely, $t_{on1}, t_{on2}, t_{off1}$, and t_{off2} .

In Fig. 2:

$$t_{off1} = \frac{T_C}{4} - a, \quad t_{on1} = \frac{T_C}{4} + a$$

$$t_{on2} = \frac{T_C}{4} + b, \quad t_{off2} = \frac{T_C}{4} - b$$

According to the triangular similarity:

$$\frac{a}{\frac{T_C}{4}} = \frac{U_M \sin \omega t_1}{U_S} \tag{1}$$

$$\frac{b}{\frac{T_C}{4}} = \frac{U_M \sin \omega t_2}{U_S} \tag{2}$$

According to Equality (1) and Equality (2) generation into $t_{off1}, t_{on1}, t_{on2}$, and t_{off2} , we get

$$t_{on1} = \frac{T_C}{4} (1 + M \sin \omega t_1) \tag{3}$$

$$t_{on2} = \frac{T_C}{4}(1 + M \sin \omega t_2) \quad (4)$$

In Equality (3) and Equality (4), $M = U_M/U_S$ is the ratio of the sinusoidal peak to the triangular peak. It is called modulation degree and its value ranges between 0 and 1. The higher the setting is, the smaller the harmonic amplitude, so more effective the harmonic suppression.

As shown in Fig. 2, sampling time is divided into double sampling and odd sampling, the two sampling times are given below:

$$t_1 = \frac{T_C}{2}k \quad (k = 0, 2, 4, \dots, 2N - 2) \quad (5)$$

$$t_2 = \frac{T_C}{2}k \quad (k = 1, 3, 5, \dots, 2N - 1) \quad (6)$$

The ratio of triangulation frequency to sine wave frequency is given below:

$$N = \frac{fc}{f} = \frac{1}{T_C f} \quad (7)$$

and

$$\omega t_1 = 2\pi f t_1 = \frac{\pi k}{N} \quad (k = 0, 2, 4, \dots, 2N - 2) \quad (8)$$

$$\omega t_2 = 2\pi f t_2 = \frac{\pi k}{N} \quad (k = 1, 3, 5, \dots, 2N - 1) \quad (9)$$

Equality (8) and Equality (9) is generated into Equality (3) and Equality (4), and we get

$$t_{on1} = \frac{T_C}{4} \left(1 + M \sin \frac{\pi k}{N}\right) \quad (k = 0, 2, 4, \dots, 2N - 2)$$

$$t_{on2} = \frac{T_C}{4} \left(1 + M \sin \frac{\pi k}{N}\right) \quad (k = 1, 3, 5, \dots, 2N - 1)$$

Therefore, the pulse width of SPWM wave is obtained as

$$t_{on} = t_{on1} + t_{on2}$$

4 DSP Realize

TMS320LF2407A DSP has two event managers and each event manager has two timers, three comparison units, three capture units, and an incremental photoelectric encoder interface. We can use the periodic value of the timer cycle register and the comparison value of the comparator to implement the SPWM. The high integration of DSP can minimize the use of peripheral circuits and reduce the cost.

In the design, a half sine of 800 Hz is first generated, and then a directional signal is added to resolve the first half cycle and the latter half cycle of the sinusoidal signal, thereby generating a 400 Hz SPWM waveform.

The carrier frequency can be set according to the system itself. In this case, the frequency of modulation wave is 400 Hz, and the system clock frequency is 40 MHz. The modulation M is chosen to be 0.9. Using timer 1 and comparison unit 3 to produce SPWM, the pulse width is output by the pin CMPR3, and the direction signal is output by the pin IOPB2.

The program consists of a main program and a timer underflow interrupt service program. The main program is a loop, and waits for a period interrupt of the timer 1 to occur. A comparison value of the comparator in the next timer period is calculated in the timer underflow interrupt service program, so that the pulse width is determined and a direction signal is given.

In order to improve the calculation speed, the sinusoidal value in the calculation is obtained by the method of lookup table, and the sine value of 0–180° is provided in Table [8]. Interrupt program setup steps are shown in Fig. 3:

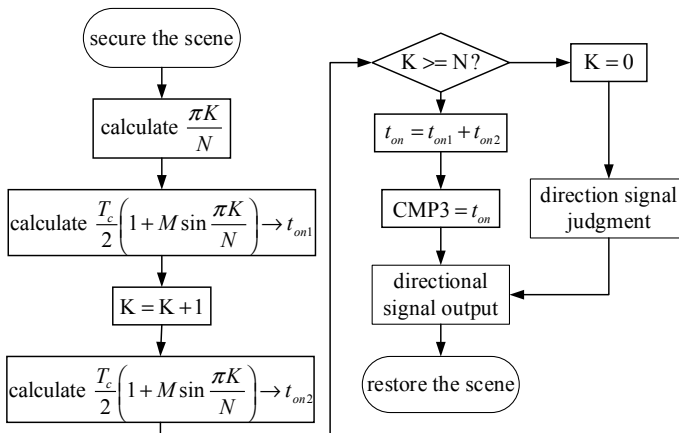


Fig. 3. SPWM generation step

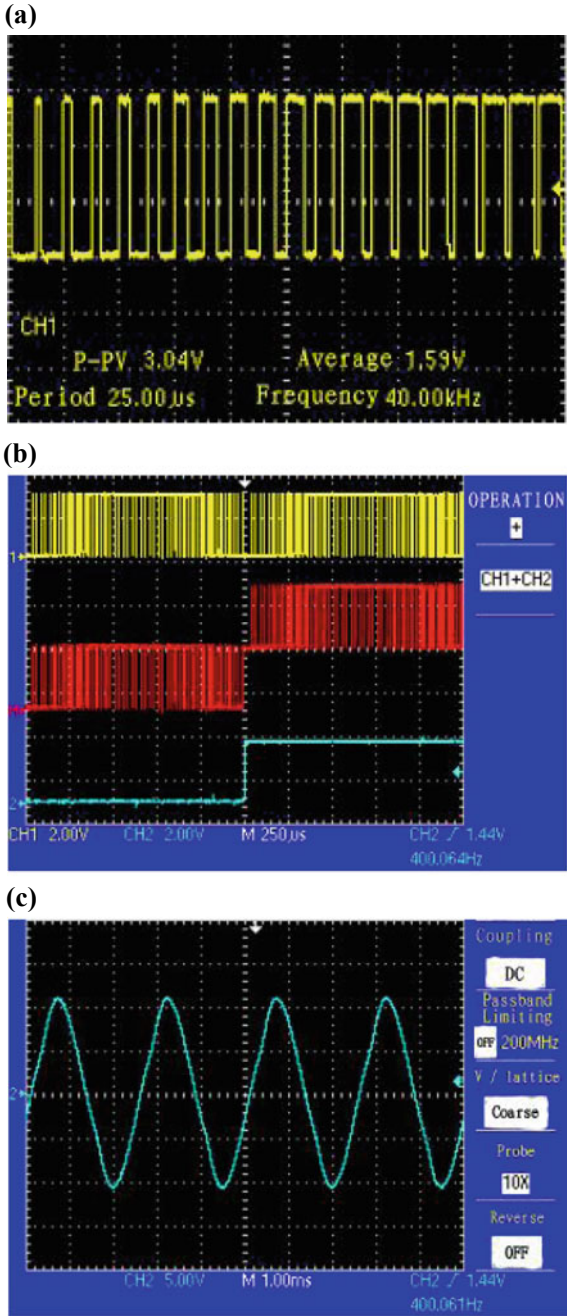


Fig. 4 a 800 Hz SPWM wave. b SPWM waveform with directional signal. c Filtered waveform

5 Results Analysis

The output waveform of 800 Hz SPWM is shown in Fig. 4a. The frequency of the carrier is 40 kHz. In Fig. 4b, the top of the waveform is 800 Hz SPWM, and below is the direction signal, which can be seen at a frequency of 400 Hz. In the middle of the picture is the 400 Hz SPWM with directional signal. Figure 4c is the sine wave shape obtained by filtering. The sine wave period is 400.061 Hz, which meets the design requirements.

6 Conclusion

Research shows that with TMS320LF2407A as the core, and using the asymmetric rule to generate synchronous excitation signal SPWM, we can get a high precision result. The high-speed operation ability of DSP ensures the good real time. Highly integrated structure results to the simplest peripheral circuit, so the system structure is simple. Using DSP to generate SPWM wave is convenient for the development of the system.

References

1. Pang, H.: Research on the Measuring Method of Shaft Angle of Self-Tuning Machine Based on DSP. Shandong University, Jinan (2006)
2. He, J.: Modern Weather Radar. Chengdu University of Information and Technology (2004)
3. Creating a sine modulated PWM signal using the TMS320F240 EVM, Application Report: SPRA411, Revised. Jan 2009
4. TMS320LF/LC240xA DSP controllers reference guide: system and peripherals. Literature Number: SPRU357B, Revised. Dec 2001
5. Procesador digital de señal DSP: TMS320LF240x: arquitectura y aplicaciones. Universidad Politécnica de Valencia (2003)
6. Zhu, J.U., Zhang, B.T.: Research on mixed programming with C language and assemble language based on TMS320LF240x. Comput. Inf. Technol. (2007)
7. Huang, A., Cheng, D.: Applied research on the interrupt of TMS320LF240x series DSP. J. Hubei Automot. Ind. Inst. (2005)
8. Wang, X., Wang, L.: DSP Control of Motor. Beijing University of Aeronautics and Astronautics Press, Beijing (2004)



Sensitivity of Radar Variables to Signal-to-Noise Ratio in Dual-Polarization Weather Radar

Xiao-yi Wang^(✉) and De-bin Su

School of Electronic Engineering, Chengdu University of Information
Technology, Chengdu 610225, China
Xiaoyi.Wang_cn@outlook.com

Abstract. Dual-polarization Doppler weather radar variables such as differential reflectivity (Z_{DR}), differential propagation phase shift (Φ_{DP}), specific differential phase (K_{DP}), and the co-polar correlation coefficient (ρ_{HV}), which are closely related to the characteristic of hydrometeors. The precision of these variables determines the accuracy for quantitative precipitation estimation (QPE) and hydrometeor classification. By analyzing the relationship between the polarimetric radar variables and the signal-to-noise ratio (SNR), it is found that when the SNR decreases to a certain threshold, polarimetric variables become more sensitive to the decrease in the SNR, and the precision of these variables become deteriorated with the decrease of the SNR. This threshold usually is closely related to the radar performance, for the X-band dual-polarization radar used in this paper, the SNR threshold is about 20 dB. The results of this study can provide some reference for radar system performance evaluation and data quality control.

Keywords: Dual-polarization Doppler weather radar · The Signal-to-Noise ratio · Radar performance

1 Introduction

Dual-polarization Doppler weather radar is currently the most advanced weather radar. Unlike conventional Doppler weather radars, by transmitting and receiving horizontally and vertically polarized electromagnetic waves, Dual-polarization Doppler weather radar can obtain the polarimetric radar variables, which closely related to the properties of clouds and precipitation, namely differential reflectivity (Z_{DR}), differential propagation phase shift (Φ_{DP}), specific differential phase (K_{DP}) and the co-polar correlation coefficient (ρ_{HV}). These radar variables can be effectively used in hydrometeor classification [1, 2].

It requires high data quality when applying dual-polarization Doppler weather radar to the hydrometeor classification. Some studies have indicated [3, 4] that in order to ensure radar data quality, compared to the reflectivity Z_H , the polarimetric radar variables: Z_{DR} , Φ_{DP} , K_{DP} , and ρ_{HV} require a high signal-to-noise ratio (SNR) of the radar echo signal. To enhance the SNR, one of the most effective measures is to

increase the radar transmission power, but the increase can be limited due to the current technical condition. Usually, the standard of transmit power used in China is: 750 kW for S-band, 250 kW for C-band, and 80 kW for X-band.

Based on the performance parameters of a domestic X-band dual-polarization weather radar, this paper discusses the sensitivity of polarimetric radar variables to the SNR and provides some reference for future radar performance evaluation and data quality control.

2 Instrument

The dual-polarization weather radar (714XDP) studied in this paper was developed by the Cold and Arid Regions Environmental and Engineering Research Institute of the Chinese Academy of Sciences and the State-owned No. 784 Radar Research Institute. The frequency band of this radar is X-band, and the maximum detection range is 150 km. The radar can acquire polarimetric parameters are Z_H , Z_{DR} , Φ_{DP} , K_{DP} , and ρ_{HV} ; Radar scan mode includes PPI, RHI, and volume scanning, more radar performance parameter details can be seen in Table 1.

Table 1. Main performance parameters of the 714XDP Radar

No.	Parameter	Content
1	Antenna diameter	2.4 m
2	Beam width	0.98° (H), 0.98° (V)
3	Cross polarization level	≤ 32 dB
4	Antenna gain	45.4 dB
5	Scan mode	PPI, RHI, VCP
6	Polarization type	Linear H and V
7	Frequency	9370 ± 20 MHz
8	Peak power	81 kW
9	Pulse width	1 μs
10	PRF	500–2000 Hz(1 μs)
11	Elevation/Azimuth resolution	0.3°/1°
12	Minimum received power	−106 dBm (1 μs)
13	Bin number	1000
14	Range bin length	150 m

3 The Polarimetric Radar Variables and SNR: Sensitivity Analysis

According to the performance parameters of the 714XDP radar, Fig. 1 shows the scatter plots of the SNR and the polarimetric radar variables Z_H , Z_{DR} , K_{DP} , and ρ_{HV} . From Fig. 1, it can be found that the accuracy of the polarimetric radar variables is closely related to the SNR. In Fig. 1a, the reflectivity Z_H exhibits a clear linear relationship with the SNR and decreases with SNR decrease. When the SNR is less than

15 dB, the Z_H become decentralized, but it still maintains a clear linear relationship with the SNR. However, for the polarimetric variables Z_{DR} , K_{DP} , and ρ_{HV} , for the high SNR, their values are mostly concentrated in the range of meteorological echo (Table 2). And when the SNR is lower than about 20 dB, the concentration of these value begins to decline, and more points start to diffuse toward the maximum and minimum values. In terms of data accuracy, this means that the error of the data increases.

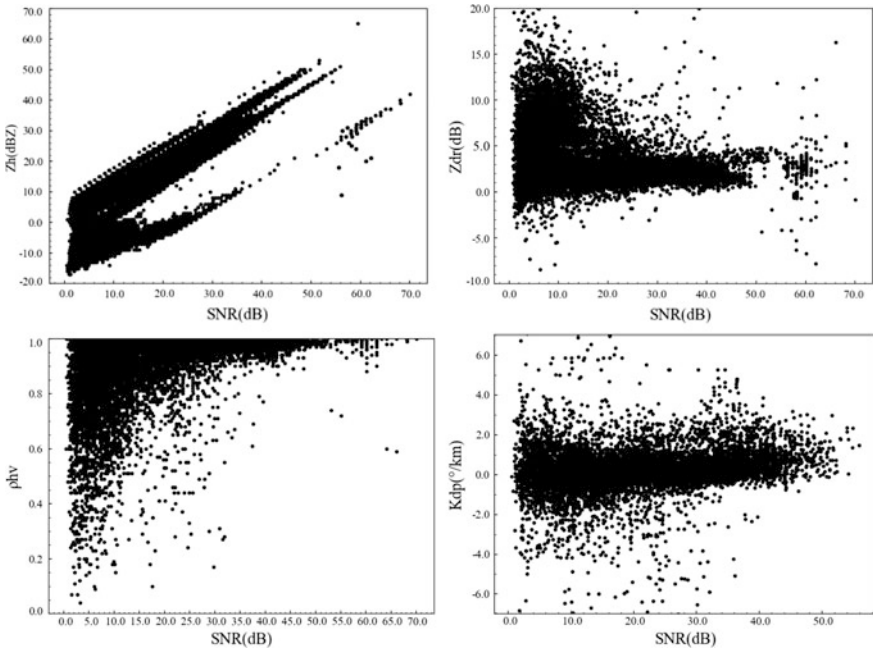


Fig. 1. Relationship between SNR and Z_H , Z_{DR} , ρ_{HV} , K_{DP} of 714XDP radar

Table 2. The range of polarimetric variables for meteorological echoes

	Z_{DR} (dB)	ρ_{HV}	K_{DP} ($^{\circ}/\text{km}$)
Max num	4.0	1.00	3.0
Mini num	-2.0	0.80	0.0

The reason for this result is mainly due to the fact that the calculation of polarization parameters is related to radar echo signal power and noise power [5]. When the SNR is higher, which means the echo signal is higher, the fluctuation of the noise has little influence on the calculation of the polarization parameter, so data obtained is more stable at this time, however, when the noise increases to some extent, which means the SNR decreases. With the decrease of SNR, the influence of the noise power on the

calculation of the polarization parameters cannot be neglected. As a result, the calculation result of the polarization data is unstable and the data error increases. For reflectivity Z_H , its value is only affected by the horizontal channel gain and noise power of the polarization radar, therefore the SNR has little effect on it. Moreover, we have also analyzed the sensitivity of the texture parameter $SD(\Phi_{DP})$ to the SNR (Fig. 2), the texture parameter $SD(\Phi_{DP})$ is defined in Eq. (1) in the Eq. (1), the average “ave(Φ_{DP})” is calculated by a movefilter window of 1.2 km, the “ $N_{\Phi_{DP}}$ ” indicates the number of range bins used for the move filter (the number used in this paper is 8). The $SD(\Phi_{DP})$ denotes that the degree of fluctuation of Φ_{DP} within a certain distance in a ray, it plays an important role in distinguishing between meteorological and non-meteorological echoes.

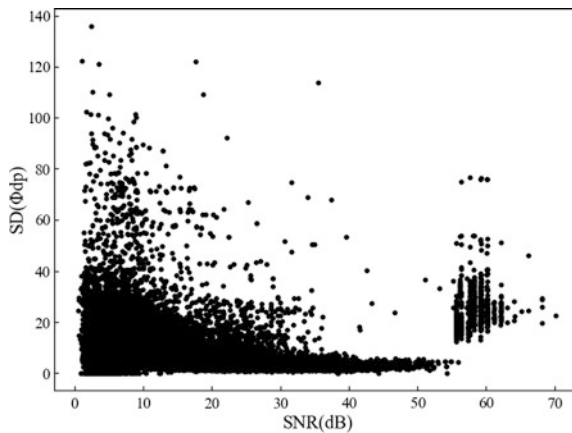


Fig. 2. Relationship between the texture parameter $SD(\Phi_{DP})$ and SNR of the 714XDP radar

$$SD(\Phi_{DP}) = \sqrt{\frac{\sum_{bin=0}^7 [\Phi_{DP} - \text{ave}(\Phi_{DP})]^2}{N_{\Phi_{DP}}}} \tag{Eq.1}$$

From Fig. 2, it can be found that as the SNR decreases, $SD(\Phi_{DP})$ shows a significant trend of increase, and its uncertainty becomes raised. Same to other polarimetric variables, When the SNR higher than about 20 dB, $SD(\Phi_{DP})$ is stable and concentrated within 10 degrees, when the SNR below about 20 dB, the increment of $SD(\Phi_{DP})$ indicates that the variation of the Φ_{DP} along the detection direction becomes higher. The main causes of this variation are: the outer noise power variation of the radar system and the phase noise which arised from the inside of the radar system.

4 Conclusions and Discussions

In this paper, based on the data of the 714XDP polarization radar, the relationship between radar detection variables Z_H , Z_{DR} , ρ_{HV} , K_{DP} and the signal-to-noise ratio of radar echo signal (SNR) is briefly shown and the texture parameter $SD(\Phi_{DP})$ of the polarimetric parameter Φ_{DP} is analyzed too. The result shows that for the 714XDP polarization radar when the SNR is less than about 20 dB, the radar polarimetric variables become sensitive to the change of SNR, and the data accuracy will gradually be deteriorated with the decrease of SNR, then radar data become unreliable. The value of the SNR threshold that causes the accuracy of polarization data to begin to degrade is mainly related to the performance of the radar system receiver, if the radar receiver performance is stable, then the SNR threshold may be lower than the threshold of the 714XDP radar. However, it means the stability of the radar receiver is poor and needs to be adjusted when the SNR threshold is too large.

References

1. Aydin, K., Seliga, T.A., Balaji, V.: Remote sensing of hail with a dual linear polarization radar. *J. Clim. Appl. Meteorol.* **25**(10), 1475–1484 (1986)
2. Straka, J.M., Zni, D.S., Ryzhkov, A.V.: Bulk hydrometeor classification and quantification using polarimetric radar data: synthesis of relations. *J. Appl. Meteorol.* **39**(8), 1341–1372 (2000)
3. Lei, L., Zhang, G., Doviak, R.J., et al.: Multilag correlation estimators for polarimetric radar measurements in the presence of noise. *J. Atmos. Ocean. Technol.* **29**(6), 772–795 (2012)
4. Lim, S., Cifelli, R., Chandrasekar, V., et al.: Precipitation classification and quantification using X-band dual-polarization weather radar: application in the hydrometeorology testbed. *J. Atmos. Ocean. Technol.* **30**(9), 2108–2120 (2013)
5. Zhao, G., Wan, Z., Jia, W., et al.: Research on effective detection range of dual-polarization weather radar. *J. Plateau Meteorol.* **35**(1), 244–250 (2016)



Analysis of Characteristics of Atmospheric Structure Constant of Refractive Index Based on Wind Profiler Radar in Precipitation

Yating Li^(✉), Debin Su, and Xingang Fan

Chengdu University of Information Technology, No. 24 Block 1, Xuefu Road,
Chengdu, China
yatinglil1993@qq.com

Abstract. The atmospheric structure constant of refractive index is an important parameter to describe the turbulent characteristics of the atmosphere. Using the ROBS data obtained by the Sichuan Provincial Meteorological Bureau in the Yanyuan, Dayi, Xinjin, Xindu, and Longquanyi districts, the real-time sampling of ROBS data combined with the hourly precipitation data provided by conventional meteorological observation stations was used for statistical analysis. Study the change trend of atmospheric structure constant of refractive index in the precipitation process in different regions. The results show that the convective movement during the precipitation process and the drag effect of the precipitation particles on the surrounding atmosphere, with the occurrence and development of the precipitation process, atmospheric structure constant of refractive index has a significant increasing trend, indicating the significance of precipitation. The vertical velocity and C_n^2 are consistent with the change of precipitation, which can better reflect the change of precipitation intensity. The C_n^2 value corresponding to the maximum precipitation time is two to five orders larger than that before and after the occurrence of precipitation.

Keywords: Atmospheric structure constant of refractive index · Wind profile radar · Precipitation · Characteristic analysis

1 Introduction

Random turbulent movements in the atmosphere cause random changes in the refractive index of the atmosphere. The atmospheric structure constant of refractive index is an important parameter that represents the atmospheric optical turbulence intensity, and its size is proportional to the backscattered energy of the wind profile radar to detect turbulence [1]. The relationship between the spatial and temporal variations of atmospheric turbulence intensity and the surrounding environmental conditions is very complex. In the atmospheric boundary layer affected by the complex surface of the earth and various weather conditions, C_n^2 is mainly affected by unstable convection caused by ground thermal radiation. In the free atmosphere, C_n^2 is mainly affected by the temperature gradient and the wind speed gradient [2]. Chun et al. [3] statistically analyzed data from April to May 2014 in the Chengdu area, showing that

the characteristics of C_n^2 of daily changes are obvious and there are differences in the time of change. Sun et al. [4] statistically analyzed a large number of sounding experimental data and found that C_n^2 of Hefei area conforms to lognormal distribution within 25 km, and the magnitude of day–night variation was 0.8.

Wind profile radar is a remote sensing device that uses atmospheric turbulence to scatter electromagnetic waves and other physical quantities to detect atmospheric wind. It can provide a variety of data products based on wind fields. Its basic data products include radial velocity, spectral width, signal-to-noise ratio, horizontal wind direction, horizontal wind speed, vertical velocity, and atmospheric structure constant of refractive index that reflects atmospheric turbulence. Ma et al. [5] used wind profile radar data to study changes in troposphere height and precipitation cloud during the Meiyu period; Muschiniski et al. [6] studied the atmospheric turbulent structure using the power spectrum data of the return signal in the vertical direction of the boundary layer wind profile radar, and analyzed the average of C_n^2 vertical structure.

The research on the atmospheric structure constant of the refractive index of wind profile radar is mainly focused on the height distribution and temperature and humidity analysis. There are few studies on the change characteristics of atmospheric structure the constant of refractive index during precipitation. This paper mainly selects precipitation events in various regions, and studies the change characteristics of atmospheric structure constant of refractive index.

2 Data

The data from this paper are from the ROBS data of the wind profile radar at the Yanyuan, Dayi, Xinjin, Xindu, and Longquanshan stations and the time-to-precipitation data from the China Meteorological Data Network [7] (All data in this paper adopts the world time). Among them, the wind profile radars at the Longquanyi, Dayi, Xinjin, and Xindu sites are TWP3 boundary layer wind profile radars, and the wind profile radars at the Yanyuan site are TWP8-L tropospheric wind profile radars. The radar performance indicators are shown in Table 1. Wind Profiler product data files of ROBS include sample height, horizontal wind direction, horizontal wind speed, vertical wind speed, horizontal reliability, vertical reliability. The high resolution of the low mode product is 60 m, and the high mode product height resolution is 60 m or 120 m. ROBS data is every 5-min interval data.

3 Results and Discussion

This article selects a rainfall process covering Yanyuan, Dayi, Xinjin, Xindu and Longquanyi from October 21 to October 23, 2017. The selected area has the same climate, different terrain. The general characteristics of the terrain in Yanyuan County are surrounded by high-altitude valleys and central hill basins. Generally, the altitude is 2300–800 m, and the highest altitude is 4393 m. It belongs to the subtropical monsoon

Table 1. Boundary layer wind profile radar parameters

Types index	Radar	
	TWP3 wind profile radar	TWP8-L wind profile radar
Working frequency (MHz)	1270–1295	1270–1295
Maximum detection height (km)	≥ 3	≥ 6
Minimum detection height (m)	≤ 100	≤ 100
Wind speed resolution (m/s)	0.2	0.2
Wind direction resolution (°)	0.5	0.5
High resolution	Low mode is 60 m, high mode is 60 m or 120 m	Low mode is 60 m, high mode is 60 m or 120 m
Time resolution	Less than or equal to 3 min for 3 beam operation and 5 min for 5 beam operation	Less than or equal to 3 min for 3 beam operation and 5 min for 5 beam operation
Measurement performance	Wind speed measurement range: 0–60 m/s Wind direction measurement range: 0–360°	Wind speed measurement range: 0–60 m/s Wind direction measurement range: 0–360°

climate zone, with small annual temperature difference and large daily temperature difference. Dayi and Xinjin regions form the main body of the western plains of Sichuan, with cloudy fog and short sunshine. Xindu and Longquanyi are located in the western part of the Sichuan Basin. The landforms mainly consist of mountains, hills, and flat dams with four seasons distinct. Figure 1a, b are the timing profiles of the wind

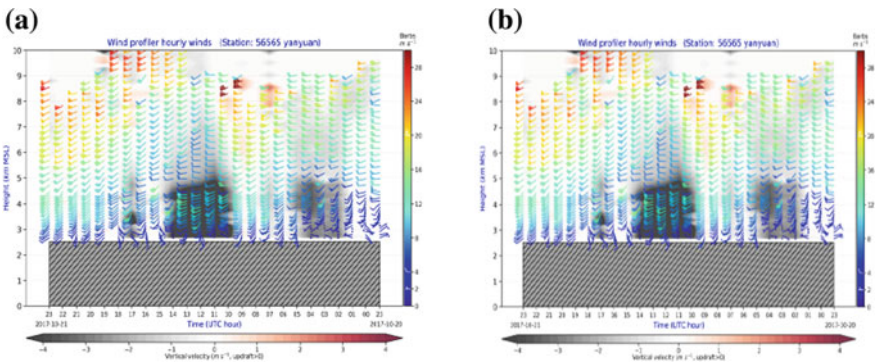


Fig. 1. Timeline of the vertical profile of the wind vector of Yanyuan: **a** October 21 **b** October 22

vector vertical profiles from October 21 to October 22, 2017 in the Yanyuan area respectively.

Analyze C_n^2 by the hourly precipitation corresponding to the vertical wind speed, as shown in Figs. 2 and 3, respectively, the Yanyuan’s 2190 m vertical wind speed profile, the Yanyuan’s 2190 m height hourly precipitation and timing diagram of C_n^2 . As observed in Fig. 1a, the vertical wind speed has a large amplitude at 03:00–09:00 with a range of $-1-5$ m/s, and the wind speed increases rapidly at 09:00–11:00. At the time of 11:00, the maximum value reached 6.7 m/s on the 21st, and the wind speed quickly dropped to -1 m/s from 11:00 to 16:00, and rose first and then decreased from 16:00 to 20:00. The change range was $0-3$ m/s, fluctuates around 0 m/s between 20:00–24:00. Figure 1b is followed by 1a. The amplitude of wind speed increases from hour to hour from 00:00 to 07:00, and the wind speed fluctuates greatly from 07:00 to 10:00. It decreases first and then rises. The maximum value is 4.9 m/s at 10:00, after which it rapidly decreases. From 12:00 to 17:00, the wind speed changes around 0 m/s. After 17:00, the value increases slowly, and the range of change is $0-4$ m/s. Corresponding to the observations in Figs. 2a, b and 3a, b, the changes of C_n^2 in precipitation and vertical wind velocity are basically consistent. The precipitation and it increased with the increase of vertical wind speed. The increase was concentrated at 12:00–14:00 and increased to a maximum at 14:00. The hourly precipitation reached 2 mm, reaching 10^{-8} . On the 22nd, the hourly precipitation decreased by 2 orders of magnitude, and the hourly precipitation remained at about 0.5 mm, staying at about 10^{-12} . The analysis shows that the change trend of vertical wind speed and hourly precipitation at the same height in Xindu are consistent, and the range of change is $10^{-18}-10^{-4}$. It can be seen from the figure that C_n^2 is less sensitive to precipitation than vertical velocity, but has a correlation with precipitation intensity.

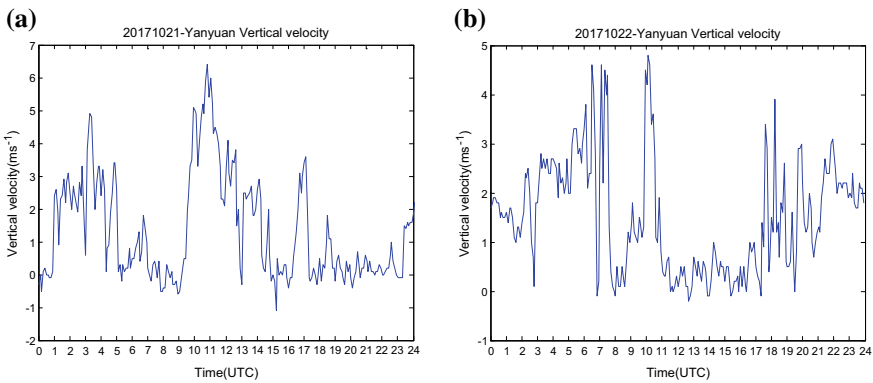


Fig. 2. Timeline of vertical wind speed profile of Yanyuan: **a** October 21 **b** October 22

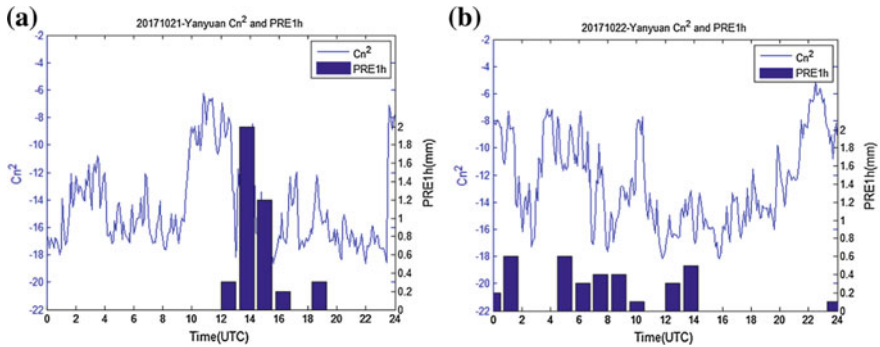


Fig. 3. Hourly precipitation and timing diagram of Yanyuan: **a** October 21 **b** October 22

According to the site classification, the hourly timing chart of precipitation and C_n^2 for Dayi, Xinjin, Xindu, and Longquanyi were made on October 22 respectively. Figure 4a is a timing chart of the site of Dayi. Hourly precipitation at 00:00–04:00 h increases first and then decreases from 0.2 to 0.4 mm. At the same time, it also shows a trend of increase and decreases at the same time. The change range is 10^{-14} – 10^{-8} . As the precipitation decreases, the value also slowly decreases and finally fluctuates around 10^{-14} . Figure 4b is a timing diagram of the Xinjin station. Hourly precipitation at 0–15 h has a relatively large span, which increases first at 00:00–03:00, then decreases at 03:00–06:00, and then starts to increase sharply at 07:00–14:00, increasing to amplitude. 1.4 mm, greatly reduced at 14:00–15:00, dropped to 0.2 mm. Value fluctuations are consistent with changes in hourly precipitation, ranging from 10^{-22} to 10^{-10} . Figure 4c, d is a timing chart for the Xindu and Longquanyi stations. The daily change in hourly precipitation is more consistent with the daily change trend. The maximum hourly precipitation 1.6 mm in the Xindu occurred at 13 o'clock, and the maximum value also appeared at the same time, with an amplitude of 10^{-7} . At hour 0–15, the hourly precipitation of Longquanyi fluctuates at 0.8 mm, the value fluctuates at 10^{-9} , the maximum hourly precipitation appears at 15 o'clock, and the amplitude is 10^{-6} . The hourly precipitation and value decrease greatly after 15 o'clock. The value fluctuates around. In sum, the hourly precipitation coincides with the overall change pattern.

According to the development of precipitation, we selected different rain moments to compare different levels of C_n^2 . Figure 5a shows the time sequence of C_n^2 at the Yanyuan station. The selected time is at 06:00 on the 21st, 13:00 on the 21st, and 04:00 and 20:00 on the 23rd. As shown in the figure, at the same altitude, the rain increases. As C_n^2 increases, the value of C_n^2 at different moments increases with the increase of rain. At 13 o'clock, the hourly amount of precipitation is maximum and the C_n^2 value reaches 10^{-7} . The value of pre-precipitation is similar to the overall size at the end of the precipitation. The 2 km height is a maximum of two orders of magnitude.

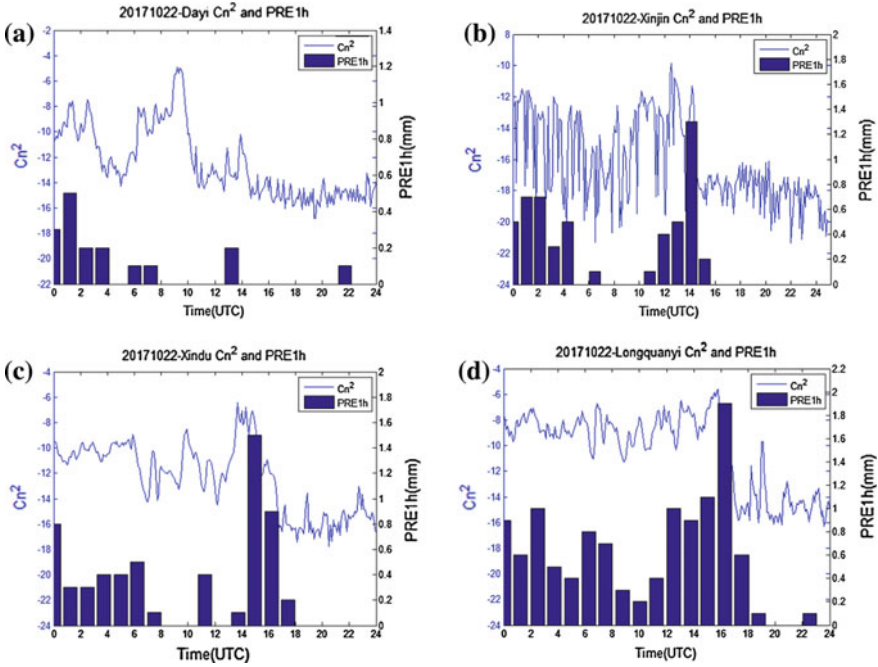


Fig. 4. Hourly precipitation and time series: **a** Dayi Station **b** Xinjin station **c** Xindu station **d** Longquanyi station

Figure 5b is the timing diagram of the C_n^2 moment at Dayi station. The selection time is at 06:00 and 13:00 on the 21st, 11:00 and 18:00 on the 22nd. The C_n^2 value of 13:00 is larger than the other three moments by two to five orders of magnitude. Figure 5c is a timing diagram of C_n^2 moment at Xindu station. The selected time is at 03:00 and 17:00 on the 21st, 18:00, 21:00 on the 22nd, and the same 17 o'clock value is larger than the other three moments by two to five orders of magnitude. With the height increase, there is a clear decrease.

In summary, the analysis of each site, with the occurrence and development of the precipitation process, there is a clear trend of increase, and the distribution of the characteristics of the vertical wind speed, hourly precipitation changes. The C_n^2 value corresponding to the maximum precipitation time is two to five orders larger than that before and after the occurrence of precipitation.

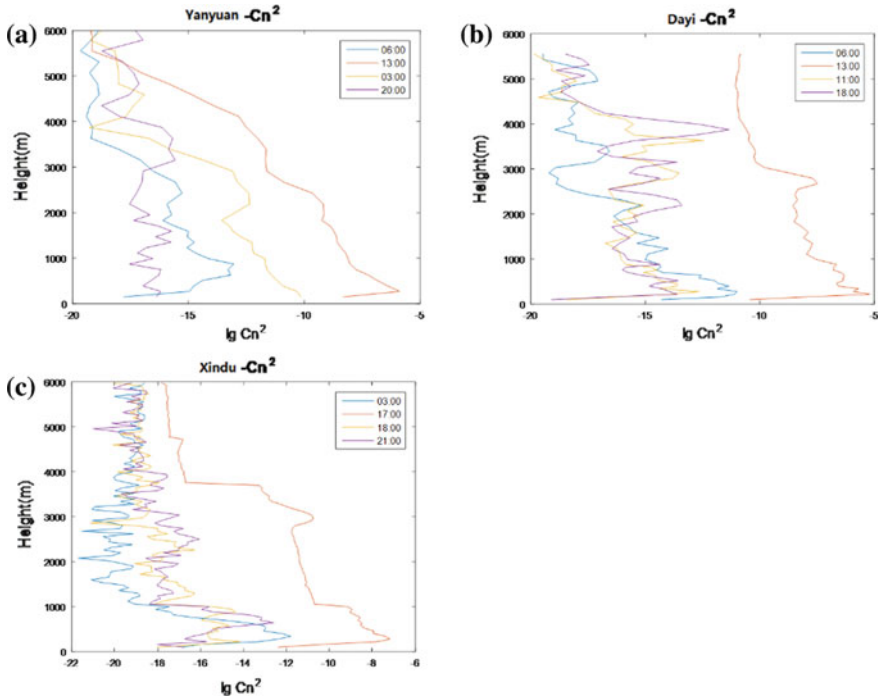


Fig. 5. Time sequence diagram of C_n^2 a Yanyuan station b Dayi station c Xindu station

4 Conclusion

In this paper, through the data analysis of the characteristics of atmospheric refractive index distribution profile in Yanyuan, Dayi, Xinjin, Xindu and Longquanyi, the following conclusions have been drawn: Convective movement in the process of precipitation and the drag effect of precipitation particles on the surrounding atmosphere, with the development of precipitation process, C_n^2 has a significant increase. The vertical velocity and C_n^2 are consistent with the change of precipitation, which can better reflect the change of precipitation intensity. The C_n^2 value corresponding to the maximum precipitation time is two to five orders larger than that before and after the occurrence of precipitation. Due to limited observation data, this article only studied a small number of cases and obtained some preliminary conclusions. In the future, a more comprehensive and in-depth statistical analysis will be conducted around the atmospheric refractive index structure constants to improve the reliability and accuracy of the conclusions.

Acknowledgments. Thanks to Sichuan Province Bureau of Meteorology for providing wind profile data, thanks to Chengdu University of Information Technology, Chengdu University of Geosciences, based on a dense ground-based observation network for nowcasting technology research project (Grant No. KYTZ201529), and Beijing Normal University's mesoscale meteorological observation network for surface rainfall observation and simulation. Project (Grant No. 2015H12) Funding.

References

1. Ecklund, W.L., Carter, D.A., Balsley, B.B., et al.: Field tests of a lower tropospheric wind profiler. *Radio Sci.* **25**(5), 899–906 (1990)
2. Zheng, R., He, P.: Research on the detection of atmospheric refractive index structure constants by wind profile radar. *Chin. J. Atmos. Sci.* **32**(1), 133–140 (2008)
3. Chun, Q., Wu, X., Wang, H., Wang, P.: Statistical analysis of refractive index structural constants in the near-surface of Chengdu area. *J. Atmos. Environ. Opt.* **10**(5), 368–375 (2015)
4. Sun, G., Ningquan, W., Xiao, L.: Statistical analysis of the height distribution of atmospheric refractive index structure constant C_n^2 . *J. Atmos. Environ. Opt.* **6**(2), 000083–88 (2011)
5. Ma, Z., Chen, D., Zhou, W.: Numerical simulation experimental study on the detection of tropospheric wind profiles by VHF radar. *Chin. J. Atmos. Sci.* **9**(2), 113–118 (1985)
6. Muschinski, A., Sullivan, P.P., et al.: First synthesis of wind-profiler signals on the basis of large-eddy simulation data. *Radio Sci.* **6**(12), 1437–1459 (1999)
7. China Meteorological Data Network [DB/OL]. <http://data.cma.cn/site/index.html>



Modeling and Simulation of Auto Parts Production Line Based on Petri Net

Shuqi Jin¹(✉), Shaohua Cui², and Chenglin Zhao¹

¹ Key Laboratory of Universal Wireless Communications, MOE, Beijing University of Posts and Telecommunications, No. 10 Xitucheng Road, Beijing, China

starryhp@163.com

² China Petroleum Technology & Development Corporation, Beijing, China

Abstract. The production efficiency of small and medium auto parts manufacturing enterprises has been restricted by unreasonable production processes arrangement, high equipment load rate, high quantity of work in process (WIP), etc. To cope with these problems, this paper proposes a novel modeling method to analyze production line. The production line model is established by the method that combination of object-oriented Petri net (OOPN) and time transitions. Then the proposed model is validated based on incidence matrix and state equation analysis method, by which the bottleneck of the production line is identified. Simulation results are provided to validate that the proposed method can accurately reflect the production process and the bottleneck.

Keywords: Production line modeling · OOPN · Incidence matrix · Production bottleneck

1 Introduction

The production line is the core of manufacturing, and the manufacturing competitiveness is affected by production cycle, quality and cost of the products. So in order to produce better quality products with lower cost in a short time, modeling is necessary.

At present, there are many methods for manufacturing process modeling, such as IDEF, object-oriented, queuing theory [1–5]. But they have obvious deficiencies in system calculation and performance analysis. Meanwhile, computer simulation software from the earliest common programming language, to the text mode simulation software, to the graphical construction model [6], has been well applied to the manufacturing process simulation [7–9]. However, they lack theoretical analysis.

This paper has both theoretical calculation and simulation verification. It proposes a novel modeling method, then the established model is theoretically analyzed. Finally, the analysis results are verified by simulation.

2 Petri Net

2.1 Basic Petri Net

A basic Petri net system is a triad [2],

$$PN = (P, T, F) \tag{1}$$

where P and T are respectively the place set and the transition set. $F \subseteq (P \times T) \cup (T \times P)$ is a directed arc set to represent the flow of resources in the net.

2.2 Object-Oriented Petri Net

Object-oriented Petri net system is a multivariate group,

$$OOPN = (P, T, F, OP, \&, W, M_0) \tag{2}$$

where OP represents the object net; and $\&$ represents the relationship between the basic net and the object net; W is the set of weights, M_0 is the initial state identity.

Because we just focus on the number of the objects, so for OOPN model, we can use the number to describe the observed objects.

3 OOPN Modeling and Analysis

3.1 Introduction of Camshaft Production Line

The production process about the camshaft production line is shown in Fig. 1.

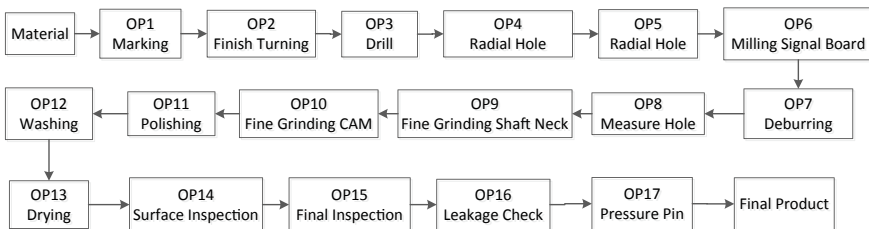


Fig. 1. Production process of camshaft

3.2 Establishing of OOPN Model

According to the production process of the camshaft, the whole process is divided into three stages: the rough machining stage, the finishing stage, and the detection stage. OOPN models of whole and each stage are shown in Figs. 2, 3, 4 and 5. In Fig. 2, P_{01} and P_{02} represent the material place and product place respectively.

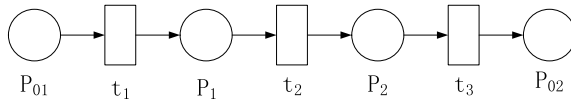


Fig. 2. Whole OOPN model

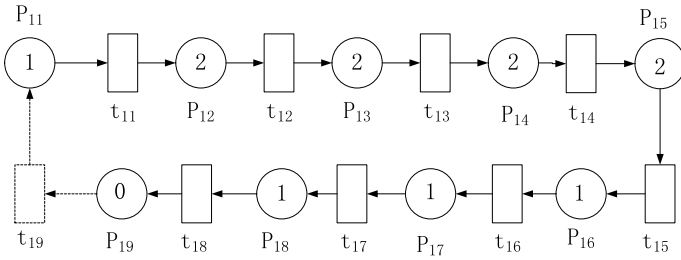


Fig. 3. OOPN model of rouge machining stage

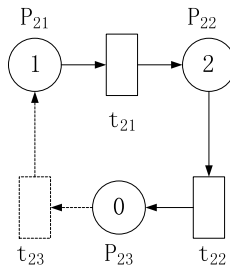


Fig. 4. OOPN model of finish machining stage

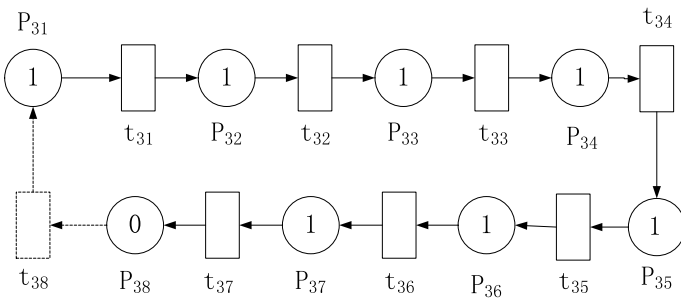


Fig. 5. OOPN model of detection stage

3.3 Validation of OOPN Model

The state equation of the established OOPN model is

$$M = M_0 + A^T X \tag{3}$$

where, A, X, and M represent incidence matrix, transition vector, and the state vectors respectively. If the established OOPN model is suitability, the production system can be returned to the initial state:

$$A^T X = 0 \tag{4}$$

Based on the incidence matrix, the basic solutions of the three stages of camshaft production line are calculated by Eq. (4):

$$x_1 = (1, 1, 1, 1, 1, 1, 1, 1, 1)^T \tag{5}$$

$$x_2 = (1, 1, 1)^T \tag{6}$$

$$x_3 = (1, 1, 1, 1, 1, 1, 1, 1, 1)^T \tag{7}$$

According to the calculation, there are nonzero solutions in the three processing stages, so the OOPN model is suitable, the established models are able to run from the beginning of the production to finish and return to the initial state, which can realize repetitive production cycle.

3.4 Bottleneck Analysis

The production cycle is the time after a series of transitions to return to the initial state, namely,

$$\tau \geq y^T (A^-)^T D x / y^T M_0 \tag{8}$$

where x satisfies the formula (4), is called T invariant, y satisfies Ay = 0, is called S invariant, D = diag(d_j) is the diagonal matrix of each transition time.

According to formula (8), the production cycle of the rough machining stage, the finishing stage, and the detection stage is: τ₁ = 31 s, τ₂ = 26 s, τ₃ = 23 s, respectively. As we can see, the production cycle of rough machining stage is the longest, and the subsequent production stages are impacted, so it is the bottleneck of the system and affects the production efficiency of the whole production line.

To take full advantage of process time, output and tack time data, and to more intuitively understand the working state of equipment on the production line, the next chapter maps the OOPN model to Plant Simulation software.

4 Simulation and Analysis

The places and directed arcs in OOPN model are converted into entity and line of Plant Simulation respectively and set the parameters of entities to represent the transition process. Then simulation time is set to one month. Finally, the statistical results of the working status of each device are shown in Fig. 6.

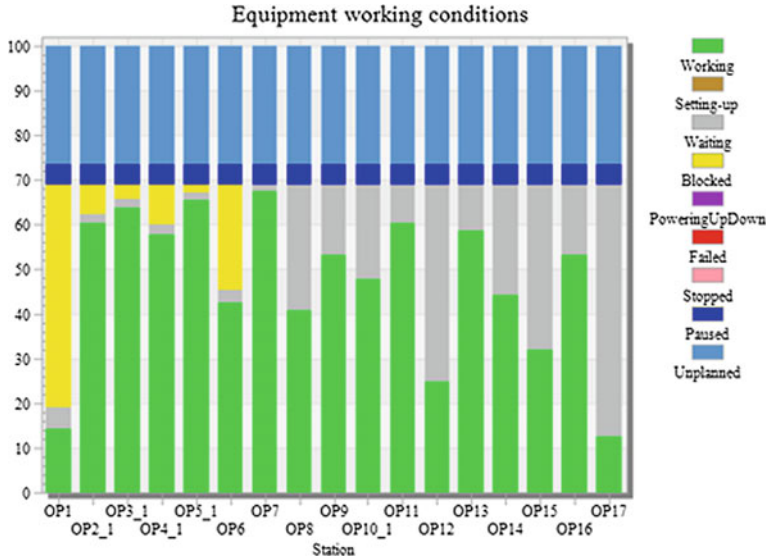


Fig. 6. Working status of each device

As evidently shown in Fig. 6, the load rate of OP7, OP5_1 and OP3_1 is as high as 67.18%, 65.42%, and 67.18% respectively. And for the first six workstations, there are obvious blocking phenomena, particularly the blocking rate of the OP1 process is up to 49.86%. The result indicated that the bottleneck of the whole production line appeared in the rough machining stage.

The result of simulation analysis is the same as that of the modeling analysis, which indicates that the proposed method that integrating OOPN hierarchical modeling with time transition can describe the actual production process well and can find the bottleneck of the production line accurately.

5 Conclusion

In the process of producing auto parts, there is a large difference in the load degree of different equipment, so the production line is imbalanced. In order to solve the problem, this paper takes the camshaft production line as an example, proposing the OOPN hierarchical model method and adding time in transitions. Then use the incidence

matrix and state equation analysis method to prove the accessibility of established OOPN model and to determine the bottleneck of the production line. Finally, the theoretical analysis results of OOPN model verified through simulation analysis. It provides a theoretical basis for improving the production line.

Acknowledgements. This work is supported by Intelligent Manufacturing Integrated Standardization and New Model Application Project (2017ZNZX03).

References

1. Sun, X.: Modeling and Analysis of Aircraft Assembly Line. Nanjing university of aeronautics and astronautics, Nanjing (2011)
2. Wu, Z.: Introduction to Petri Net. China Machine Press, Beijing (2006)
3. Taghinezhad-Niar, A., Javadzadeh, T., Farzinvash, L.: Modeling of resource monitoring in federated cloud using Colored Petri Net. In: IEEE International Conference on Knowledge-Based Engineering and Innovation, pp. 0577–0582 (2017)
4. Jung, M., Kraft, K., Wehn, N.: A new state model for DRAMs using Petri Nets. In: IEEE International Conference on Embedded Computer Systems: Architecture, Modeling, and Simulation, pp. 221–226 (2017)
5. Wang, Q., Wen, L.: Modeling and optimization of aircraft assembly line based on Petri Net. *J. Zhejiang Univ.* **49**(7) (2015)
6. Chaichompoo, O., Thongtak, A., Vatanawood, W.: Transformation of time Petri Net into Promela. In: IEEE International Conference on Telecommunication Systems Services and Applications, pp. 1–5 (2017)
7. Klebanov, B., Mufazalov, A., Myasoedov, I.: Use of plant simulation for improvement technological and business processes of metallurgical manufacture. In: IEEE Chinese Control Conference, pp. 9681–9684 (2016)
8. Klebanov, B.I., Aksenov, K.A., Antonova, A.S., Panteleyeva, Y.S., Krymov, E.A.: Using plant simulation for simulation and optimization of logistics processes. *Mod. Probl. Sci. Educ.* (2014)
9. Li, H., Wang, S.: Simulation study on small and medium sized automobile welding production line based on EM-Plant. *J. Chang. Univ.* **23**(6) (2013)



A Radar Electromagnetic Environment Sensing Method Based on Cyclic Spectral Algorithm

Jurong Hu¹(✉), Yu Zhang¹, Xujie Li¹, Xiaoyong Ni²,
and Evans Baidoo¹

¹ The School of Computer and Information, Hohai University, Nan Jing, China
hujurong@sina.com

² University of Electronic Science and Technology of China, Cheng Du, China

Abstract. In this paper, a radar electromagnetic environment sensing method based on the cyclic spectral algorithm is discussed, which can be used to acquire the spectrum information of radar signals and distinguish them. This paper uses the second-order cyclostationary detection algorithm based on the spectral correlation function (SCF) to obtain the cyclic spectral. The estimation of SCF is and the estimation precision by calculating deviation and variance of SCF are displayed. In the simulation, a scenario of radar electromagnetic environment is presented by transmitting Linear Frequency Modulation signals (LFM) and Amplitude Modulation signals (AM). Simulation results indicate that the cyclic spectral algorithm can not only sense the spectrum information of signals but also judge the type of signal. Therefore, the bandwidth of the interference information can be detected. The simulation results show that this method is highly preferred for radar electromagnetic environment sensing even under low signal-to-noise ratio (SNR) circumstance.

Keywords: Radar electromagnetic environment · Second-order cyclostationary · Detection · SCF · MIMO radar

1 Introduction

Multiple-Input Multiple-Out-put (MIMO) Radar is designed to combat channel fading and improve receiver performance through various methods [1]. Multiple receivers are used to detect targets and sense the complex electromagnetic environment. In this paper, to improve the anti-interference ability of MIMO radar, we consider to detect the range of frequency of the interference by sensing the radar electromagnetic environment. Then waveforms are designed and transmitted by adjusting the range of frequency to avoid the band of interference information.

The electromagnetic environment of MIMO radar is pretty complex for its devise noise and poor prior knowledge. In this paper, cyclostationary detection technique is utilized to sense spectrum information as in the radar communication system almost all signals show a cyclic stabilization [2–6]. The technology of cyclostationary detection is highly preferred for spectrum sensing in low SNR due to its robustness to the

uncertainty of noise. It can also exploit signal spectral feature by identifying spectral correlation peaks in the SCF which includes cyclic and angular frequency domains.

The remainder of this paper is organized as follows. Section 2 describes the cyclostationary feature detection. Section 3 calculates the estimation precision of SCF. Section 4 gives the simulation results. Finally, the paper is concluded in Sect. 5.

2 Second-Order Cyclostationary Detection Algorithm

Define received signal

$$z(t) = \sum_{i=1}^q s_i(t) + n(t) \quad (i = 1, 2, \dots, q) \tag{1}$$

where $n(t)$ represents the noise and $\sum_{i=1}^q s_i(t)$ represent the combination of signals which include LFM, AM, and so on. t and q represents time and the number of signals.

The autocorrelation function $R_{zz}(t, \tau)$ of $z(t)$ can be defined as follows:

$$R_{zz}(t, \tau) = E\{z(t)z^*(t + \tau)\} \tag{2}$$

where τ stands for the time offset. $z^*(\cdot)$ represents the conjugate of $z(\cdot)$ and $E\{\cdot\}$ aims to calculate the mean.

The autocorrelation function $R_{zz}(t, \tau)$ of $z(t)$ can also be expanded as follows by using Fourier series:

$$R_{zz}(t, \tau) = \sum_{p=-\infty}^{\infty} R_{zz}^{\alpha}(\tau) e^{j2\pi\alpha t} \tag{3}$$

where p is the number of accumulations. Thus we have

$$R_{zz}^{\alpha}(\tau) = \lim_{T \rightarrow \infty} \int_{-T/2}^{T/2} z(t)z^*(t + \tau) e^{-j2\pi\alpha t} dt \tag{4}$$

where $R_{zz}^{\alpha}(\tau)$ is the cyclic autocorrelation function (CAF) of $z(t)$. T is period and α is the second-order cycle frequency which equals p/T_0 .

The Fourier transform of the CAF can be defined as SCF as follows:

$$\begin{aligned}
 S_{zz}^\alpha(f) &= \int_{-\infty}^{\infty} R_{zz}^\alpha(\tau) e^{-j2\pi f\tau} d\tau \\
 &= \int_{-\infty}^{\infty} \left\{ \lim_{T \rightarrow \infty} \int_{-T/2}^{T/2} z(t) z^*(t + \tau) e^{-j2\pi\alpha t} dt \right\} e^{-j2\pi f t} dt \\
 &= \lim_{T \rightarrow \infty} \lim_{\Delta t \rightarrow \infty} \int_{-\Delta t/2}^{\Delta t/2} \frac{1}{T} Z(t, f + \alpha/2) Z^*(t, f - \alpha/2) dt
 \end{aligned} \tag{5}$$

where $Z(\cdot)$ is the Fourier transform of signal $z(t)$. Frequency $(f + \alpha/2)$ and frequency $(f - \alpha/2)$ belong to two spectral components of $z(t)$. Δt represents the time interval.

The sketch for calculating the SCF of signal $z(t)$ can be summarized in Fig. 1. SCF provides more possibilities for detecting cyclostationary signals and it works well when distinguishing cyclostationary signals from the noise since cyclostationary signals have peaks in certain positions of SCF diagrams.

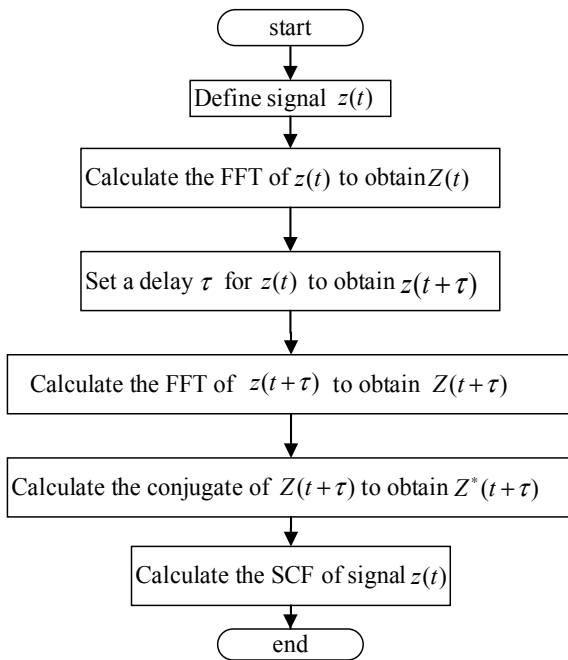


Fig. 1. The sketch for calculating the SCF of signal $z(t)$

3 The Estimation of SCF

We consider converting the estimation of SCF to the estimation of the autocorrelation function as the autocorrelation function’s estimation theory is mature. It’s an estimation process to obtain autocorrelation function $R_{zz}(t, \tau)$ of $z(t)$. $R[m, n]$ represents the discretization form of $R_{zz}(t, \tau)$. t and τ represent starting time and the time offset respectively. $t = mT_1$ and $\tau = nT_2$, where m and n are discrete points and T_1 and T_2 are sampling period. The cyclic autocorrelation function $R^{\Omega_1}[n]$ comes from the discrete Fourier transform (DFT) of $R[m, n]$ about m as follows:

$$R^{\Omega_1}[n] = \sum_{m=-(M-1)}^{M-1} R[m, n]e^{-j\Omega_1 m} \tag{6}$$

where $\Omega_1 = \omega_1 T$, Ω_1 is the digital angular frequency and ω_1 represents analog angular frequency. $\omega_1 = 2\pi f_1$, f_1 is angle frequency and M is the number of observations.

The SCF $S^{\Omega_1}[\Omega_2]$ comes from the DFT of $R^{\Omega_1}[n]$ about n as follows:

$$S^{\Omega_1}[\Omega_2] = \sum_{n=-(N-1)}^{N-1} R^{\Omega_1}[n]e^{-j\Omega_2 n} \tag{7}$$

where $\Omega_2 = \omega_2 T$, Ω_2 and ω_2 represent the digital and analog angular frequency separately. $\omega_2 = 2\pi f_2$ and N is the number of observations. Thus

$$\begin{aligned} S^{\Omega_1}[\Omega_2] &= \sum_{n=-(N-1)}^{N-1} \sum_{m=-(M-1)}^{M-1} R[m, n]e^{-j\Omega_1 m} e^{-j\Omega_2 n} \\ &= \sum_{n=-(N-1)}^{N-1} \sum_{m=-(M-1)}^{M-1} R[m, n]e^{-j(\Omega_1 m + \Omega_2 n)} \end{aligned} \tag{8}$$

Calculate the mean of $R[m, n]$ about m as follows:

$$E_m\{R[m, n]\} = \frac{P - m}{P} R[m, n] \tag{9}$$

where P is the number of observations. Calculate the mean of (9) about n as follows:

$$E_{m,n}\{R[m, n]\} = E\{R[m, n]\} = \frac{P - m}{P} \cdot \frac{P - n}{P} R[m, n] \tag{10}$$

Calculate the mean of (8) about $S^{\Omega_1}[\Omega_2]$ and put (10) into it as follows:

$$\begin{aligned}
 E\{S^{\Omega_1}[\Omega_2]\} &= \sum_{n=-(N-1)}^{N-1} \sum_{m=-(M-1)}^{M-1} E\{R[m, n]\} e^{-j(\Omega_1 m + \Omega_2 n)} \\
 &= \sum_{n=-(N-1)}^{N-1} \sum_{m=-(M-1)}^{M-1} \frac{(P-m)(P-n)}{P^2} R[m, n] e^{-j(\Omega_1 m + \Omega_2 n)}
 \end{aligned} \tag{11}$$

Thus, the deviation of $S^{\Omega_1}[\Omega_2]$ can be showed as follows:

$$\text{bia}\{S^{\Omega_1}[\Omega_2]\} = \sum_{n=-(N-1)}^{N-1} \sum_{m=-(M-1)}^{M-1} \frac{mn - P(m+n)}{P^2} \hat{R}[m, n] e^{-j(\Omega_1 m + \Omega_2 n)} \tag{12}$$

where $\hat{R}[m, n]$ is the estimation of discrete autocorrelation function $R[m, n]$.

It is obvious that $\text{bia}\{S^{\Omega_1}[\Omega_2]\}$ is the asymptotic unbiased estimation of $S^{\Omega_1}[\Omega_2]$ when P approaches to ∞ for fixed m and n . For the fixed P , the mean of $\hat{R}[m, n]$ is close to $R[m, n]$ only when m and n is far less than P . It is very difficult to calculate the variance of $S^{\Omega_1}[\Omega_2]$. When signal $z(t)$ is a Gaussian random variable, whose mean is μ and variance is σ^2 as follows: $z \sim N(\mu, \sigma^2)$. Then the variance of $S^{\Omega_1}[\Omega_2]$ can be defined as

$$\text{var}\{S^{\Omega_1}[\Omega_2]\} = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4 \tag{13}$$

The mean square error (MSE) can be defined as

$$\text{MSE}\{S^{\Omega_1}[\Omega_2]\} = \text{var}\{\hat{S}^{\Omega_1}[\Omega_2]\} - \{\text{bia}\{\hat{S}^{\Omega_1}[\Omega_2]\}\}^2 \tag{14}$$

Substitute (12) and (13) into (14), the MSE of $S^{\Omega_1}[\Omega_2]$ can be showed as follows:

$$\begin{aligned}
 \text{MSE}\{S^{\Omega_1}[\Omega_2]\} &= \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4 \\
 &\quad - \left\{ \sum_{n=-(N-1)}^{N-1} \sum_{m=-(M-1)}^{M-1} A \times \hat{R}[m, n] e^{-j(\Omega_1 m + \Omega_2 n)} \right\}^2
 \end{aligned} \tag{15}$$

where $A = [m - P(m+n)]/P^2$.

The smaller the MSE of $S^{\Omega_1}[\Omega_2]$ is, the more precise estimation of $S^{\Omega_1}[\Omega_2]$ is. The estimation sketch of $S^{\Omega_1}[\Omega_2]$ can be summarized in Fig. 2.

4 Simulation and Analysis

In this section, simulation results are presented and they testify the ability of acquiring the spectrum information of cyclostationary detection algorithm in the complex electromagnetic environment.

Define $s_{lfm}(t) = s_1(t) + s_2(t) + s_3(t)$ where $s_1(t)$, $s_2(t)$ and $s_3(t)$ are three different signals. Define their frequency rate $K_1 = 3 \times 10^{12}$, $K_2 = 1.5 \times 10^{12}$, $K_3 = 1.5 \times 10^{12}$,

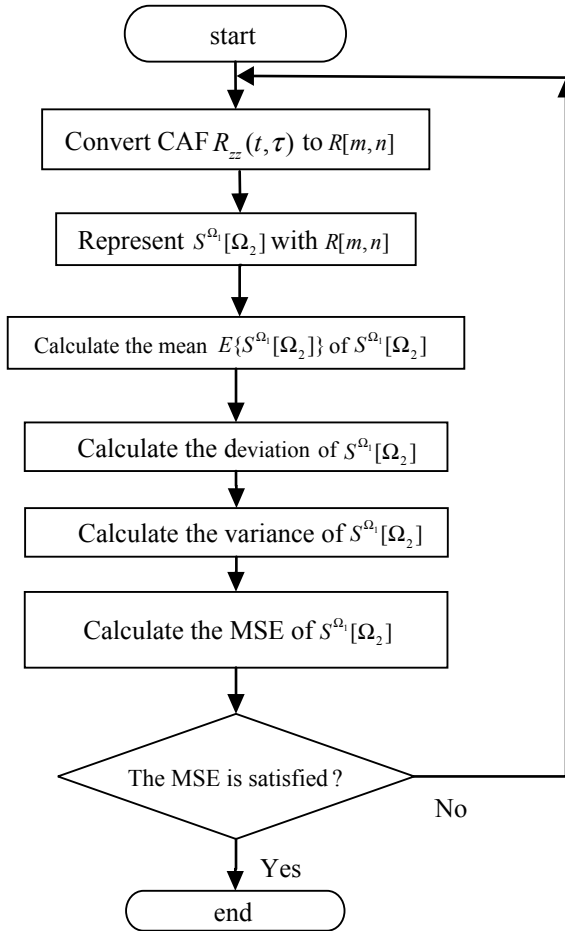


Fig. 2. The sketch of estimation of $S^{\Omega_1}[\Omega_2]$

their initial frequency $f_1 = 0$ MHz, $f_2 = -7.5$ MHz, $f_3 = 10$ MHz, and their bandwidth $B_1 = 30$ MHz, $B_2 = 15$ MHz, $B_3 = 15$ MHz. Figure 3 depicts the SCF of signal $s_{lfm}(t)$. It is a symmetric process to calculate the SCF of signal $s_{lfm}(t)$. Thus, the simulation is symmetric about $\alpha = 0$. To observe clearly, the contour map of Fig. 3 is shown in Fig. 4. It depicts that the spectrum of signal $s_{lfm}(t)$ focuses on the $\alpha = 0$. Because of the influence of initial frequency of $s_1(t)$, $s_2(t)$ and $s_3(t)$, the distribution of signal $s(t)$ from left to right is $s_2(t)$, $s_1(t)$, $s_3(t)$ and there are frequency overlaps between the signals. The frequency overlap between $s_2(t)$ and $s_1(t)$ is shown in area A of Fig. 4. The length of area A represents the length of frequency overlap.

The frequency overlap between $s_1(t)$ and $s_3(t)$ is shown in area B of Fig. 4. The length of area B represents the length of frequency overlap. The gradient of line 1a–4a can be used to judge the magnitude of the frequency rate between signals. Line 1a

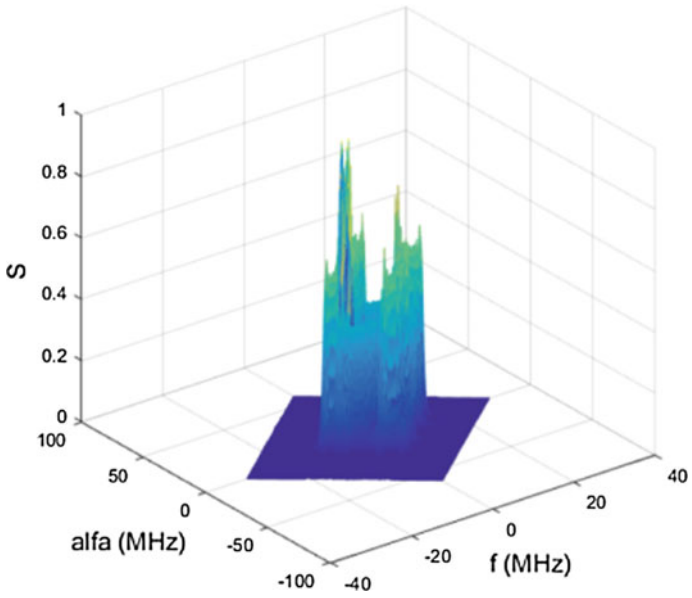


Fig. 3. The SCF of signal $s_{fm}(t)$

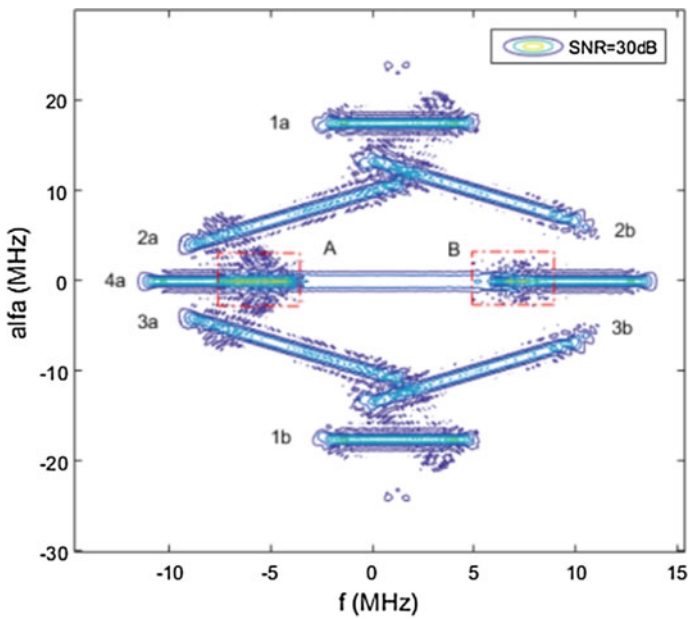


Fig. 4. The contour map of SCF of signal $s_{fm}(t)$

represents the frequency rate between $s_2(t)$ and $s_3(t)$, and the reason why line 1a is parallel to line 4a is that $K_2 = K_3$. Line 2a represents the frequency rate between $s_2(t)$ and $s_1(t)$. The effect is shown in Fig. 4 as the result of $K_2 < K_1$. Similarly, line 3a represents the frequency rate between $s_1(t)$ and $s_3(t)$. The effect is shown in Fig. 4 as the result of $K_1 > K_3$. It shows that the analysis of SCF can be used to sense the information of signal such as bandwidth, frequency, frequency rate, and frequency overlap.

To analyze the influence of the SNR on the cyclostationary detection algorithm, we select the $SNR = 5$ dB and $SNR = 20$ dB, other simulation conditions are the same as Fig. 3. Simulation results in Fig. 5 indicate that the cyclostationary detection algorithm is highly preferred for spectrum sensing. We can obtain the spectrum information such as the frequency overlap and the frequency rate even through the SNR is low.

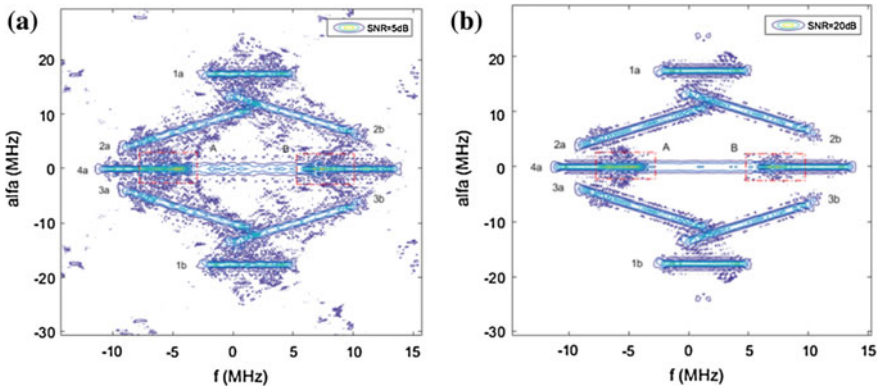


Fig. 5. The influence of the SNR on the cyclostationary detection algorithm

There are various interference signals in electromagnetic environment of radar. Simulation is performed to illustrate the ability of cyclostationary detection algorithm, which can be used to distinguish the frequency range among different kinds of interference signals. Figure 6 shows the SCF of different signals. An AM signal is named $s_1(t)$ whose carrier frequency f_c is 10 MHz and the baseband frequency f_b is 5 Hz. A LFM signal is named $s_2(t)$ whose bandwidth B is 11.25 MHz, the frequency rate K is 2.25×10^{12} and the initial frequency f_0 is 15 MHz. Define $s_{am-lfm}(t) = s_1(t) + s_2(t)$.

The SCF of signal $s_{am-lfm}(t)$ are shown in Fig. 6 when $SNR = 5$ dB and $SNR = 20$ dB. We use the contour maps which indicate that there are the spectral characteristic at $alfa = \pm 2f_c$ (point A and point C) and $f = \pm f_c$ (point B and point D) of $s_1(t)$. The spectral characteristic of signal $s_2(t)$ concentrates on the location of line 3a, the length of line 3a equals the bandwidth of signal $s_2(t)$. It is a symmetric process to calculate the SCF of signal. Thus, the simulation is symmetric about $alfa = 0$. The lines 1a–2b exist because of the combination of AM signal $s_1(t)$ and LFM signal $s_2(t)$. The location and the length of lines 1a–2b change with the frequency rate and initial frequency of $s_2(t)$. There is a frequency overlap at point D about $s_1(t)$ and $s_2(t)$.

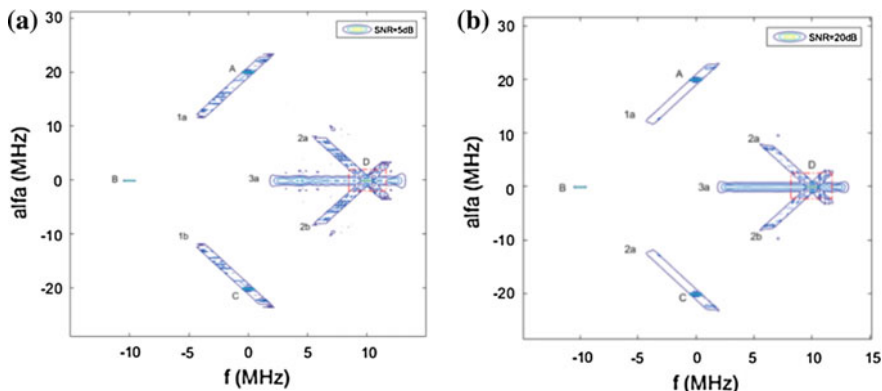


Fig. 6. The contour map of SCF of signal $s_{am-lfm}(t)$ in different SNR

Simulation results indicate that the cyclostationary detection algorithm can distinguish different signals even $SNR = 5$ dB. With the SNR increasing, the discernment becomes better and better. We can analyze the SCF of complex electromagnetic signals to obtain the spectrum information such as bandwidth of signal, initial frequency, frequency rate, frequency overlap, and so on.

Figure 7 shows the comparison of the Cyclic Spectral Algorithm and the method of short-time Fourier transform (STFT). The experiment is finished by the LFM signal and the AM signal. Define the $SNR = 20$ dB. As illustrated in Fig. 7, with the decrease of bandwidth, the STFT Algorithm gets weaker and weaker in distinguishing signals. However, the Cyclic Spectral Algorithm can detect signals clearly.

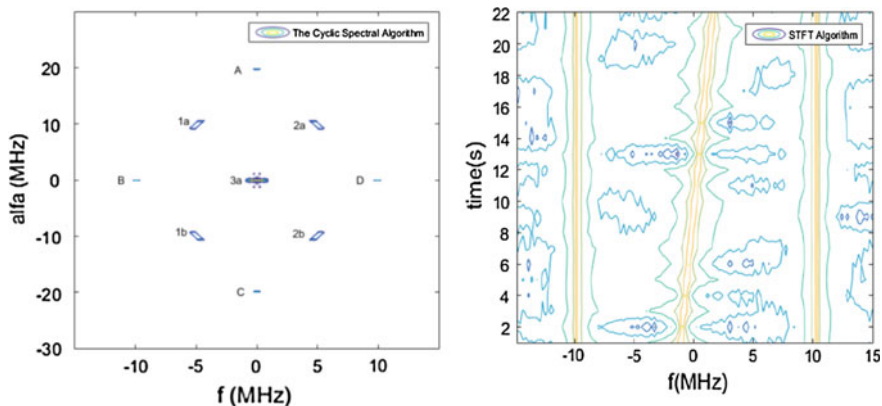


Fig. 7. Comparison of radar electromagnetic environment sensing method

5 Conclusion

This paper presents a radar electromagnetic environment sensing method based on the cyclic spectral algorithm to sense the spectral information by calculating the SCF. A method to calculate the estimation precision of SCF is also introduced. Simulation results indicate that the type of signal can be obtained and the bandwidth, frequency rate, initial frequency, and frequency overlap can also be known from the curve of SCF in low SNR. Thus, this method can detect the interference frequency effectively in the complex electromagnetic environment. However, the scenario of radar electromagnetic environment is simple. In the future, complex scenario will be further considered.

References

1. Yu, S., Wang, X.: Joint spectrum sensing in distributed MIMO systems. In: Vehicular Technology Conference, pp. 1–4. IEEE (2011)
2. Huang, G., Tugnait, J.K.: On cyclostationarity based spectrum sensing under uncertain gaussian noise. *IEEE Trans. Signal Process.* **61**(8), 2042–2054 (2013)
3. Damavandi, M.A., Nader-Esfahani, S.: Compressive wideband spectrum sensing in cognitive radio systems based on cyclostationary feature detection. In: International Conference on Next Generation Mobile Applications, Services and Technologies, pp. 282–287. IEEE (2016)
4. Zhang, T., Yu, G., Sun, C.: Performance of cyclostationary features based spectrum sensing method in a multiple antenna cognitive radio system. In: Wireless Communications and Networking Conference, WCNC, pp. 1–5. IEEE (2009)
5. Yawada, P.S., Wei, A.J.: Cyclostationary detection based on non-cooperative spectrum sensing in cognitive radio network. In: IEEE, International Conference on Cyber Technology in Automation, Control, and Intelligent Systems, pp. 184–187. IEEE (2016)
6. Kandeepan, S., Baldini, G., Piesiewicz, R.: Experimentally detecting IEEE 802.11n Wi-Fi based on cyclostationarity features for ultra-wide band cognitive radios. In: IEEE, International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 2315–2319. IEEE (2009)



Performance Analysis of Mid-Far Infrared Wave in Satellite-Ground Link

Meng Jing^(✉), Li Shuai, Lin Qingqing, and Liu Shuai

Qian Xuesen Laboratory of Space Technology, Beijing 100094, China
mengjing@qxslab.cn

Abstract. Based on the basic theory of laser communication in the atmosphere, the effects of wavelengths, transmission distance, and visibility on communication link are studied; considering the restriction parameters in the satellite-ground communication link, the outage probability, fade statistic, intensity fluctuation and bit error rate of OOK/BPSK modulation are derived; the link budget is discussed; the results show that mid-far infrared wave has better transmission performance in atrocious atmosphere compared with the near-infrared wave, which could become an optimization choice in satellite-ground downlinks.

Keywords: Communication in satellite-ground link · Mid-far infrared wave · Link performance · BER

1 Introduction

Due to the advantages of large capacity, high speed, high concealment, and high anti-interference ability, optical communication has become the development trend of large capacity space communication in the future. A series of pilot projects have been carried out in space communication using the near-infrared (0.8–1.55 μm) laser [1–4], the transmission rate is also a few Mbps to 5.625 Gbps; with low bit error rate coherent laser communication is realized. However, haze and turbulent environment will seriously affect the near-infrared communication performance, under the constraints of the atmospheric model, searching for the available optical frequency band has always been the primary concern of space laser communication [5].

In recent years, the longwave infrared laser is becoming more practical [6]. Andrew et al. established a preliminary long wave infrared laser system, using a quantum cascade laser with a power of 180mW, adjustable wavelength and working at room temperature to achieve the average transmitted rate 155 Mbps (up to 1 Gbps) [7]. Donald et al. established an experimental system with a maximum modulated frequency of 500 MHz with 5 km transmitted distance and the laser output power exceeds 1 W [8].

This paper is organized as follows: the communication characteristic in atmospheric and the influence of different link parameters are studied in Sects. 2 and 3 analyze the outage characteristics of infrared laser satellite-ground link and the BER performance under OOK/BPSK modulation; the link budget of satellite-ground downlink using mid-infrared is carried out in Sect. 4; Finally, the conclusions are drawn in Sect. 5.

2 Basic Theory

2.1 Transmission Loss

The atmospheric transmittance of slant path can be defined as [9]

$$T = \exp(-\sec \varphi \int_0^H \alpha(\lambda, h) dh) \quad (1)$$

here λ is the wavelength, φ is the zenith angle, H is the transmitted distance, $\alpha = \alpha_a + \alpha_s$ is the total attenuation coefficient, α_a is the absorption coefficient, α_s is the scattering coefficient. For the atmospheric absorption, there are eight higher transmittance bands between the 0.72–15 μm , known as the atmospheric window, with the increase of altitude, atmospheric absorption effect decrease obviously. For the atmospheric scattering, Rayleigh-scattering and Mie-scattering are classified according to the difference of laser wavelength and particle radius.

The Rayleigh-scattering occurs when the diameter of particles is much smaller than the wavelength, and its attenuation coefficient is

$$\sigma_m = 2.677 \times 10^{-17} \times Pk^4/T \quad (2)$$

here P is the atmospheric pressure, T is the absolute temperature, k is the wave number.

The Mie-scattering occurs when the diameter of particles is the same as the wavelength, and its attenuation coefficient is

$$\sigma_a = 3.91/R_v \cdot (\lambda_0/\lambda)^q \quad (3)$$

here R_v is the visibility (km), λ is the wavelength, $\lambda_0 = 0.55 \mu\text{m}$, q is the wavelength correction factor which related to the visibility, can be represented as

$$q = \begin{cases} 1.6 & R_v > 50 \text{ km} \\ 1.3 & 6\text{km} < R_v \leq 50 \text{ km} \\ 0.16R_v + 0.34 & 1\text{km} < R_v \leq 6 \text{ km} \\ R_v - 0.5 & 0.5\text{km} < R_v \leq 1 \text{ km} \\ 0 & R_v < 0.5 \text{ km} \end{cases} \quad (4)$$

The Rayleigh-scattering and Mie-scattering attenuation coefficients vary with laser wavelength are shown in Fig. 1, the attenuation caused by the molecular can be ignored in mid-far infrared band compared to the near-infrared wave; and in the same visibility conditions, the longer the wavelength, the smaller the attenuation of the Mie-scattering.

2.2 Atmospheric Turbulence

Turbulence is a random change of the refractive index due to the variations in atmosphere temperature and pressure, which can cause the laser beam jitter and the intensity fluctuation of the received signal. In turbulence environment, the scintillation index of the received intensity is [10]

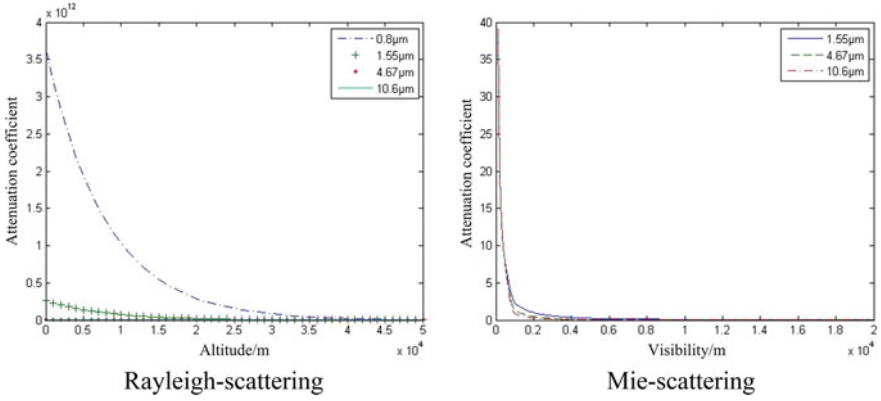


Fig. 1. The scattering attenuation coefficient varies with the wavelength

$$\sigma_I^2 = 2.25\mu_1 k^{7/6} (H - h_0)^{5/6} \sec^{11/6} \varphi \quad (5)$$

$$\mu_1 = \int_{h_0}^H C_n^2(h) \left(\frac{h - h_0}{H - h_0} \right)^{5/6} dh \quad (6)$$

here H is the transmitted distance, h is the altitude, h_0 is the ground altitude, k is the wave number, φ is the slant path zenith angle, C_n^2 is the refractive-index structure constant, according to the Hufnagel-Valley model [11]:

$$C_n^2(h) = A_{hv} \exp(-h/100) + 0.00594(v_v/27)^2 (10^{-5}h)^{10} \times \exp(-h/1000) + 2.7 \times 10^{-16} \exp(-h/1500) \quad (7)$$

here A_{hv} is the refractive-index structure constant near the ground, v_v is the vertical path wind speed, typical values are $A_{hv} = 1.7 \times 10^{-14} \text{ m}^{-2/3}$, $v_v = 21 \text{ m/s}$. The scintillation index in satellite-ground link varies with the wavelength is shown in Fig. 2 with $H = 36000 \text{ km}$, $h_0 = 500 \text{ m}$, $\varphi = 30^\circ$. It is shown that the longer the wavelength, the lower the scintillation index, and also the lower requirement for the aperture of the receiving antenna.

3 Performance Analysis in Satellite-Ground Laser Link

3.1 Outage Characteristics

The fluctuation of the received intensity follows the lognormal distribution in weak turbulence, and follows the gamma-gamma distribution in medium and strong turbulence, which probability density function can be defined as [12]

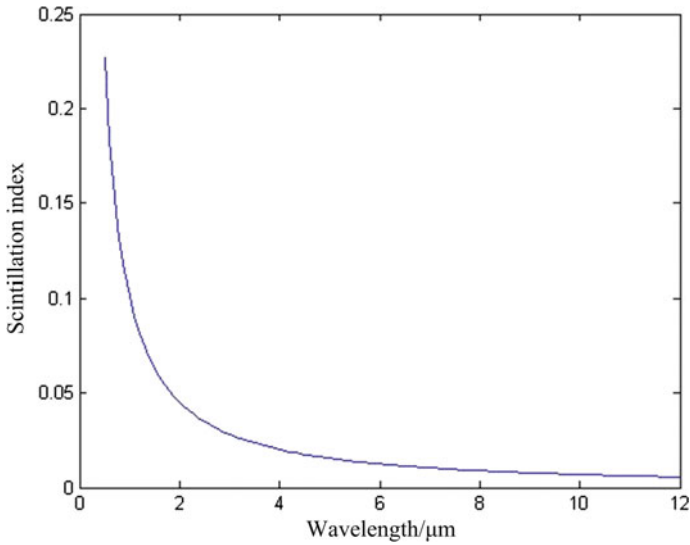


Fig. 2. The scintillation index of received intensity varies with the wavelength

$$p_I(u) = \frac{2}{\Gamma(\alpha)\Gamma(\beta)u} (\alpha\beta u)^{(\alpha+\beta)/2} K_{\alpha-\beta}(2\sqrt{\alpha\beta}u) \tag{8}$$

here u is the normalized intensity, $K(\cdot)$ is the second type modified Bessel function, the parameter α and β are

$$\alpha = \frac{1}{\exp[\sigma_{\ln X}^2(D)] - 1}, \quad \beta = \frac{1}{\exp[\sigma_{\ln Y}^2(D)] - 1} \tag{9}$$

here $\sigma_I^2(0, L; D) = \exp[\sigma_{\ln X}^2(D) + \sigma_{\ln Y}^2(D)] - 1$, $\sigma_{\ln X}^2(D)$ and $\sigma_{\ln Y}^2(D)$ are the scintillation index with the fixed receiver aperture.

The outage probability is defined as the probability that the system bit error rate is greater than the target bit error rate, the outage probability can be expressed as

$$\Pr_{fade} = \int_0^{I_T} p_I(u) du \tag{10}$$

here I_T is the intensity threshold. The expected fading can be defined as

$$\langle n(I_T) \rangle = \frac{2\sqrt{2\pi\alpha\beta\sigma_I^2(0, L; D)}v_0}{\Gamma(\alpha)\Gamma(\beta)} (\alpha\beta I_T)^{(\alpha+\beta-1)/2} K_{\alpha-\beta}(2\sqrt{\alpha\beta}I_T) \tag{11}$$

The average fading time can be defined as

$$\langle t(I_T) \rangle = \text{Pr}_{fade} / \langle n(I_T) \rangle \tag{12}$$

3.2 BER of OOK/BPSK Modulation

Under the effects of atmospheric turbulence, the average bit error rate of OOK and BPSK can be expressed as

$$\text{Pr}(E) = \langle BER \rangle_{OOK} = \frac{1}{2} \int_0^\infty p_I(u) \text{erfc} \left(\frac{\langle SNR \rangle u}{2\sqrt{2}} \right) ds \tag{13}$$

$$\text{Pr}(E) = \langle BER \rangle_{BPSK} = \frac{1}{2} \int_0^\infty p_I(u) \text{erfc} \left(\frac{\langle SNR \rangle u}{\sqrt{2}} \right) ds \tag{14}$$

here $\text{erfc}(\cdot)$ is the error function complement, $\langle SNR \rangle$ is the average signal-to-noise ratio, which can be expressed as

$$\langle SNR \rangle = \frac{SNR_0}{\sqrt{\frac{1}{1 + 1.33\sigma_R^2 \Lambda^{5/6}} + \sigma_I^2(0, L; D) SNR_0^2}} \tag{15}$$

here SNR_0 is the signal-to-noise ratio without turbulence.

3.3 Numerical Results

3.3.1 The Fading Probability and Average Fading Time

The fading probability and average fading time of the satellite-ground link vary with the received intensity are shown in Fig. 3 with receiver aperture is 10 cm, the longer the wavelength, the lower the fading probability and average fading time of the link.

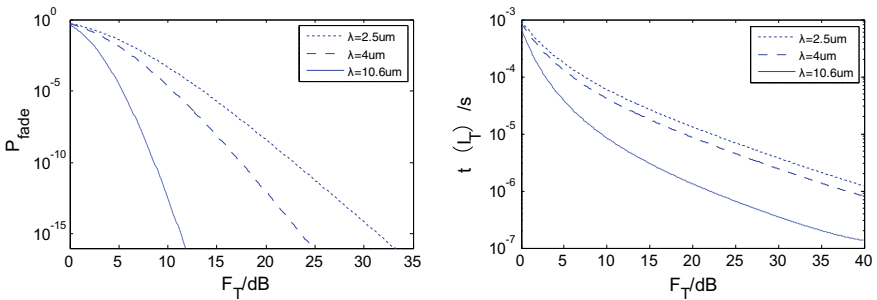


Fig. 3. The variation of outage probability and average fading time with the intensity

3.3.2 Relation of BER and SNR with OOK/BPSK Modulation

The satellite-ground laser link performance with the relation of BER and SNR under OOK/BPSK modulation is shown in Fig. 4. The BPSK modulation method is obviously superior to the OOK, especially in longwave infrared which can achieve better than 3 dB gain compared to the OOK modulation.

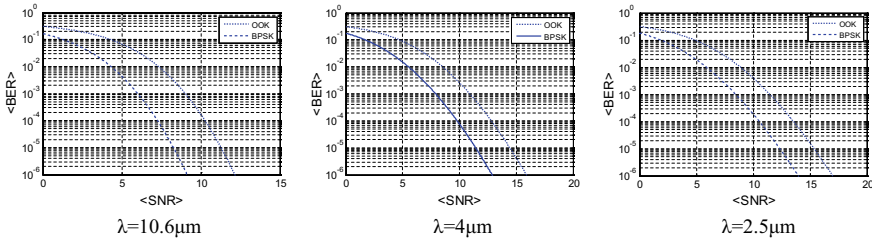


Fig. 4. Relation of BER and SNR with OOK/BPSK modulation

4 Link Budget in Satellite-Ground Down Link

The average received power of satellite-ground downlink can be calculated as

$$M = P_t + G_T + \tau_t + L_{FSL} + G_R + \tau_r + \eta_{dd} - S_{req} - \tau_{am} - L_{FM} \tag{16}$$

here M is the link margin, P_t is the output power, G_T and G_R is the gain of transmitted and received antenna, respectively, τ_t and τ_r is the transmittance of transmitted and received antenna, respectively, L_{FSL} is the free space loss; η_{dd} is the efficiency of the detector, S_{req} is the minimum received power request, τ_{am} is the atmospheric loss; L_{FM} is the scintillation loss.

Choose 4 μm mid-infrared wave for satellite-ground downlink budget analysis, the laser output power is 1 W, the transmittance is 0.6, transmitted distance is 40,000 km, transmitted antenna aperture is 250 mm, received antenna aperture is 1 m, the link margin is usually requested at least more than 3 dB, so it is required the received detector has a more than 50 dBm sensitivity in order to establish the satellite-ground downlink.

5 Conclusions

The performance of satellite-ground laser communication with different wavelengths is studied, considering the transmitted characteristics in turbulent environment, the longwave infrared has better performance than the near-infrared, which could breakthroughs the limitations of near-infrared laser communication in atrocious atmosphere. Improving the power of longwave infrared laser is a key problem to be solved in practical application. The future work will focus on the influence of satellite platform dynamics environment, noise environment and APT performance on the mid-far infrared laser communication in satellite-ground downlinks.

References

1. Seel, S., Kämpfner, H., Heine, F., et al.: Space to ground bidirectional optical communication link at 5.6 Gbps and EDRS connectivity outlook. In: IEEE Aerospace Conference, pp. 1–7 (2011)
2. Edwards, B.L., Israel, D., Wilson, K., et al.: Overview of the laser communications relay demonstration project. ASTIA Documents **1303**, 1–13 (2012)
3. Bohmer, K., Gregory, M., Heine, F., et al.: Laser communication terminals for the European Data Relay System. In: Proceedings of SPIE, vol. 8246, p. 82460D (2012)
4. <http://www.nsoas.gov.cn/HY2AZhuanti/HY2A/hy202.html>
5. http://www.nasa.gov/offices/oct/strategic_integration/technology_roadmap.html
6. Gutowska, M., Pierscińska, D., Nowakowski, M., et al.: Transmitter with quantum cascade laser for free space optics communication system. Bull. Pol. Acad. Sci. **59**(4), 419–423 (2011)
7. Pavelchek, A., Trissel, R.G., Plante, J., et al.: Long-wave infrared (10 μm) free-space optimal communication system. In: Proceedings of SPIE on Optical Science and Technology, pp. 247–252 (2004)
8. Hutchinson, D.P., Richards, R.K., et al.: All-weather long-wavelength infrared free space optical communications. In: International Symposium on Optical Science and Technology, pp. 44–49 (2002)
9. Hamid, H.: Deep Space Optical Communications. Jet Propulsion Laboratory California Institute of Technology Press, California (2005)
10. Beland, R.R.: Propagation Through Atmospheric Optical Turbulence. The Infrared and Electro Optical Systems Handbook. SPIE Optical Engineering Press, Bellingham (1993)
11. Wu, Z.S., Wei, H.Y., Yang, R.K., et al.: Study on scintillation considering inner-and outer-scales for laser beam propagation on the slant path through the atmospheric turbulence. Prog. Electromagn. Res. **80**, 277–293 (2008)
12. Al-Habash, M.A., Andrews, L.C.: Mathematical model for the irradiance probability density function of a laser beam propagating through turbulent media. Opt. Eng. **40**(8), 1554–1562 (2001)



Seawater Antenna for High-Frequency Surface Wave Radar

Linwei Wang, Changjun Yu (✉), and Haorong Wang

The School of Information and Electrical Engineering, Harbin Institute of Technology at Weihai, Weihai, China
yuchangjun@hit.edu.cn

Abstract. Liquid antennas can be easily reconfigured, which means that the antennas can work at different frequency by adjusting the height of liquid surface. This paper presents that seawater antennas can be used in high-frequency surface wave radar. First, performance of seawater antennas is analyzed using Ansys HFSS. Then the measured results of a real seawater antenna are presented. The results show good agreement with the simulation and can be used in a range of 4–15 MHz.

Keywords: Seawater antenna · Liquid antenna · High frequency surface wave radar

1 Introduction

In recent years, liquid antennas get a lot of attention. Initially, liquid antennas appeared among amateur radio operators. In 2002, N9ZRT made a cylindrical antenna using salt water for high frequency communication [1]. In 2005, Ewananovil [1] demonstrated a Marine monopole antenna with a frequency of 1.3 GHz. In 2012, Tam [2] made various injection antenna systems using water pumps. In 2012, Lei [3] analyzed the influence of solution conductivity and dielectric constant of the base on antenna performance. In 2014, Hua [4] improved the radiation efficiency of seawater antenna by loading metal discs on the top of the feeding probe. Hua [5] designed a seawater half-loop antenna for VHF maritime wireless communications in 2015.

Liquid antennas have great reconfigurable ability, which is convenient to vary characters by changing the size and shape of the antenna. When the radar system stops working the liquid inside antennas can be pumped out, which improves the stealthiness level [6]. Furthermore, water is nearly transparent, which makes it difficult to be observed. Seawater is one of the most popular materials because it is easy to obtain with relatively low cost.

High-frequency surface wave radar (HFSWR) operates in 3–15 MHz, in this frequency band, the conductivity of seawater is 4–6 S/m, and the relative dielectric constant is 81, consequently, seawater can be regarded as a perfect conductor. Besides, HFSWR antennas need to cover a relatively wide band, which results in significant difficulty in design. The reconfigurability of seawater antennas has a future prospect in HFSWR applications.

In this paper, a design of seawater antenna for HFSWR receiving is investigated. First, we analyze the influence of structure on the performance of the seawater antennas. Then the measured results are given compared with the simulation. The results demonstrate that the seawater antenna performs well in frequency from 4 to 15 MHz.

2 The Structure of the Seawater Antenna

Figure 1 shows the structure of the seawater antenna. The antenna consists of a plexiglass tube, a POM base, and a water pump. The external diameter of the tube is d_1 , and the inner diameter is d_2 . The height of seawater is H , the height of POM base is H_1 . The antenna needs to be mounted on a ground screen.

Figures 2 and 3 are the reflectance coefficient and impedance characteristic of the seawater antenna compared with the perfect conductor (PEC) antenna. When the height is 4 m, radius is 10 cm, the seawater and PEC antennas have similar performance, while the impedance of seawater antenna is higher, and the resonant frequency is lower.

The seawater antenna can change the resonant frequency by adjusting the height of the seawater. Figures 4, 5 and 6 show the HFSS simulation of different seawater height. The resonant frequency of the seawater antenna will rise if the height of water surface reduces.

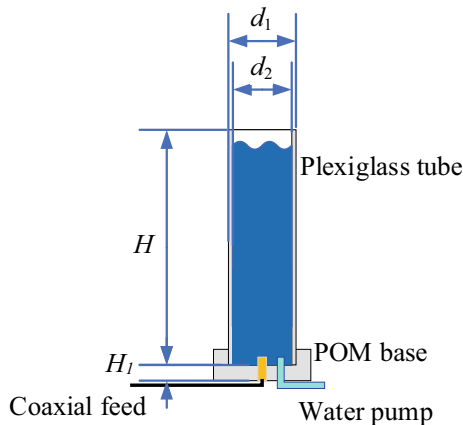


Fig. 1. Structure of the seawater antenna

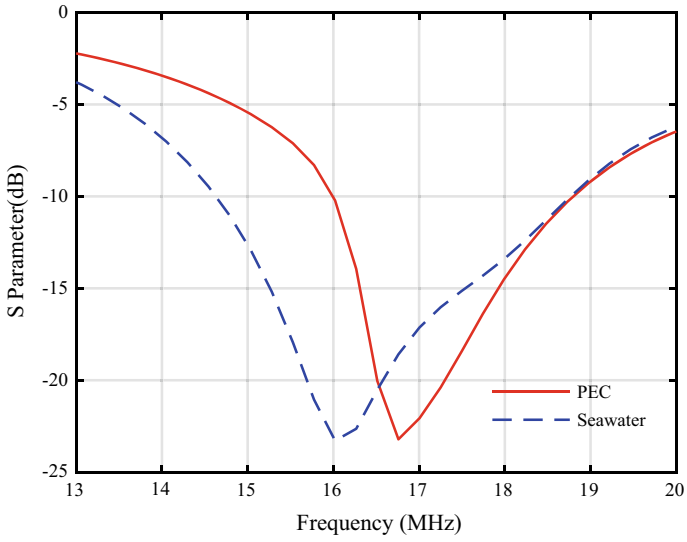


Fig. 2. Reflectance coefficient of antenna

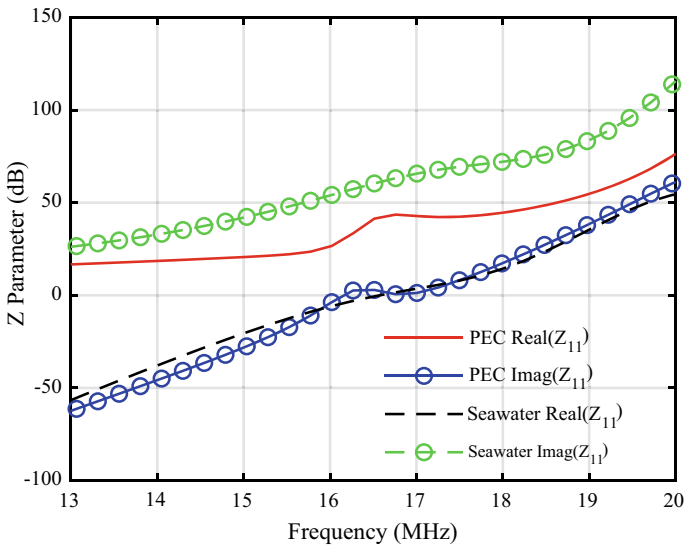


Fig. 3. Impedance characteristic of antennas

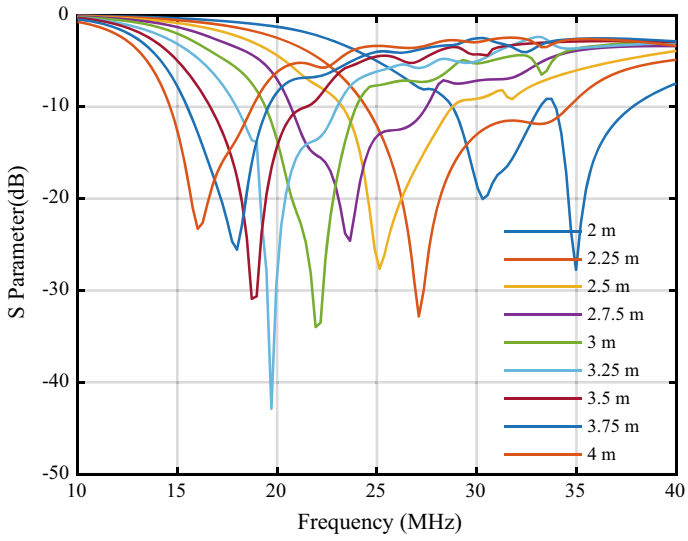


Fig. 4. S parameters with different seawater height of the antenna

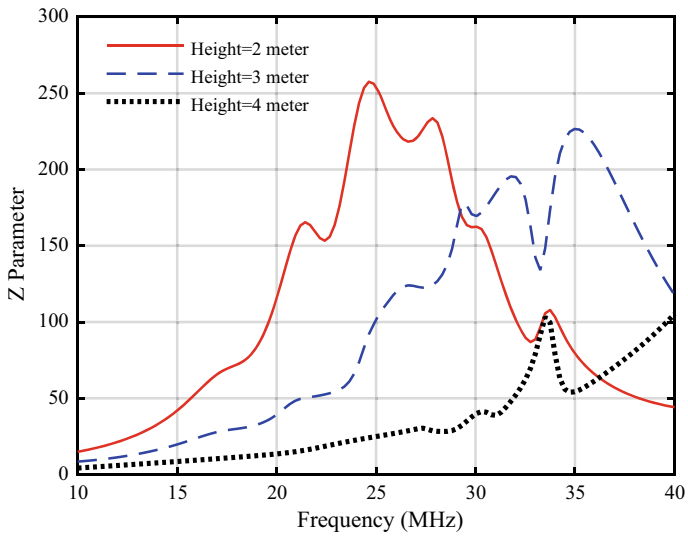


Fig. 5. Real (Z_{11}) with different seawater height of the antenna

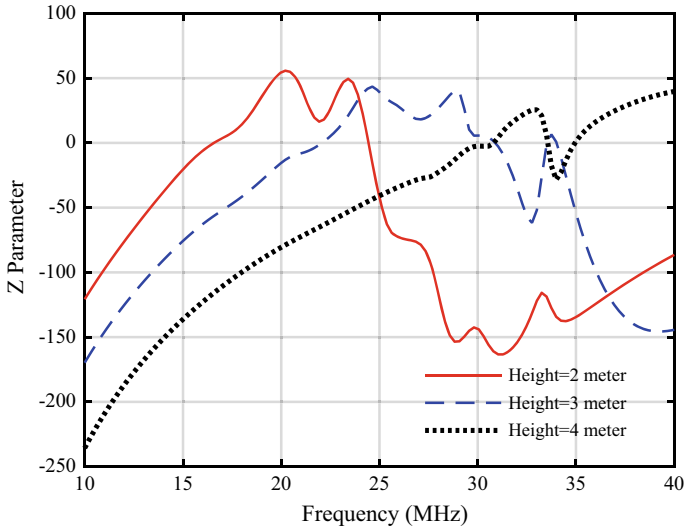


Fig. 6. Imag (Z_{11}) with different seawater height of the antenna

3 Measurement

The seawater antenna fabricated is shown in Fig. 7. The maximum height of the antenna $H = 4$ m. Plexiglass tube external diameter $d_1 = 22$ cm, and inner diameter $d_2 = 20$ cm. The height of POM base $H_b = 20$ cm. The antenna was set up at the beach, with a 10 cm spacing grid ground screen. We use sodium chloride (NaCl) solution whose concentration is about 1%.

Figure 8 shows the measured results of the seawater antenna. In order to make the antenna to operate at the frequency 4–15 MHz, the antenna matching is adjusted roughly. In Fig. 8, the curves with an asterisk is simulated results, while the curves without asterisk are measured results. The measured and simulated results have the same trend. When the height of seawater varies from 4 to 2 m, the resonant frequency changes from 8 to 12 MHz, and the available frequency ($S_{11} < -10$ dB) is 4–16 MHz.

The seawater antenna gain is about -20 dBi, as shown in Fig. 9, which is higher than the antenna in a practical HFSWR system, whose detection range reaches 200 km. Consequently, the seawater antenna satisfies the requirement of HFSWR receiving antennas. The value of gain is higher at 11 MHz may due to the cross-coupling from nearby antennas whose height is about 6 m.



Fig. 7. Photograph of the fabricated seawater antenna

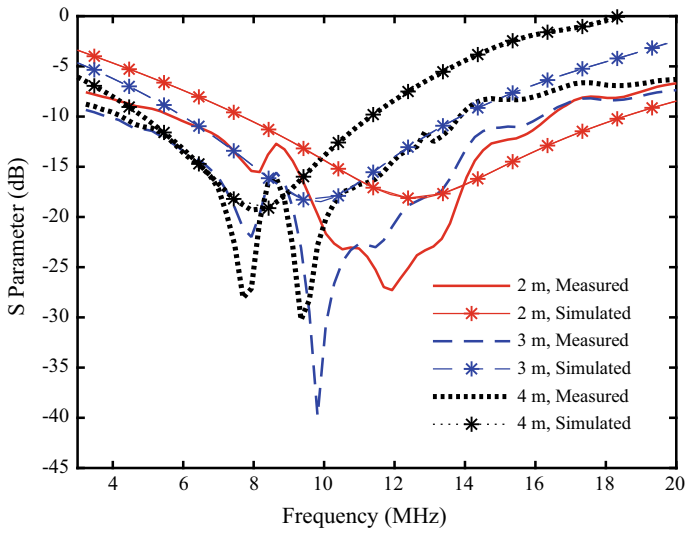


Fig. 8. Measured and simulated reflection coefficients of the seawater antenna

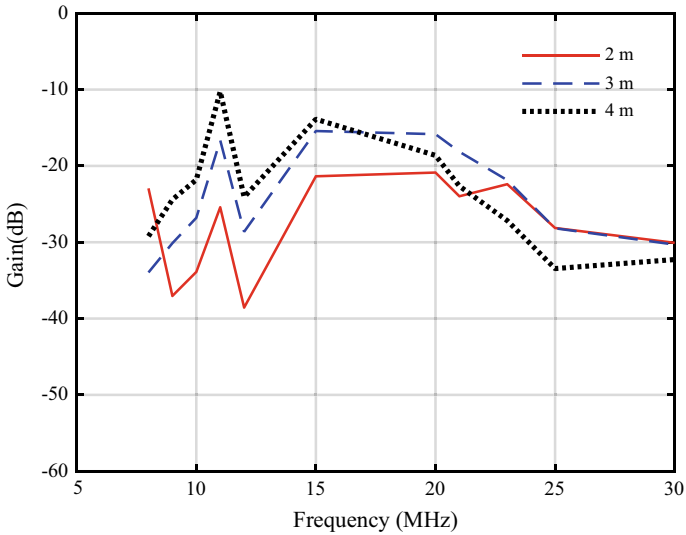


Fig. 9. Gain of the seawater antenna

4 Conclusion

This paper has presented the idea of application of seawater antennas on HFSWR. The measured results of the seawater antenna agreed with the simulation. The center frequency of the antenna can be altered by controlling the height of the seawater. The resonant frequency of the antenna can be turned from 8 to 12 MHz, and the bandwidth is 15.6% (4–16 MHz, $S_{11} < -10$ dB). The results show the availability of using seawater antennas on HFSWR. The reconfigurability of seawater antenna meets the requirement of wide bandwidth of HFSWR, and reduces the possibility of being detected. Other characteristics such as low cost and high transparency are also useful for HFSWR.

Acknowledgements. This work was supported in part by the National Key R&D Program of China under Grant 2017YFC1405202, in part by the National Natural Science Foundation of China under Grant 61571159 and Grant 61571157, and in part by the Public Science and Technology Research Funds Projects of Ocean under Grant 201505002.

References

1. Paraschakis, E., Fayad, H., Record, P.: Ionic liquid antenna. International workshop on antenna technology (2005)
2. Tam, D.W.S.: Electrolytic fluid antenna. US, US8169372 (2012)
3. Xing, L., Huang, Y., Alja'afreh, S.S., Boyes, S.J.: A monopole water antenna. *IEEE Trans. Antennas Propag.* **7363**, 1–4 (2012)

4. Hua, C., Shen, Z., Lu, J.: High-efficiency sea-water monopole antenna for maritime wireless communications. *IEEE Trans. Antennas Propag.* **62**(12), 5968–5973 (2014)
5. Hua, C., Shen, Z.: Sea-water half-loop antenna for maritime wireless communications. *IEEE Trans. Antennas Propag.* 231–232 (2015)
6. Fan, R.: Research on water antennas for wireless communications. Doctoral dissertation, South China University of Technology (2016)



An Underwater Sensor Networks Based Cooperative Positioning System for Falling Water Containers

Manyu Xu and Ying Wang^(✉)

College of Information Science and Technology, Dalian Maritime University,
Dalian 116026, China
wangying@dlmu.edu.cn

Abstract. In the process of shipping, container overboard falling accidents occur frequently. To detect the submerged containers, this paper proposes an underwater acoustic sensor network based detection system, which performs the positioning task through a multi-beacon nodes cooperative method to improve the robustness and accuracy of the system. The simulation results show that the proposed multi-beacon nodes cooperative positioning system can effectively solve the problem of link break between the sensing nodes and the beacon nodes due to the severe underwater environment and extend the detecting area with improved positioning accuracy.

Keywords: Drowning container · Underwater sensor networks · Localization

1 Introduction

As part of the global trade, approximately 130 million containers loaded with goods are shipped worldwide each year. However, due to the bad weather and unpredictable sea conditions, etc., it is inevitable that containers fall into the sea and submerge. Recently, in March 2018, the Maersk Shanghai container ship lost about 76 containers in North Carolina. One of the containers was equipped with 3 tons of dangerous chemicals, sulfuric acid, which may cause serious pollution to the ecological environment [1]. It is important to salvage the drowning containers which carry dangerous goods in time. It should be noted that a quick and accurate detecting and positioning system for the drowning container is a prerequisite for timely salvage.

Underwater acoustic sensor networks technology has been utilized to detect and localize the submerged containers [2]. However, in the scheme proposed in [2] there exist some shortcomings, such as un-optimized deployment of the detecting sensor nodes and high probability of the link break between the detecting sensor node and the beacon node mounted on the container. In this paper, an improved container positioning scheme for a three-dimensional environment is considered. In particular, a fixed detecting nodes deployment method is employed to cover a large area with a small number of detecting nodes. Furthermore, to extend the coverage area of the detecting nodes and improve the positioning accuracy and robustness, a cooperation scheme between the beacon nodes are presented.

The paper is organized as follows: Sect. 2 describes the multi-sensor collaborative positioning schemes based on the existing system model in 3D space, and Sect. 3 compares the results and analyzes the proposed schemes. Section 4 gives the conclusions.

2 Multi-beacon Nodes Cooperative Positioning System

2.1 Node Deployment

When building a UASNs system, sensor nodes deployment is one of the most important tasks. Especially when the underwater sensor nodes transmit the sensing information to the surface sink node through multiple-hop relays, a good deployment strategy can ensure the network connectivity of the system, and lays a good foundation for the subsequent positioning and real-time tracking submerged containers.

In order to optimize the network performance such as network connectivity and coverage, this paper chooses the tetrahedral deployment scheme in a deterministic way so that the requirements of high coverage and positioning accuracy can be satisfied with fewer sensor nodes [3]. With the help of the autonomous underwater vehicle and surface unmanned ship technologies, sensor nodes can be placed in the designated locations in a controlled manner. Tetrahedron deployment scheme is superior to the random deployment and cube deployment methods in reducing positioning errors and improving positioning ratio, while maintaining the average number of adjoining anchor nodes and network connectivity. A tetrahedron deployment unit is shown in Fig. 1. Four sensing nodes or anchor nodes are placed at the vertices of a regular tetrahedrons, and the possibly exiting submerged containers may randomly lay somewhere in the monitoring space. The number of the tetrahedron deployment unit for covering the whole detecting area can be determined by the volume of the interested area and the underwater acoustic communication distance.

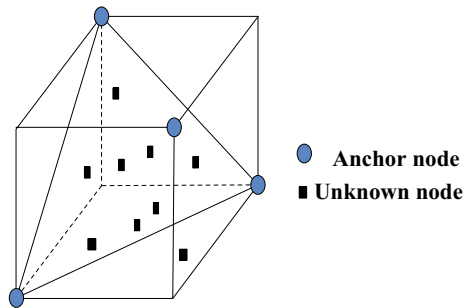


Fig. 1. The tetrahedron deployment unit

2.2 Improved Cooperative Positioning Scheme

Several beacon nodes with the ability of underwater acoustic communication are installed at the different vertex of a container, and each beacon node is assigned a unique ID number. After a container falls into water, the trigger unit will activate the corresponding beacon nodes, and the monitoring center will manage the deployment of sensor nodes in the accident water area. Then, the anchor nodes will communicate with the beacon nodes installed on the container to obtain the distance information between the anchor nodes and the beacon nodes.

Due to the complex underwater environment, in order to reduce the inability to locate the possibly existing container, multi-beacon nodes cooperative positioning scheme is used to improve the existing system.

First, we use the method of multi-beacon nodes, and several beacon nodes are installed on a container in advance. When some beacon nodes cannot communicate with the anchor nodes successfully, other beacon nodes may still have good links with the anchor nodes. As a result, the container can be located. Second, in order to solve the problem that all the beacon nodes on a container cannot communicate with the anchor nodes in a tetrahedron deployment unit, we convert the beacon nodes that have been located into anchor nodes, and such beacon nodes broadcast the their assistant location information containing their accuracy index. Subsequently, when an unlocated beacon node that cannot be located by the normal anchor nodes receives the assistant locating information, it selects the beacon node with the highest accuracy index to calculate its own location. The accuracy index is calculated as follows [4]:

$$\eta_i = 1 - \frac{\sum_{k=1}^3 |(u_i - x_k)^2 + (v_i - y_k)^2 + (w_i - z_k)^2 - l_{ik}^2|}{\sum_{k=1}^3 (u_i - x_k)^2 + (v_i - y_k)^2 + (w_i - z_k)^2} \quad (1)$$

where (u_i, v_i, w_i) is the location coordinate of the i th beacon node, (x_i, y_i, z_i) is the i th anchor node's coordinate. l_{ik} is the measured distance between the i th beacon node and the k th anchor node.

Suppose that two acoustic beacon nodes are installed on a container, then under normal conditions, two sets of measurement data will be transmitted back to the monitoring center for each container. Then, the two sets of data can be used to perform position calculation independently by the tetrahedron positioning algorithm [5], and the positioning results are further fused to reduce the positioning error. On the other hand, if one beacon node fails to communicate with the anchor nodes due to the severe underwater acoustic communication condition and cannot join the positioning network, the other beacon node may still have communication links to the anchor nodes. If both beacon nodes lose links to the anchor nodes, they can try to use the assistant anchor nodes switched from other located beacon nodes. Thus, an improved positioning performance can be achieved.

3 Simulation Results

Simulation for the proposed container positioning system is performed. Figure 2 describes the relationship between the positioning error and the total number of anchor nodes with one beacon node on each container. It can be observed that the tetrahedron deployment is better than the cube deployment and random deployment scheme on the positioning error performance. Figure 3 shows the comparison of the positioning

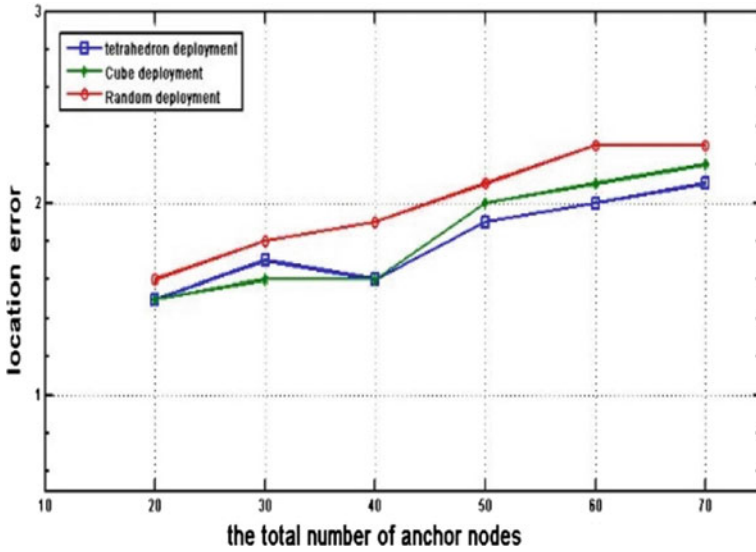


Fig. 2. Location errors of different deployment schemes

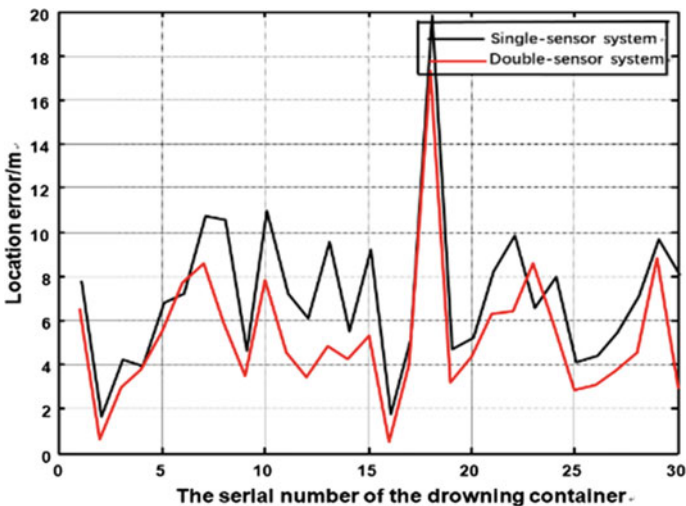


Fig. 3. Location errors with different beacon nodes on each container

performance between two positioning system with one and two beacon nodes on each container, respectively. The simulation results show that the positioning error of the two beacon nodes system is significantly lower than that of the single beacon node system.

4 Conclusion

An improved collaborative localization method is applied to perform the drowning container positioning task. The unstable underwater acoustic communication link may influence the positioning performance. The proposed multi-beacon nodes based cooperative positioning system takes the advantage of multi-beacon nodes installed on each container to adapt to the severe underwater environment. The simulation results show that the proposed scheme can improve the robustness and accuracy of the positioning system with relatively small number of detecting sensor nodes.

Acknowledgments. This work was supported in part by Fundamental research Funds for Central Universities under grant No. 3132016318.

References

1. How to avoid the loss of containers at sea. <http://www.maysun56.com/Article/rhbmjzxzhs.html>
2. Wang, Y., Liu X.: A detection and positioning method and system of the drowning container based on time-frequency addressing. China. Patent 105548999 A (2016)
3. Han, G., Zhang, C., Shu, L., Rodrigues, J.J.P.C.: Impacts of deployment strategies on localization performance in underwater acoustic sensor networks. *IEEE Trans. Ind. Electron.* **62**(3), 1725–1733 (2015)
4. Zhou1, Z., Cui1, J-H., Zhou, S.: Localization for large-scale underwater sensor networks. *Netw.* **4479**, 108–119 (2007)
5. Lee, J.-K., Kim, Y., Lee, J.-H., Ki, S.-C.: An efficient three-dimensional localization scheme using trilateration in wireless sensor networks. *IEEE Commun. Soc.* **18**(9), 1591–1594 (2014)

Internet of Things



Wisdom Farm Internet of Things Software Design and Selection Program

Chaoyan Man, Li Guo^(✉), Yang Gao, and Yu Zhang

School of Information Science and Engineering, Dalian Polytechnic University,
Dalian, China
guoli@dlpu.edu.cn

Abstract. Internet of Things (IOT) technology is another leap forward for the world information industry following the computer and the Internet, providing an unprecedented opportunity for the rapid development of agricultural modernization. This article describes the concept and architecture of agricultural Internet of Things (IoT), analyzes the key technologies of agricultural IoT from the three levels of information perception layer, network transport layer, and application layer. It analyzes the key technologies of IoT and proposes the Agricultural Internet of Things System architecture design. And pointed out that the middleware server software architecture, through the test MQTT agreement, identify the appropriate message mechanism. The development of agricultural internet of things will play a greater role in accelerating the integration of informatization and agricultural modernization and promoting the development of agricultural modernization.

Keywords: Agricultural internet of things · Information perception · Middleware · MQTT protocol

Wisdom Farm uses sensors to collect data and equipment for the greenhouses and installs corresponding controlled devices that change the greenhouse environment in an effort to keep the plants in the best environmental condition. The controlled devices include drip irrigation equipment, roller blinds, ventilators. These are hardware devices in the perceptual layer. They have a variety of data structures, electrical properties of different properties, and multiple types of transport protocols. The overall process of agricultural site data monitoring as shown in Fig. 1.

Taking the application of IoT environment as an example, taking full account of the low power consumption of sensor devices and low network bandwidth, our entire system adopts lightweight publish/subscribe system based on Internet of Things environment. The overall architecture design is shown in Fig. 2.

Figure 2 depicts the architecture design and interaction between the publisher (lower computer), lightweight messaging broker (middleware), and subscriber (upper computer) of a lightweight publish/subscribe system. The lightweight publish/subscribe system under the Internet of Things environment is mainly divided into three parts, among which subscribers (PC) include farm monitoring system, control system, and early warning system; lightweight message broker layer (middleware) is the system's server, responsible for publishing and subscription registration, subscription

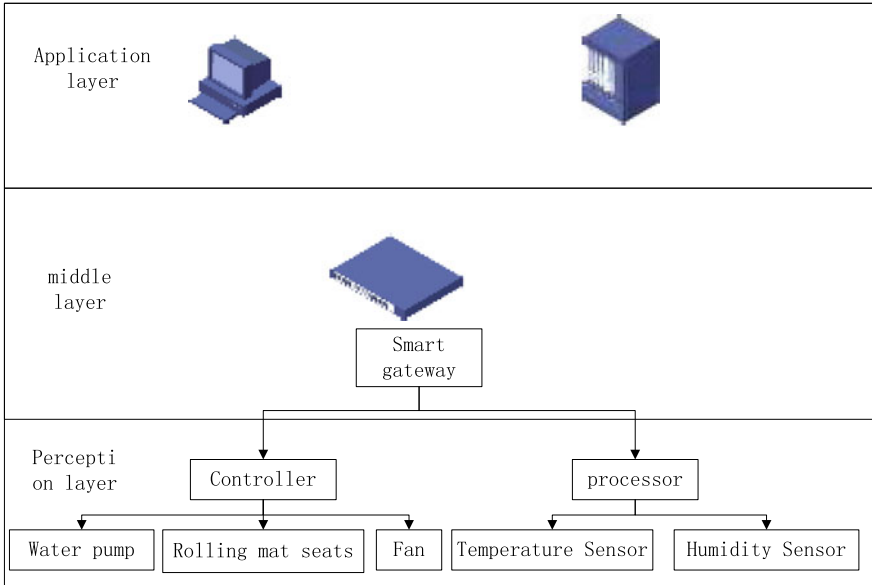


Fig. 1. The overall process of farm data monitoring

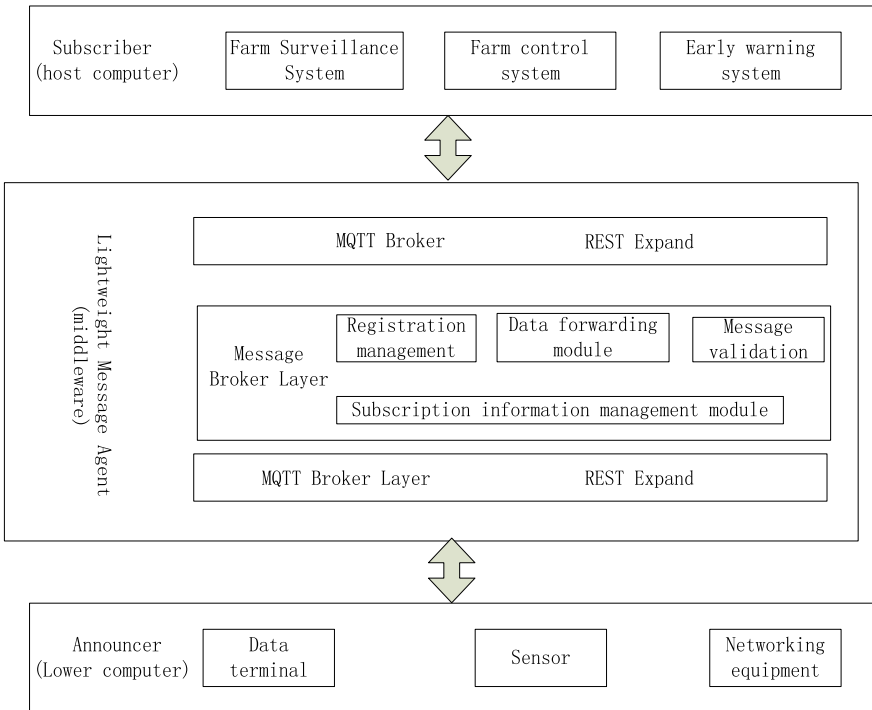


Fig. 2. System architecture design

information management, message matching, the implementation of the message storage and forwarding functions; Publisher to the remote terminal, for example, including mobile data terminals, sensors, and intelligent networking equipment, responsible for the acquisition and release of information.

In publish/subscribe systems, client participants include publishers and subscribers. The publisher and the subscriber of the message are decoupled, that is, the publisher of the message does not know the presence of the subscriber, nor does the subscriber need to know which publisher the message originated from, and neither the publisher nor the subscriber has a reference to the other party. Publishers and subscribers also decoupled in time, when the publisher sends the message, does not require subscribers online, when subscribers go online, they can still receive subscription news. The message broker can guarantee that the publishers and subscribers decoupled in time and space can accurately forward the message according to the message subscription mode.

The Lightweight Message Agent layer is the core functional layer of the publish/subscribe system, implemented as an MQTT Message Broker, responsible for connecting the publisher of the message with the subscriber of the message. The MQTT Message Agent is a server-side implementation that supports the MQTT Lightweight Transport Protocol. In this system, multiple MQTT message agents operate in a distributed system with network bridging mode, which implements message load balancing. The MQTT Message Broker is divided into two parts, the MQTT Broker Interface Layer and the Message Broker Message Service Layer. The MQTT Broker interface implementation layer provides the publisher or subscriber with interfaces for publishing messages, subscribing to messages, unsubscribing, and the like, and provides RESTful interface extension to the outside so that publishers and subscribers can conveniently send and subscribe to the HTTP request sent by a browser via a browser.

As shown in Fig. 2, Message Broker messaging service layer consists of a registration management module, a subscription information management module, a data forwarding module, and a message verification module, and also includes data stores such as subscription tables and routing tables. Among them, the registration management and subscription information management module is responsible for managing and maintaining subscriber information, and is responsible for real-time update subscription table, when new subscription information to join or cancel the subscription, in the subscription table to do the appropriate additions and deletions. The data forwarding module is responsible for forwarding the message posted by the publisher to the corresponding subscribers. The message checking module is responsible for verifying the validity of the message, and only the processed data forwarding module can obtain the message.

To achieve this, as shown in Fig. 3.

Cloud server software architecture design is based on the choice of server software running platform and monitoring system communication protocol. According to the stability requirements of the server software, choose the Linux operating platform, the server, and the device and other terminals using the MQTT protocol, cloud server software architecture shown in Fig. 3.

Linux platform is used by server platforms, because Linux system has been the first choice for high-performance server systems, the system kernel has many open-source functions. Agricultural Surveillance System Server software design focuses on the design of the software hierarchy and the related technologies used in communication,

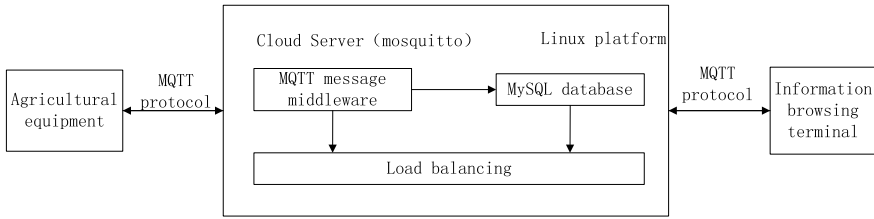


Fig. 3. Middleware server software architecture

as well as the core business processes for writing server software using the Linux system network programming and I/O reuse models. The core of the system cloud server software is designed with improved MQTT message middleware and deployed MySQL database. According to the characteristics of MQTT middleware, an appropriate load balancing algorithm is adopted to achieve cloud server load balancing.

System testing before the need to build a system test platform. Test platform set up includes two parts, server software server platform structure and test equipment terminal deployment. To ensure that the test environment is real and effective, cloud server software testing is conducted using agricultural equipment.

Agricultural remote monitoring system cloud server software based on cloud server platform deployment does not need to consider server hardware related deployment. By the use of cloud server to meet the deployment of server software requirements, the design will be deployed in the cloud server software Aliyun server, load balancing using Ali cloud load balancing module. Through the purchase of Alibaba Cloud computing services, according to the need to configure the server parameters, late if the configuration can't meet the needs of the relevant configuration can be changed to ensure system stability. Cloud server software platform single cloud server configuration is shown in Table 1.

Table 1. System test content list

Name	Configuration
Operating system	CentOS 7.2 ×64
Internal storage	8 GB
Bandwidth	100 Mbps
Hard disk	512 G

In order to ensure the stability of the cloud server, the cloud server software adopts cluster deployment. The single machine test in the cluster is deployed according to the structure in Table 1, and the cloud server configuration can be configured according to the actual scale of the agricultural device when it is put into use.

Deployment content in the cloud server mainly includes MQTT message middleware deployment, MySQL database deployment, and load balancing cluster deployment.

The use of agricultural remote monitoring system to test the basic functions of cloud server software and the function of each module and performance testing standards. Table 2 shows the main contents of the test.

Table 2. The main contents of the test

Test module	Test items
User login module	User login password
	User name and login password of MySQL database
Information communication	Post a message
	Subscribe
Access control	The user corresponding to the theme
	The user corresponding theme list
Data encryption	Encrypted data in transit
Message persistence	The subject and the message in the MySQL database
The server status monitoring	Subscribe to a specific theme to check the server status

The functional test does the testing according to the functions of each part of the cloud server software. The basic requirement of the test is whether the cloud service software can meet the requirements of the agricultural remote monitoring system and the business logic of each part is correct.

First, we use the online farm equipment to initiate the connection to the server software, and release the theme information. We can use the mobile client to subscribe to the theme information and check the running status of the online agricultural equipment. The message received by the mobile client shows that the server can receive and send the message correctly and the message is communicated correctly.

Server status monitoring subscribes related topics through the online agriculture equipment in the mobile phone client; selects server through the phone client; then views the server parameters.

On the premise that the function is correct, we need to do the server software performance testing. Test scheme is divided into two kinds, namely, a single machine test and multiple machine deployment test. Performance testing uses MQTT clients simulates the way of concurrent access to the server. After a simulated MQTT client establishing a connection with a server, the MQTT client sends data to the server and simulates multiple client subscription messages. Sending and receiving messages are attached to their times. So we can calculate and get the time when the messages are sent and received.

The performance test items of a single server include the number of server connections, server CPU, server memory, size of sending data and response time, the test results are shown in Table 3. The test results in Table 3 are the results when the MQTT client connection server publishes a message with QoS = 2. Different QoS settings, the test results will be different.

Table 3. Performance test results of a single server

The number of connections	CPU server utilization (%)	Server memory utilization (%)	The size of a single client sending the data	Completion time of subscription (s)
1000	1.5	1.2	5 KB	2
	1.6	15.4	50 KB	10
	2.1	62.4	1 M	19
	2.3	86.9	5 M	113
5000	8.1	3.1	5 KB	8
10,000	14.2	4.5	5 KB	21
20,000	26.3	9.1	5 KB	53
30,000	37.4	16.7	5 KB	94
40,000	46.2	20.4	5 KB	97
50,000	67.6	26.3	5 KB	104

After the test, the number of MQTT messaging middleware connections in a single server software can reach around 150,000. Before the improvement of Mosquitto, the test results show that the number of connections is 50,000 or so, the improvement is obvious. If a client issues a message, the server's CPU, memory, and the sending data will all consume server resources. The server CPU utilization is mainly affected by the number of connections, the more connections, the higher the server CPU usage. Due to the different content of the message posted by the client, the resource usage of the server is different. In particular, the memory of the server increases with the increase of sending data by a single client, and changes greatly. It can be seen from the MQTT protocol is suitable for sending short message data, not large files.

Server cluster testing, which is dynamically allocated by load balancing to each node server in the cluster. Cluster deployment is tested with five servers, one of which deploys a load balancing server, while the other four make up the server cluster, two of which are central node servers. The MQTT client connection first accesses the load balancing module, then allocated by the load balancing module to the server node.

The test items include cluster connection number, average server CPU utilization, average server memory utilization, sending data size, and response time. The client connection number of the cluster test is the same as the number of client connections tested by a single server. Under this premise, the response speed of the test cluster is tested, and the test results are shown in Table 4. As you can see from the test results, the cluster server deployment can share the server's connections to each node server. With the same number of connections, the CPU and memory usage of each node server is lower than that of a single server connection. When the number of client connections is large, the time for the client to subscribe is larger than that of the single server.

Table 4. Performance test results of cluster server

Connections number	Server CPU utilization (%)	Server memory utilization (%)	Sending data size of single client (KB)	Subscription time (s)
1000	0.3	0.4	5	3
5000	2.1	0.6	5	4
10,000	3.8	1.1	5	6
20,000	6.9	2.4	5	12
30,000	10.4	3.1	5	20
40,000	12.7	4.6	5	27
50,000	17.5	2.8	5	34

1 Conclusion

This paper designs cloud server software for the agricultural remote monitoring system. The software design starts from the demand analysis of agricultural remote monitoring system, then from system architecture design to detail and implementation, finally the system test proves the rationality of the design of the server software. Cloud server software provides a messaging platform for the agricultural remote monitoring system. The user can remotely monitor agricultural equipment and check the relevant data of the device through the mobile terminal. By connecting agricultural equipment to the Internet and conducting remote monitoring, this design is conducive to the intelligent upgrading of the agricultural field and conforms to the trend of future agricultural intelligence.

References

1. Daoliang, L.: Introduction to Agricultural Internet of Things. Science Press, Beijing (2012)
2. BieHua, W.: The application of the internet of things agriculture to the needs of modern scientific instruments. *Mod. Sci. Instrum.* **3**, 5–6 (2010)
3. Tang, K.L., Wang, Y., Liu, H., et al.: Design and implementation of push notification system based on the MQTT. In: International Conference on Information Science and Computer Applications. ISCA 2013, pp. 116–119 (2013)
4. Heng, R., Yue, M., Haibo, Y., et al.: Message push server based on MQTT protocol. *Comput. Syst. Appl.* **23**(3), 77–80 (2014)
5. Xing, W.: Software Architecture Design, pp. 43–64. Electronic Industry Press, Beijing (2007)
6. Forta Writting, B., Xiaoxia, L., Ming, Z. (translation): MySQL will be known. People Post Press, Beijing, pp. 2–56 (2009)
7. Yan, Bo, Yan, Chang, Ke, Chenxu, et al.: Information sharing in supply chain of agricultural products based on the internet of things. *Ind. Manage. Data Syst.* **116**(7), 1397–1416 (2016)
8. Kshetri, Nir: The economics of the internet of things in the Global South. *Third World Q.* **38**(2), 311–339 (2017)

9. Sun, X., Yang, Y., Guo, H.: Applications of near field communication of internet of things in supply chain information system of agricultural products. *Trans. Chin. Soc. Agric. Eng.* **30** (19), 325–331 (2014)
10. Brewster, Christopher, Roussaki, Ioanna, Kalatzis, Nikos, et al.: IoT in agriculture: designing a Europe-wide large-scale pilot. *IEEE Commun. Mag.* **55**(9), 26–33 (2017)
11. Benaissa, S., Plets, D., Tanghe, E., et al.: Internet of animals: characterisation of LoRa sub-GHz off-body wireless channel in dairy barns. *Electron. Lett.* **53**(18), 1281–1283 (2017)
12. Wang, JunPing, Duan, ShiHui, Shie, YouKang, et al.: Multi-objects scalable coordinated learning in internet of things. *Pers. Ubiquit. Comput.* **19**(7), 1133–1144 (2015)
13. Gardiner, S.: Business re-imagined: IIoT is set to transform energy, agriculture, manufacturing and more. *Maclean's* **128**(38), 44 (2015)



Architecture Design of Modern Marine IoT Cloud Server Software

Ling Yu¹(✉), Yang Gao², Chaoyan Man², and Yu Zhang²

¹ Network Information Center, Dalian Polytechnic University, Dalian, China
yuling@dlpu.edu.cn

² School of Information Science and Engineering, Dalian Polytechnic University, Dalian, China

Abstract. Internet of things in the field of marine fisheries develops very fast. The current fishery servers have the problems of insufficient concurrency capability, low stability, small scalability, and difficulties in maintaining, which cannot meet the requirements of stable data transmission in high concurrency mode. Since most fishing equipment is embedded and has limited computing resources, remote monitoring and control systems for fisheries require that communication modules can be easily integrated into the equipment and occupy a small amount of equipment computing resources. This article introduces the MQTT protocol subscription/publishing features, and designs the cloud server software architecture.

Keywords: Internet of things · Marine fisheries · Remote monitoring of fisheries · MQTT agreement

The MQTT protocol is an application layer protocol based on the TCP/IP protocol. The client and the server transmit the messages by subscribing/publishing the theme of the message, the content of which is the load of the message theme. The MQTT server is called message middleware, and the message middleware is the application software running on the server designed for the MQTT protocol.

The message transmission between MQTT clients needs to go through the message middleware. The message middleware has the forwarding function. The MQTT client and the message middleware use the publish/subscribe mode to transmit messages, and the transmission mode is as shown in Fig. 1.

In Fig. 1, after the three MQTT clients establish a connection with the messaging middleware, each MQTT client subscribes to and publishes the messages according to their own business needs. The MQTT client 1 posts the message to the message middleware, which publishes the message to the client 2 and the client 3 subscribing to the message according to the subscription. Similarly, client 2 and client 3 publish messages in the same mode as client 1 posts messages.

Cloud server software architecture design is based on the choice of server software running platform and monitoring system communication protocol. According to the stability of the server software requirements, select the Linux platform, server and device and other communication protocol between terminals using MQTT protocol, cloud server software architecture shown in Fig. 2.

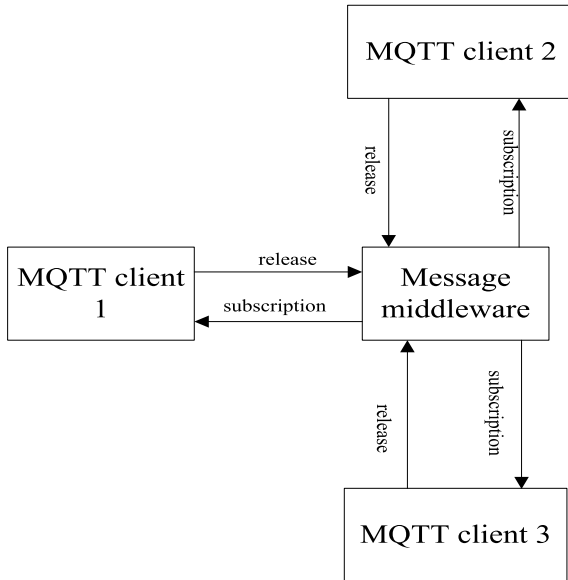


Fig. 1. MQTT publish/subscribe model

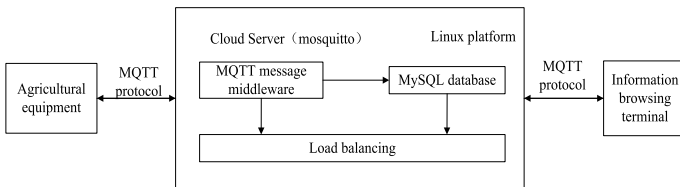


Fig. 2. Cloud server software architecture

According to the demand of the Marine fishery remote system, the function of the cloud server software demand mainly includes seven parts, respectively for the user login validation, business processing, message persistence, access control, data encryption, load balancing, and server status monitoring, as shown in Fig. 3.

The Cloud server software running in the Internet environment and the corresponding IP address on the Internet opening to the outside world are the ways to have a risk of malicious programs to connect server. Designing the user login authentication module for specific users to connect to the cloud server is the basic requirement of the server software, so the Marine fisheries have cloud server monitoring and control system software to design the login authentication module. This module is the first step on the device to connect server software, the user through a user name and password for connection to the server, the server to verify the user name and password. If the user name and password are correct through validation, user successfully logs into the server; Otherwise, the user login to the service software fails. Cloud server to authenticate a user name and password should be a timely response, because Marine

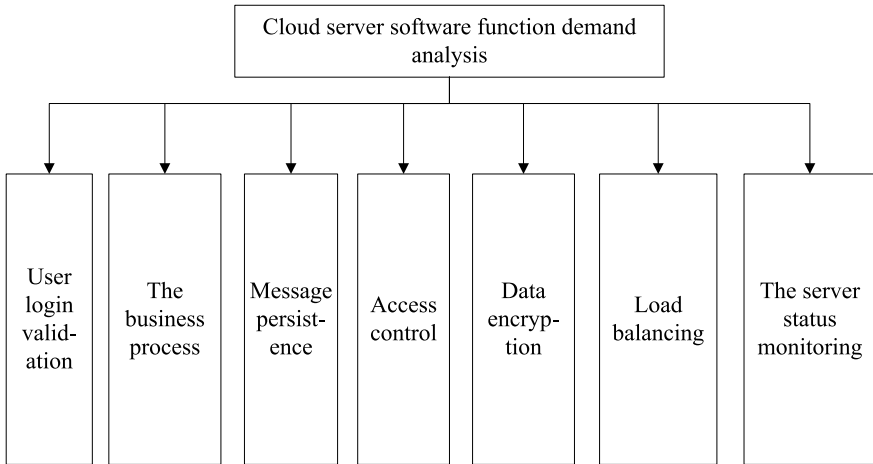


Fig. 3. Industrial monitoring system server software functional requirements

fishery equipment is higher to the requirement of real-time response speed and can directly affect the data transmission speed, so the user name and password need to deposit the response quickly to the database or file.

After a user successfully logs into the server, it sends data to the server or obtains data. The core of the cloud server software is to handle the user's connection and data services. As the number of industrial devices in marine fisheries remote monitoring systems is dynamically changing, server software needs to consider situations where high-concurrent access to marine fisheries equipment is required and it should have the capacity to handle high-concurrent client connections.

The cloud server software design should consider the storage of data between the device and the cloud server. The remote monitoring of the marine fishery has the demand of viewing the historical information of the device. The server software needs to provide an interface for implementing this function. After the data sent by the device to the server is persistently stored in the database after being processed by the server software, the mobile terminal can view the content of the message sent by the device to the server by accessing the database.

Cloud server software needs to add access rights to the device or user. Because marine fisheries are relatively independent devices, messaging needs to be independent of each other and need to be tailored to the needs of a particular device to access only specific message content. The cloud server software needs access control so that the device subscribes to and receives related message contents according to the permission list. This method can limit the access rights of the device. Only the users or devices with super privileges can view the contents of all the messages processed by the server.

User or device data are uploaded to the server via the Internet. The Internet is an open network. Data can be transmitted on the network in clear text without any data processing. To ensure data security, the data needs to be encrypted. Server software and client data encryption needs to be considered on the basis of the communication protocol to achieve, and also need to consider the computing resources of the limited end of the encryption algorithm or protocol carrying capacity.

The ability of a single server handling the connection of the device is limited. When the number of equipment connections is large, we need to consider the load balancing among multiple servers. Load balancing not only includes the topology of load balancing and load balancing algorithm, but also includes the data synchronization between multiple servers. If we put data persistence into a database, the database will support distributed configuration. Then using a load balancing algorithm to dynamically configure the connection between the device and the server ensures the data consistency across multiple servers.

Server status monitoring is an important part to ensure the stable operation of server software. The items that the server software needs to monitor include the number of device connections, server memory, and hard disk usage. By monitoring these contents, you can get a real-time control of server health status and troubleshoot problems related to server software operation before the server fails.

Server software design meets the functional requirements of the marine fisheries remote monitoring system also need to meet the performance requirements. Performance requirements mainly include three aspects, one is real-time, the second is the carrying capacity, the third is the coupling of system and device-side program, as shown in Fig. 4.

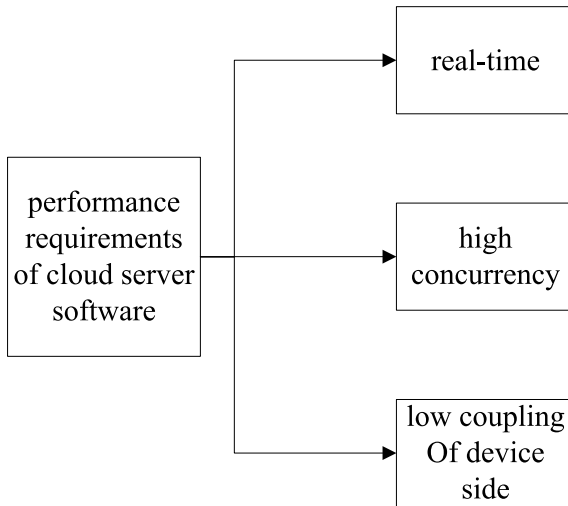


Fig. 4. Performance requirements map of cloud server software

1 Conclusion

This article describes two parts of the cloud server software, which are requirement analysis and architecture design. First we introduce the remote monitoring and control system of marine fisheries, and analyze its working mode. According to the requirements of monitoring system in marine fishery, the requirements of cloud server software are analyzed, and the cloud server software is analyzed in detail from function and

performance aspects. On the basis of requirement analysis, the architecture of cloud server software is designed, including the choice of operating platform, communication protocol, database, and so on. We choose cloud server software as MySQL database, according to the requirements of the storage requirements, we design MySQL database storage structure.

References

1. Davoli, L., Protskaya, Y., Veltri, L.: An anonymization protocol for the internet of things. In: 2017 International Symposium on Wireless Communication Systems (ISWCS), Beijing (2017)
2. Bertino, E., Islam, N.: Botnets and internet of things security. *Computer*. **50**, 76–79 (2017)
3. Ng, Irene C.L., Wakenshaw, Susan Y.L. et al (2017) The Internet-of-Things: Review and research directions. In: *International Journal of Research in Marketing*. Vol.34: 3–21
4. Caputo, A., Marzi, G., Pellegrini, M.M.: The internet of things in manufacturing innovation processes. *Bus. Process Manag. J.* **22**, 383–402 (2016)
5. Garlock, T.M., Camp, E.V., Lorenzen, K.: Using fisheries modeling to assess candidate species for marine fisheries enhancement. *Fish. Res.* **186**, 460–467 (2017)
6. Essington, T., Ciannelli, L., Heppell, S., et al.: Empiricism and modeling for marine fisheries: advancing an interdisciplinary science. *Ecosystems* **20**, 237–244 (2017)
7. Rolim, F.A., Ávila-da-Silva, A.O.: Effects of marine protected areas on fisheries: the case of São Paulo State, Brazil. *Latin Am. J. Aquat. Res.* **44**, 1028–1038 (2016)
8. Klemas, Victor: Fisheries applications of remote sensing: an overview. *Fish. Res.* **148**, 128–136 (2013)
9. Watson, G.J., Murray, J.M., Schaefer, M., et al.: Bait worms: a valuable and important fishery with implications for fisheries and conservation management. *Fish Fish.* **18**, 374–388 (2017)
10. Stephen, C.M., Paul, J.D., Thomas, L.C., et al.: Approaches to fully documented fisheries: practical issues and stakeholder perceptions. *Fish Fish.* **16**, 426–452 (2015)
11. Wood, G., Lynch, T.P., Devine, C., et al.: High-resolution photo-mosaic time-series imagery for monitoring human use of an artificial reef. *Ecol. Evol.* **6**, 6963–6968 (2016)
12. Chassot, E., Bonhommeau, S., Reygondeau, G., et al.: Satellite remote sensing for an ecosystem approach to fisheries management. *ICES J. Mar. Sci.* **68**, 651–666 (2011)



Three-Dimensional Structure Measurement and Optimization Method of Indoor Scene Based on Single Image

Ronghe Wang^(✉), Xinhai Zhang, Bo Zhang,
Jianning Bi, and Xiaolei Guo

National Engineering Laboratory for Public Security Risk Perception and Control
by Big Data (PSRPC), China Academy of Electronics and Information
Technology, China Electronics Technology Group Corporation,
No. 11 Shuangyuan Rd, Shijingshan District, Beijing 100041,
People's Republic of China
wangronghe@buaa.edu.cn

Abstract. In the paper, we detect 2D surfaces and recognize 3D structures in images, and generate occluded part of three-dimensional structure. We detect objects vanish points and geometric segmentation lines, and generate the vector and normal plane. We measure the objects boundaries which have been detected. In the process, we have found the best linear segmentation positions, generated buildings, and other 3D models in room. We propose some effective assumptions to recognize the object model from images by geometric reasoning. At the same time, we put forward structure prediction technology combined with the volume reasoning by parameter representation of spatial objects. We detect scene model relationships of each other by combining image-rich appearance with geometric features. We proposed image geometric modeling grammar framework according to previous discriminative classifier. It is used to represent the physical structure of the visual component. The framework has broken the traditional probability texture context grammar tree model. We proposed spatial context theory and model generation rules. Finally, on the public existing data sets, we proved that we only use structure prediction and linear segmentation of scene to recover 3D structure by a lot of experiments, comparing to restoration algorithm using whole image structure, ours method can produce more credible scene entity models and space constraint relationships.

Keywords: Models detection · Geometric reasoning · Energy minimization method · Models generation · Scene grammar model

1 Introduction

Using the relationship between objects, extract 3D geometry information can be traced back to the origin of the computer vision. Forty years ago, in the early days of the computer vision, researchers extracted linear information from Block world scene [1]; however, the method cannot handle errors from the extracted line or other shape information. In information processing systems, it is an important task to recover 3D

scene from 2D indoor images. Researchers often understand visual hierarchy structure of indoor images by recovering scene probability grammar (SSH). Restoration of three-dimensional objects in a complex scene by a single image not only relates to object recognition, but also need to extract the 3D room space layout information (e.g., ceiling, floor, wall, etc.). The extraction of indoor space layout needs visual recognition technologies to provide critical geometric information. There have been some qualitative prediction methods for image parameter model.

2 Related Work

In recent years, some more practical algorithms extracting 3D structure in limiting threshold has been proposed. Such as Manhattan Worlds [2] and improved line segmentation method using Manhattan world [3], however, these methods only focus on direction of three-dimensional world and do not try to extract target information from indoor scene information. Manhattan world hypothesis [4] is used to gain layout of the single indoor image in common monocular reconstruction algorithm [5]. These methods can be classified as discriminative approaches and generative approaches.

2.1 Discriminative Approaches

Conditional random field model (CRFs) [6] focuses on each pixel semantic label (such as building, sheep, road, ship, etc.); however, these methods do not understand scene structure. Pixel level label is not enough to express the problem of object block and three-dimensional object relationship, etc. Hoiem [7] and Saxena [8] et al. put forward different methods to generate surface to explore geometric feature vector. In order to measure and recover indoor scene with structured support vector machine [9], Lee [10], Wang [11], Hedau [12] et al. modeled background, foreground, and geometric layout using different methods.

2.2 Generative Approaches

Generative approaches mainly detect the key features of image. Zhu and Mumford [13] proposed the generative method driving AND/OR graph to express visual hierarchy in order to measure model size and calculate model relationship. Han and Zhu [14] detect rectangular structure in artificial scene by the sub-tree reapplication adapter theory [15], Porway et al. [7] generate scene configuration in the C4 algorithm introduced Markov chains model, and these algorithms are typical of generative approaches.

2.3 Others Methods

Some methods modeling geometric structure from single image, and generate geometric labels and classification [16], and then found the folding line perpendicular to the surface. Lee [13] proposed an indoor structural parameter model which was calculated by building block world size constraint; however, the sampling process of this model may have errors in spatial layout assumption, resulting in the deviation of

calculation process. Hedau [17] et al. used appearance based on classifier and visual features to compute regions of no confusion in 3D scenes, and used structured methods to compute rectangular structures in images. In the literature [18], Wang et al. proposed a similar approach that did not require the use of standard contrast results to mark the chaotic parts of an image. Sugihara [8] provided a boundary optimization method based on line detection; however, this method did not generate 3D objects. Han and Zhu [9] were devoted to finding vanishing points of rectangular structures from linear structures.

3 Ours Method

In this paper, we detect straight-line and orthogonal vanishing point by the detected straight line in images. Detection building structure is divided into four steps: detecting segmentation line, detecting vanishing point, generating normal plane, and generating 3D scene configuration finally.

3.1 Detection of Three-Dimensional Orthogonal Vanishing Point and Segmentation Line

We put forward some fixed building hypothesis and tested each hypothesis and find the best match (Fig. 1).



Fig. 1. The process of detection, measurement of boundary line, and generation of normal plane

We use MATLAB own algorithm Kovese [11] to detect straight line, restore the direction of building according to three vanishing points, and generate three-dimensional models. We used the random sampling detection 3D orthogonal vanishing points in the image according to the method of Rother [12]. We also detect corners and generate normal plane according to the cost function [7], and the formula is shown as follows (Fig. 2):

$$V_A(\lambda, p) = |\lambda| \exp(-\beta/2\sigma^2)$$

3.2 Corner Generation Rules

Segmentation line detection in the scene is shown below (Fig. 3).

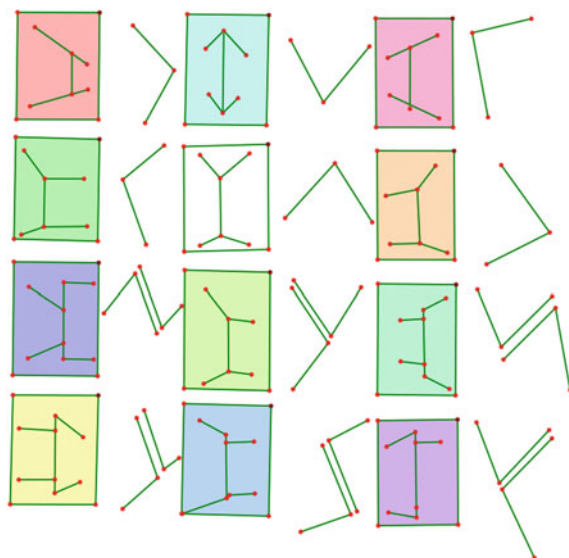


Fig. 2. The detection of segmentation line and corner



Fig. 3. The detection of segmentation line in scenes

In this paper, an additional prediction item is added in order to detect corner. The prediction item added in this paper is shown as follows (Fig. 4):

$$v_P(p) = \exp(-(\phi(\lambda)/\psi)^2)$$

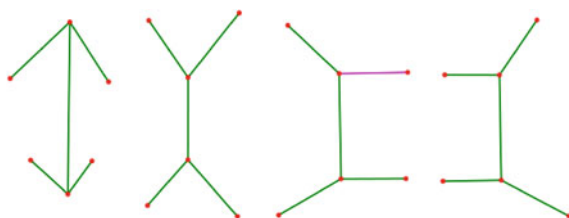


Fig. 4. The diagram of grammatical structure about corner detection

3.3 Surface Generation Rules

In order to enhance the surface detection, this paper introduces a pattern parameter to increase calculation precision. Scene middle surface generation process is shown in the following (Fig. 5).

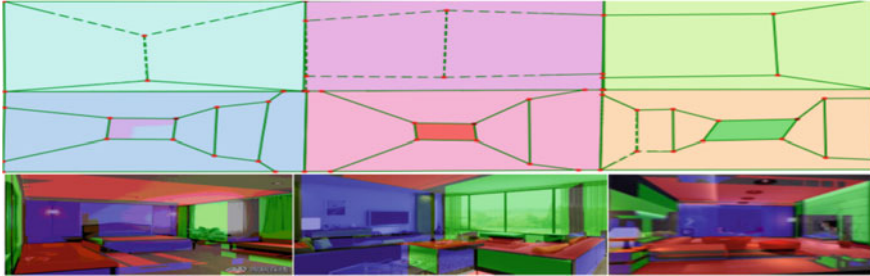


Fig. 5. The computation of the normal plane in scene

3.4 The Theory of Generating Scene Optimal Configuration

Many of the indoor environment can meet the Manhattan building blocks world theory [14], such as most of the planes are in three orthogonal directions to each other, in addition, indoor environment often has a single floor and ceiling, and the ceiling brightness is close to a constant. Another property of indoor scene is symmetry between ceiling and floor. In conclusion, this algorithm is based on the following assumptions:

1. Each object in scene has a nonzero finite volume.
2. In order to test the volume location of indoor and related items, we assume that all items are on floor.
3. In order to test objects whether in room or on the ground, we need to check the projection of them whether in room or on the ground. If they are not in room, it will remove from relevant items.
4. In order to test volume conclusion relationship of two objects, we assume that two objects are in the same plane, and then projects two objects on the plane, if the projection is together, it means they have cross volume.
5. Objects cannot intersect; therefore, the object volume of different stations is of mutual exclusion.
6. Each object should be included in the free space scene wall.
7. If given the detection hypothesis $\{x_1, x_2, \dots, x_n\}$ and related objects hypothesis $\{o_1, o_2, \dots, o_n\}$ of image, we calculate the relationship between the detection results and target according to weight of computation $R = (x_1, o_1), \dots, (x_n, o_m)$, among them, R represents the reasonable allocation of scene.

3.5 Generation Geometry Rules for Scene Configuration

It is necessary to define a series of geometric rule constraints, such as defining walls, corners, etc. This paper defines two geometric rules for generating scene: “and rules” and “or rules”. (1) “and rules” defining a plurality of vanishing point according to detection results originally belong to one object, and generating vertex and plane of an object; (2) “or rules” defining a plurality of vanishing point detection results originally belong to different objects to generate different object vertex and plane (Fig. 6).



Fig. 6. The assumptions of scene configuration according to above rule

4 Experiments

Finally, we test ours algorithm by the experimental data including 102 images of York Urban database [10], UIUC database [7, 16], and 44 images from the network are collected by Delage et al according to average error of area under ROC curve (AUC-ROC) to evaluate the quality of ours algorithm (Fig. 7).

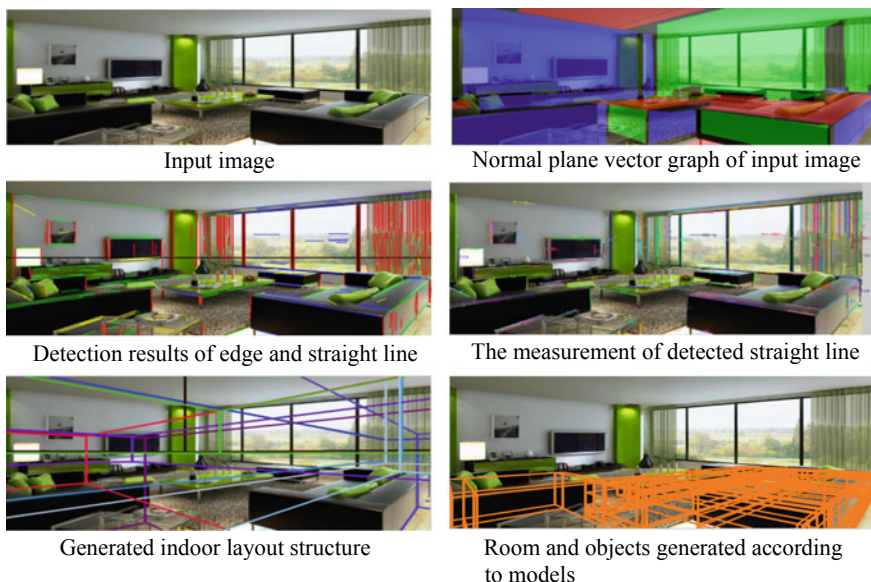


Fig. 7. The results of our experiments

The first group experiments: We test ours algorithm on indoor data set introduced in [11]. The data set consists of 314 images, including spatial layout annotated images and non-labeled cluttered layout images (209 training images and 105 test images). After testing, the average error of ours algorithm is 26.5% in chaotic scene without annotation. In the annotated scene, the average error of ours algorithm is about 16.2% (Fig. 8).



Fig. 8. We test ours algorithm on indoor data set introduced in [12]

The second group experiments: We test ours algorithm on the UIUC dataset. The UIUC dataset contains 314 cluttered indoor images, in which the standard contrast result is a label map of background layout without foreground objects. In this paper, the experimental group data has 220 images and 6 aspects scenes. The standard contrast results are manually segmented scene, and ours algorithm achieves good results on this dataset (Fig. 9).



Fig. 9. We test ours algorithm on the UIUC dataset

Table 1 is the segmentation accuracy comparison between ours algorithm and other algorithms in UIUC database.

The third group experiments: This paper collected 54 indoor images, including the case that some objects are blocked. We used the experimental design similar to the literature [14] and marked standard contrast results manually for each pixel. The average error and AUC-ROC of this algorithm in this dataset are 20% and 13%, respectively (Fig. 10).

The fourth group experiments: We tested 44 images collected by Delage et al. and other network images. Ours target focus on object against the wall and on the floor. We detect straight line in this experimental dataset, measure linear distance, generate normal plane, and then generate relevant scene configuration. The error of our method is 19.5% on this dataset (Fig. 11).

Table 1. The compared results of measurements precision on UIUC database

Experimental groups	Hoiem' method [10] (%)	Hedau' method [12] (%)	Wang' method [14] (%)	Lee' method [15] (%)	Ours method
The first group	73.50	79.90	81.40	78.80	89.66
The second group	74.56	81.87	83.80	85.36	87.32
The third group	78.62	84.89	86.87	89.67	90.30
The forth group	87.32	90.98	91.37	92.88	94.67

**Fig. 10.** We test ours algorithm on indoor images collected in this paper**Fig. 11.** We test ours algorithm on 44 images collected by Delage et al. and other network images

5 Conclusions

In this paper, we proposed an image 3D structure measurement and optimal configuration method based on vanishing point detection, normal plane generation, and geometric rule constraint. Our algorithm has multiple occlusion objects cross-checking, measuring, and generating configuration rules. At the same time, we introduced a method to generate indoor scene structure and 3D mode in single indoor image. At the same time, we proposed a more optimized reasonable allocation hypothesis of scene. In this paper, we first estimate and detect vanishing points, generate segmentation lines and normal planes in indoor images, and also our method improves volume reasoning theory.

Acknowledgements. We sincerely thank the reviewers and editors' work to this paper. This paper is supported by National Natural Science Foundation Projects of China, the national key research and development program of china: Research and development of intelligent security card port monitoring and warning platform (Grant No. 2016YFC0800507), Innovation Foundation Program of China Electronics Technology Group Corporation: Research on holographic and abnormal behavior intelligent warning technology for social security risk targets.

References

1. Gupta, A., Satkin, S., Efros, A.A., Hebert, M.: From 3d scene geometry to human workspace. In: CVPR (2011)
2. Lee, D., Gupta, A., Hebert, M., Kanade, T.: Estimating the spatial layout of rooms using volumetric reasoning about objects and surfaces. In: NIPS (2010)
3. Hedau, V., Hoiem, D., Forsyth, D.: Recovering the spatial layout of cluttered rooms. In: ICCV (2009)
4. Tu, Z., Bai, X.: Auto-context and its application to high-level vision tasks and 3d brain image segmentation. PAMI (2009)
5. Saxena, A., Sun, M., Ng, A.: Make3d: learning 3d scene structure from a single image. PAMI (2008)
6. Lee, D., Gupta, A., Hebert, M., Kanade, T.: Estimating spatial layout of rooms using volumetric reasoning about objects and surfaces. In: Advances in Neural Information Processing Systems, vol. 7, pp. 609–616. MIT Press, Cambridge, MA (2010)
7. Wang, H., Gould, S., Koller, D.: Discriminative learning with latent variables for cluttered indoor scene understanding. In: ECCV (2010)
8. Hedau, V., Hoiem, D., Forsyth, D.: Recovering the spatial layout of cluttered rooms. In ICCV (2009)
9. Han, F., Zhu, S.C.: Bottom-up/top-down image parsing with attribute grammar. PAMI (2009)
10. Porway, J., Zhu, S.C.: C4: computing multiple solutions in graphical models by cluster sampling. PAMI **33**(9), 1713–1727 (2011)
11. Saxena, M.S., Ng, A.: Make3d: learning 3d scene structure from a single image. PAMI (2008)
12. Delage, E., Lee, H., Ng, A.: A dynamic bayesian network model for autonomous 3d reconstruction from a single indoor image. CVPR (2006)
13. Han, F., Zhu, S.: Bottom-up/top-down image parsing by attribute graph grammar. In: Proceedings of International Conference on Computer Vision (ICCV) (2005)
14. Lee, D., Hebert, M., Kanade, T.: Geometric reasoning for single image structure recovery. In: Proceedings of CVPR (2009)
15. Hedau, V., Hoiem, D., Forsyth, D.: Recovering the spatial layout of cluttered rooms. In: Proceedings of ICCV (2009)
16. Hedau, V., Hoiem, D., Forsyth, D.: Thinking inside the box: using appearance models and context based on room geometry. In: ECCV (2010)
17. Denis, P., Elder, J.H., Estrada, F.: Efficient edge-based methods for estimating manhattan frames in urban imagery. In: ECCV (2008)
18. Delage, E., Lee, H., Ng, A.Y.: A dynamic bayesian network model for autonomous 3d reconstruction from a single indoor image. In: CVPR (2006)



The Intelligent Supervision System of Farm Based on “Internet + BDS + GIS”

Wei Fu ^(✉), X. R. Dong, Weiyi Shuai, Mingqi Yang, and Jun Wang

Space Engineering of University School Space Information, No 1, Ba Yi Road,
Huairou 101416, Beijing, China
1098761931@qq.com

Abstract. This paper proposes a method about the intelligent supervision system of farm based on “Internet + BDS + GIS”. Using Internet of Things (IOT) to monitor and dispatch the management of farm machinery. Based on BDS/GIS/GPRS and communication technology, the system can realize intelligent monitoring and management for farms. The system consists of four subsystems: agricultural machinery monitoring and dispatching management system, automatic driving system of agricultural machinery based on BDS, information management system and data management system. The GPRS is used to achieve instant connection, and it can also improve the channel utilization rate, transmission rate, and the quality of communication service. The agricultural machinery monitoring and dispatching management system, automatic driving system of agricultural machinery based on BDS, and information management system are based on the positioning module of the BDS, which can obtain the position information and ensure the continuous stability of the service. GIS technology is applied to farm monitoring and management, to realize real-time monitoring of farmland information and the collection and analysis of data fast, efficient, and comprehensive, as well as the monitoring and management of related personnel, machinery, and livestock. GIS makes the management more scientific and intelligent.

Keywords: Intelligent supervision system of farm · Internet · BDS · GIS · GPRS · Intelligent

1 Introduction

With the continuous improvement of the automation and information level of agricultural production, how to monitor and control the agricultural field environment convenient and effective is the focus of agricultural information research. The remote monitoring system of agricultural information is a comprehensive system which is combined with various technologies, including electronic technology, computer technology, communication technology, automatic control technology, sensor technology, and so on. It can collect and analyze the information on the farm site through the intelligent software that runs in the monitoring center. The farmers will be able to take the information of the whole farm and respond to emergencies in time and accurately. However, the traditional video monitoring system has the defects of simple function

and low degree of automation, and the existing remote monitoring system has the disadvantages of low-cost performance and complex system. How to effectively integrate related technologies in other fields and develop a high-performance and low-cost farm monitoring and management system is a problem to be solved urgently.

In view of the shortcomings of the existing system deficiencies and actual demand, this paper proposes a method about the intelligent supervision system of farm based on “Internet + BDS + GIS”. Based on the BDS, GIS, and GPRS, the intelligent monitoring and management of the farm can be realized, and it can provide the managers with comprehensive, timely and accurate services, and various information inquiries. Taking farm management as the core and demand oriented, we can not only realize the monitoring and management in the control centers, but also achieve monitoring and management based on mobile phone terminals.

2 Key Technology

2.1 BeiDou Satellite Navigation System—BDS

BDS is a satellite navigation system built and run independently by China. It can provide all-weather, all-day, and high-precision positioning, navigation, and time service for global users. As the third generation of BeiDou system (BD3) opened the global networking stage in 2018, and the application of the second generation system (BD2) continued to expand and popularize, its application in the agricultural field realized the automation of farmland management and opened a new mode of agricultural production. In the field of agriculture, the application of BDS can reasonably control the work of planting, fertilizing, irrigation and harvesting, and so on, effectively reducing the human expenditure and improving the efficiency of agricultural production. The combination of BDS and agricultural machinery can make the modernization of agriculture transform into fully mechanized automation, and with the growing maturity of BDS, it will become a reality.

2.2 Geographic Information Systems—GIS

Geographic information system, referred to as GIS, is an edge technology subject which has been developed rapidly in 1960s, and can integrate geospatial data processing and computer technology. It is a technical system of collecting, storing, managing, processing, retrieving, analyzing, and displaying the geographical distribution of space objects and related attributes and answering user questions as the main task under the computer software and hardware. From a technical point of view, the application of GIS in agriculture is mainly reflected in the following aspects:

1. As a tool for investigation and analysis of agricultural resources;
2. As a tool for the management of agricultural production; and
3. As an auxiliary decision-making tool for agricultural management.

In the field of precision agriculture, GIS is able to manage the basic information, store, analyze, and process spatial data with high spatial resolution. Thus, the spatial

distribution map of crop yield, soil properties and disease and insect pests, and other environmental factors are generated, and it can also support space-aided decision-making and output graphics and geographical statistics and field decision-making drawings.

2.3 General Packet Radio Service—GPRS

General Packet Radio Service (GPRS) is the abbreviation of general packet radio service. It is a packet switching data carrying and transmission way developed on the basis of GSM. Different network users share the same group of GPRS channels. Only when a user needs to send or receive data, will he occupy the channel resources. In this way, multiuser business reuse can make more effective use of wireless network channel resources. Through the GPRS wireless connection Internet, that is, using the GPRS module, the farmland information collector is connected to Internet through GPRS wireless network to realize the remote Internet and data transmission, which not only covers a wide coverage, but is also fast in transmission speed and has low cost and more economic benefit.

3 System Design

This paper proposes a method about the intelligent supervision system of farm based on “Internet + BDS + GIS”. Based on the BDS, GIS, and GPRS, the intelligent monitoring and management of the farm can be realized, and it can provide the managers with comprehensive, timely and accurate services, and various information inquiries. The overall structure of the system is illustrated in Fig. 1.

3.1 Agricultural Machinery Monitoring and Dispatching Management System

The structure diagram of agricultural machinery monitoring and dispatching management system is shown in Fig. 2, including vehicle BDS positioning unit and agricultural machinery monitoring and dispatching center as two parts. The vehicle terminal consists of the BDS receiving module, the GPRS transceiver module, the MCU main control module, and various sensors. The BDS receiving module is used to receive the data of the positioning satellite and the differential datum station, and calculate the geographical position coordinates of the agricultural machinery itself. All kinds of sensors acquire information about farmland temperature, humidity, and seeding depth. Through the GPRS transceiver module of the vehicle terminal, all kinds of data such as the comprehensive information like location, state, and alarm of agricultural machinery are sent to the central server with fixed IP address and stored in the center database using the GPRS network. At the same time, the corresponding alarm information is sent to the corresponding client monitoring scheduling unit. The monitoring and

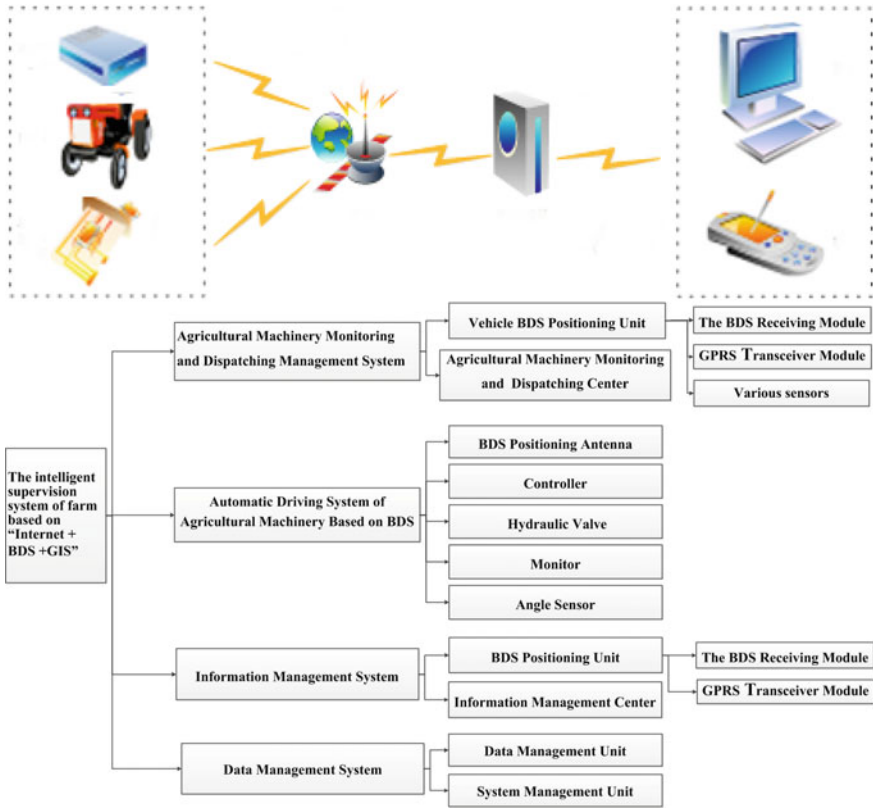


Fig. 1. The overall structure diagram of the system

dispatching unit connects with the central server through the Internet network, receives the data from the center server in real time, matches the agricultural machinery coordinate information with the electronic map, and displays the correct position of the agricultural machinery on the map, so that the monitoring and dispatching center can grasp the dynamic position information of agricultural machinery clearly and intuitively and the machine can be monitored in real time. The monitoring and dispatching center can receive the information of agricultural machinery position and the situation of agricultural machinery operation, dispatch the agricultural machinery reasonably by invoking the agricultural machinery dispatching model, and transfer the dispatching results to the central server through the Internet network, and then send the dispatching results to the corresponding agricultural machinery vehicle terminals through the GPRS network to realize the rational dispatching of agricultural machinery resources.

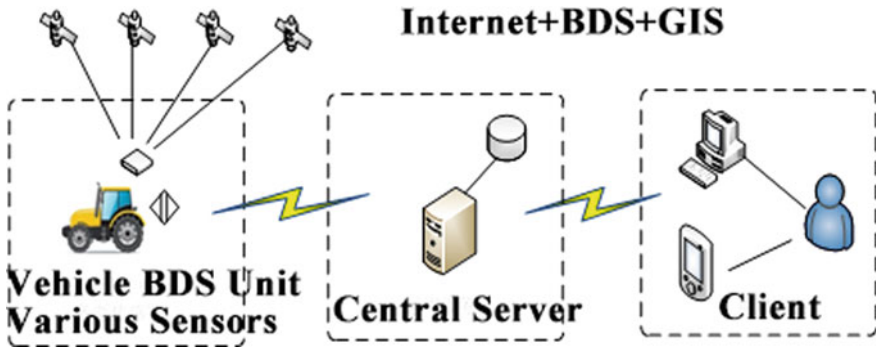


Fig. 2. The structure diagram of agricultural machinery monitoring and dispatching management system

3.2 Automatic Driving System of Agricultural Machinery Based on BDS

The automatic driving system is shown in Fig. 3, which is composed of high-precision BDS positioning antenna and main module, controller, hydraulic valve, angle sensor, and so on. The vehicle automatic driving receives the correction data from the differential reference station through the communication module. The vehicle's Beidou positioning module and fusion algorithm are used to get the accurate location speed information and the three-axis attitude sensor is used to get three direction information of the vehicle itself. The path planning and navigation trajectory control of agricultural machinery are carried out on the vehicle tablet computer using the information obtained by steering wheel angle sensor. Using the special control and starter to control the electromagnetic steering valve, the automatic driving of the vehicle can be realized, and can complete the agricultural operation, such as ground turning, raking, longing, sowing and spraying, and so on.

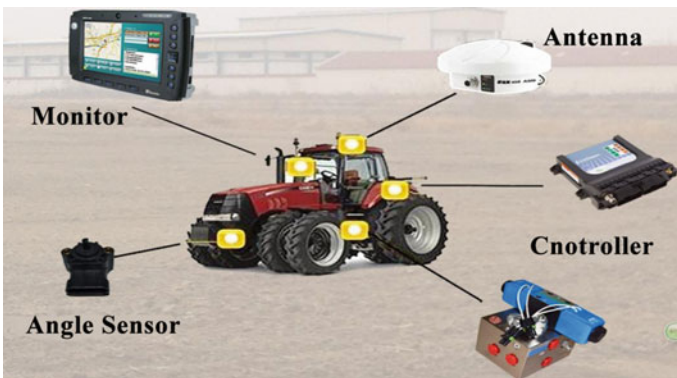


Fig. 3. The diagram of automatic driving system

3.3 Information Management System

The main words in all headings (even run-in headings) begin with a capital letter. Articles, conjunctions, and prepositions are the only words which should begin with a lower case letter. The information management system is composed of two parts, the information management center and the BDS positioning unit, as shown in Fig. 4. The BDS positioning unit is mainly equipped on staff, farm machinery, and grazing livestock to receive location data of the positioning satellite and calculate the location coordinates of the carrier (staff, farm machinery, and grazing livestock position information). Then the position information will be sent to a central server with fixed IP address through the GPRS communication module and saved in the center database. The information management unit connects with the center server through the Internet network, receives the data from the center server in real time, matches the coordinate information of the carriers with the electronic map, and displays the correct position of the carriers on the map, so that the user can clearly and intuitively grasp the dynamic position information of the related carriers and monitor them in real time. Combined with GIS technology, and using MapXtreme as the bottom development platform of map operation, the track replay, the location information display, and the monitoring of its state such as missing tracking, crossing the boundary alarm, and so on can be realized according to the location information transmitted by the positioning module. At the same time, LBS services such as command and control, information inquiry, and statistical analysis based on location information can be realized.

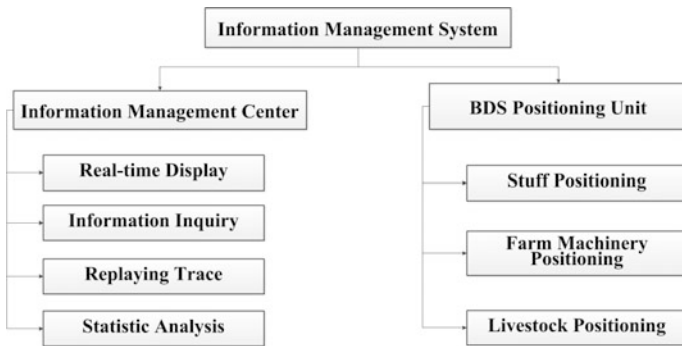


Fig. 4. The diagram of information management system

3.4 Data Management System

The data management system is composed of two parts: data management unit and system management unit, as shown in Fig. 5. Data management unit includes data receiving, parsing, storing, and processing functions. The data receiving program is written in C#. By listening to the port of the server, creating a control array, multiple connection requests can be accepted at the same time. The data sent through the GPRS is parsed in the database according to the custom data frame protocol. The system

management unit includes permission management and system maintenance function. Through the permission management, the user can access the data in the database through the data management system, and the database administrator is also able to maintain the database through the system.

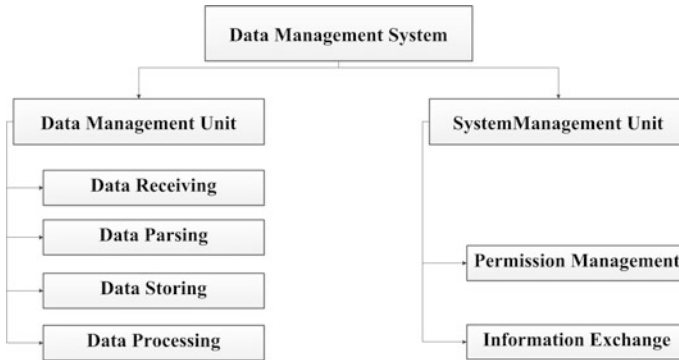


Fig. 5. The diagram of data management system

4 Conclusions

It will not adapt to the requirements of modern agriculture if the agricultural production stays on the traditional scale, production and management. Therefore, it is necessary to take the road of scale and modernization if agricultural wants to develop. By combining the BDS, GIS, mobile Internet (GPRS), and related automation control technology, communication technology, computer technology, and sensor technology, the system can provide the managers with comprehensive, timely and accurate service and various information inquiries, and help realize the intelligent monitoring and management of the farm. It is hoped to have certain guiding significance for future agriculture.

References

1. Song, C., Ding, R., Yue, P.: The discussion about the functions and applications while BeiDou navigation satellite system is applied to wisdom agriculture. *Comput. Knol. Technol.* **12**(10), 272–274 (2016)
2. Shi, G.: Application of GPS and GIS technology in monitoring system of precise agriculture. *Hubei Agric. Sci.* **50**(10), 1948–1950 (2011)
3. Yang, J.: Analysis of the application of BeiDou navigation system in agricultural field. *Sci. Technol. Innov.* **13**, 102–103 (2015)
4. Li, N., Xiaochen, F.: The application of GIS to the management of intelligent farm. *Mod. Agric.* **10**, 49–50 (2017)

5. Jiang, T., Zhao, J., Guo, R., et al.: Application of BeiDou navigation system in precision agriculture system. *Dig. Commun. World* **02**, 9–12 (2016)
6. Zhang, Y., Zhang, R., Jin, Y., et al.: Application of automatic driving system on agricultural tractor based on BeiDou navigation system. *Agric. Equip. Technol.* **42**(3), 8–11 (2016)



A Platoon-Based Vehicle-to-Vehicle Connectivity Enhancing Scheme Under Bidirectional Highway Scene

Kun Xu^(✉)

Chongqing Key Labs of Mobile Communications Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
xukun1993628@163.com

Abstract. The connectivity of vehicle-to-vehicle (V2V) communication is critical for efficient data sharing among vehicles. However, the vehicles moving in opposite direction show worse connectivity. To address this issue, a platoon-based V2V connectivity enhancing scheme under bidirectional highway scene is proposed in this paper. First, a typical bidirectional highway scene is investigated to ensure a comprehensive analysis of the connectivity probability among vehicles. Second, a time correlation expression of vehicle movement is derived based on the platoon-based expected lifetime model which is presented. Finally, compared with one popular dynamic clustering scheme in bidirectional road scenarios, the simulations validate the superior performance of the proposed scheme.

Keywords: Platoon-based scheme · Connectivity analysis · V2V communication · Bidirectional highway

1 Introduction

Currently, extensive studies have focused on the network connectivity of the Internet of Vehicles (IoVs) for realizing the related services, which are performed by message dissemination on the available links of vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I). Authors in [1] introduce generalized speed factor to analyze network connectivity in highway under various sub-scenes. In [2], the authors propose a probabilistic model for link duration in bidirectional highway and make some improvements in [3] under more complex channel environment. In [4], the authors regard vehicle mobility as Markov process for the time and spatial correlation. In [5], the connectivity also influenced by the variation of the safety distance is revealed. However, the cluster effects on the connectivity are always ignored.

There have been some literature [6–8] noticing that some consecutive vehicles with close space on the same direction can naturally be grouped into a platoon.

Kun Xu, M.S., Chongqing University of Posts and Telecommunications. His current main research interest includes Internet of Vehicles.

Platoon-based driving pattern in highway is regarded as a promising driving manner [8]. Vehicles in the same platoon share facilitate more efficient information dissemination for the higher steady moving state. And clustering movement of vehicles can be spontaneous or passive [9]. Broadly speaking, there are two types of scenarios for V2V communications [10]: vehicles traveling in opposite direction and same direction. Considering a vehicle driving along one direction will meet a great number of vehicles in the opposite direction with a short meeting time, and there would be still a good chance for information sharing if a proper V2V communication strategy is designed [11].

To address the above issues, this paper is dedicated to investigating a typical bidirectional road scenario in IoVs and exploring the connectivity performance among platoon-based vehicles driving in opposite directions. Specifically, the main contributions of this work are summarized as follows:

- (1) A platoon-based connectivity performance enhancing scheme in bidirectional highway for data sharing service is proposed, in which the connectivity probability and the expected link lifetime model of data service among platoon-based vehicles moving in the opposite direction are derived.
- (2) A comparison with one popular dynamic cluster scheme based on the proposed scheme is presented and verified.

The remainder of the paper is organized as follows: a platoon-based connectivity performance enhancing scheme under bidirectional highway scene are described in Sect. 2. In Sect. 3, the numerical results are obtained from simulations. Finally, Sect. 4 concludes this work and discusses future research direction.

2 Model

For any two vehicles i and j , if the distance is less than available transmission range R , they have probability to connect each other, which depends on the distance and vehicle density distribution on the road. Suppose the message is propagated from source to destination through V2V by multi-hop. If the distance between any vehicles i and j is longer than R , they could be disconnected, which is regarded as individual vehicle. Conversely, the vehicle i is grouped with j as a platoon. The front vehicle in each platoon is treated as header for message relaying and other members are connected to each other; hence, the communication range of one platoon is larger than individual vehicle. Sets R_1 and R_2 are the communication ranges for an individual vehicle and a platoon, respectively. And set Sd_1 , Sd_2 , and Sd_3 are the minimum safety distance for the inter-individual vehicle and the inter-platoon and the platoon member vehicles, respectively, which also satisfied $Sd_2 > Sd_1 > Sd_3$ as shown in Fig. 1.

2.1 Platoon-based V2V Connectivity Probability

In this part, the connectivity probability of vehicles inside one platoon and platoons in opposite direction under bidirectional highway is focused on. According

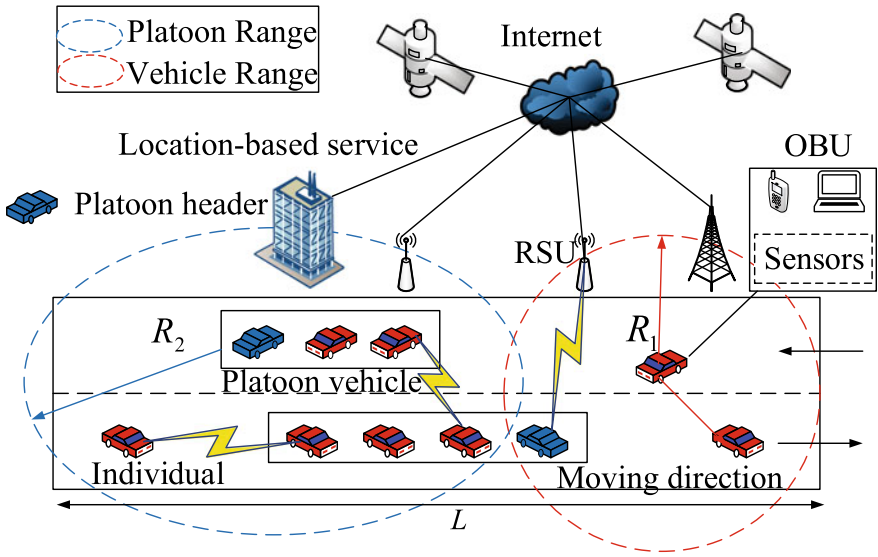


Fig. 1. Platoon-based V2V connectivity scenario

to traffic theory [12], let X represent the inter-vehicle distance between two consecutive vehicles, ρ is the vehicle density per unit road length. There is at least one vehicle in the interval with length x . The probability is given by

$$P_r \{X \leq x\} = h(x) = 1 - e^{-\rho x}. \tag{1}$$

Then, it can be found that X is independent identically distributed and obeys an exponential distribution. X_i ($i = 1, 2, \dots, N - 1$), let L express the length of road, where $N = L\rho$, and P_c be the connectivity probability of the IoVs which is given by

$$P_c = \prod_{i=1}^{N-1} P_r \{X_i \leq R\} = (1 - e^{-\rho R})^{N-1}. \tag{2}$$

Set γ as the ratio of the platoon in the road. So the connectivity probability between any two consecutive vehicles with minimum safety distance can be derived as

$$P_{all} = \left[\gamma \left(1 - e^{-\rho(R_2 - Sd_2)} \right) + (1 - \gamma) \left(1 - e^{-\rho(R_1 - Sd_1)} \right) \right]^{(N-1)}. \tag{3}$$

Set the neighbor nodes of platoon head as S_{Ni} , and all nodes are uniformly distributed on the road, the density of cluster in one way is $\rho_{platoon} = \frac{S_{Ni+1}}{2R_1}$. When the link between platoon member j and platoon leader i is connected, let P_m express the probability of platoon vehicle members as

$$P_m = \prod_{j=1}^{S_{Ni}} p \{0 \leq D_{ij} \leq R_1\} = \prod_{j=1}^{S_{Ni}} (1 - e^{-\rho R_1}) = \left(1 - e^{-\frac{(S_{Ni+1})}{2} R_1} \right)^{S_{Ni}}. \tag{4}$$

Considering the minimum safety distance Sd_3 between successive members of a platoon, the formula should be expressed as

$$P_m = \prod_{j=1}^{S_{Ni}} p \{0 \leq D_{ij} \leq (R_1 - Sd_3)\} = \left(1 - e^{-\frac{(S_{Ni}+1)(R_1 - Sd_3)}{2R_1}}\right)^{S_{Ni}}. \quad (5)$$

When the platoon-based vehicles are moving in the opposite direction road, the density of cluster becomes $\rho_{\text{platoon}} = \frac{S_{Ni}+1}{4R_1}$, because the communication range of platoon header and platoon members in one lane is $2R_1, R_2$, considering the header control of the group in bidirectional road with the maximum range as $4R_1$. Hence, let P_{platoon} express the probability of platoon vehicles in two lanes as

$$P_{\text{platoon}} = \prod_{j=1}^{S_{Ni}} (1 - e^{-\rho 2R_2}) = \left(1 - e^{-\frac{(S_{Ni}+1) \cdot R_2}{2R_1}}\right)^{S_{Ni}}. \quad (6)$$

2.2 Platoon-Based Link Lifetime Among Vehicles

In this part, according to the research on the relative motion between the vehicles in the same direction [10], the link will break when the distance between i and j is R_1 , marking the driving distance of vehicles at a time interval of t is $S(t)$, then the distance difference between i and j is computed as

$$S_j(t) - S_i(t) + X = R_1. \quad (7)$$

When a vehicle or a platoon driving is in the reverse two-lane highway, the reverse movement increases the communication range R from R_1 to $2R_1$, then

$$S_j(t) + S_i(t) - L_p = 2R_1. \quad (8)$$

Among them, L_p represents the platoon length. Before that, we have discussed the number of neighbor nodes of head vehicle i in a platoon is S_{Ni} , and L_p is expressed as

$$L_p = \sum_{i=1}^{S_{Ni}} d_{i(i+1)}. \quad (9)$$

The driving distance difference between i and j can be expressed as

$$S_j(t) + S_i(t) = \frac{1}{2}a_r t^2 + v_r t. \quad (10)$$

where $a_r = a_j + a_i, v_r = v_j + v_i$. Here, whether there is a platoon in bidirectional highway and platoon size is needed to be discussed. The first scene is that both two lanes are individual vehicles, and expected link lifetime can be computed as

$$t = \frac{-v_r + \sqrt{v_r^2 + 2a_r \cdot 2R_1}}{a_r}. \quad (11)$$

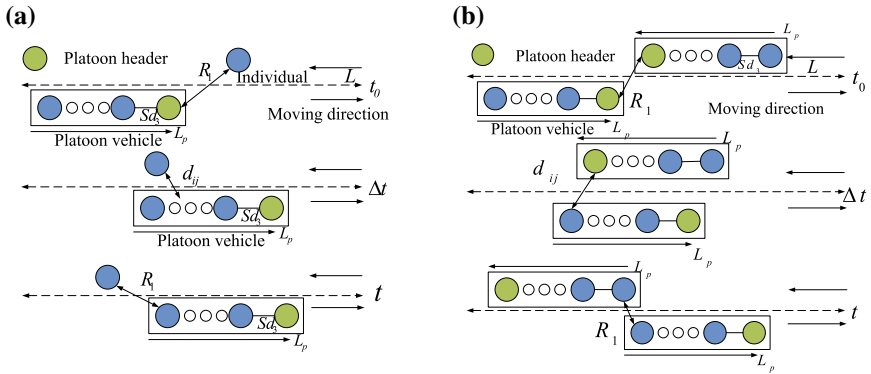


Fig. 2. a One platoon in single lane, **b** two platoons in two lanes

For the second scenario, the platoon exists in one lane or both, and link lifetime can be computed as follows and is shown in Fig. 2.

$$t = \frac{-v_r + \sqrt{v_r^2 + 2a_r(2R_1 + L_p)}}{a_r} \tag{12}$$

3 Experiment

In this section, the proposed platoon-based connectivity performance enhancing scheme in bidirectional highway is numerically simulated. The impacts of the minimum safety distance Sd_1 , Sd_2 , and Sd_3 , the vehicle density ρ , the communication ranges R_1 and R_2 , and the platoon ratio γ , the neighbor nodes of platoon head S_{Ni} , the relative velocity v_r and acceleration a_r , and platoon size L_p are analyzed. The length of the road is set as $L = 10$ km.

3.1 V2V Connectivity Probability in Platoons

Figure 3a shows that the road connectivity probability is enhanced when there are platoons, and the variation of connectivity probability is proportional to the platoon ratio γ , in which $\gamma = 0$ expresses no platoon on the road, and the performances of all cases are almost the same due to the connectivity probabilities of both approaches to 1 under the density of 0.06 veh/km. Then, Fig. 3b shows two vehicle densities ($\rho = 0.02, 0.04$) by whether considering the minimum safety distances are compared to present the whole road connectivity probability under different platoon ratio γ . It is found that the probability is decreased and is influenced by the safety distance.

Figure 4 shows V2V connectivity probability in platoons, consider the minimum safety distance Sd_3 is at least 30 m, and the platoons in opposite direction road always have higher connectivity probability than one platoon in any single lane. In addition, when the number of neighbor vehicle nodes increases,

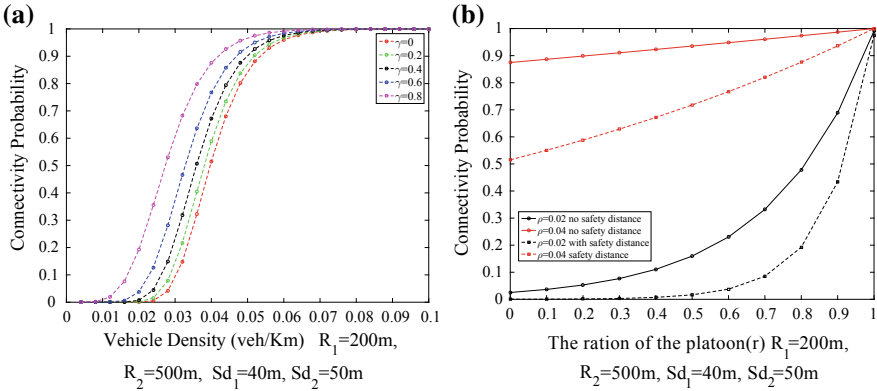


Fig. 3. **a** Platoon-based road connectivity probability, **b** minimum safety distance added in

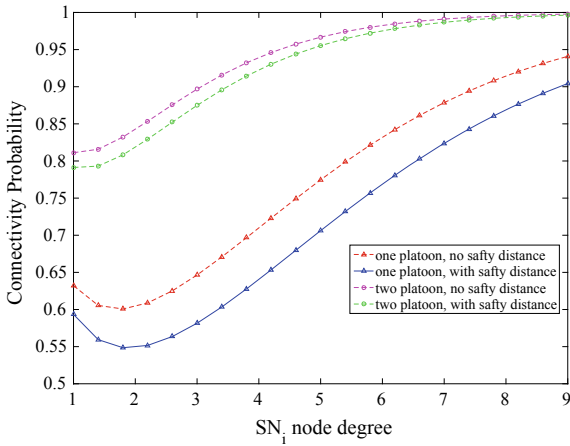


Fig. 4. V2V connectivity probability in platoons

the change in the amplitude of the connectivity probability is decreasing and is approaching to 1. And the platoons in opposite direction road always have higher connectivity probability, which may be due to the increase in the number of neighbor nodes which leads to a decrease in dynamic variable space between vehicles within the communication range. In addition, when considering the minimum safety distance, the connectivity probability is decreased.

3.2 Platoon-Based Link Lifetime Among Vehicles

Figure 5a shows four critical situations in a bidirectional highway scene which is simulated based on the V2V connectivity probability in platoons with the relative acceleration as $a_r = 1m/s^2$ is set: (1) no platoon in two lanes, (2) both smallest platoons in two lanes, (3) one biggest platoon in single lane, and (4)

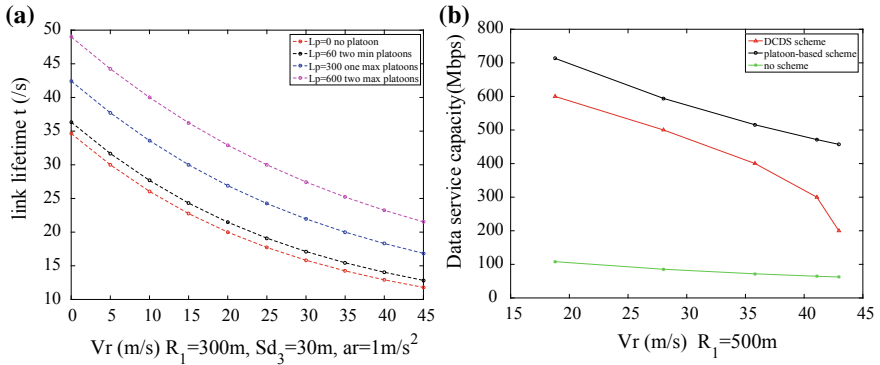


Fig. 5. **a** Platoon-based link lifetime among vehicles, **b** comparative analysis between the DCDS and platoon-based scheme

both biggest platoon in two lanes. And other results can be extended from the four basic situations. It can be seen intuitively that although the link lifetime will gradually become smaller in all cases with the increase of relative velocity and acceleration, the presence of platoons on both sides of the road can significantly improve the link lifetime among vehicles.

Figure 5b shows the comparative analysis between the dynamic clustering based data sharing algorithm (DCDS) and proposed scheme based on the analysis for the expected link lifetime, according to the parameters setting of DCDS, in which the data service capacity is one main parameter used to evaluate the connectivity performance among cluster-based vehicles in bidirectional highway with five traffic scenarios under different relative speeds computed by unit conversion as $v_r = [18.8, 28.1, 35.8, 41.1, 42.9]$ m/s, the communication range R_1 is set as 500 m. Considering the DSRC, it supports the data transmission rate between 3 and 12 Mbps for 60 and 120 km/hr vehicle speed [11]. Hence, the maximum data service capacity among the five relative speeds is presented in the figure. It is clear that the maximum data service capacity is enhanced when the platoon-based scheme is considered under the same condition. In addition, it is found that in the case of lower average relative velocity, both show little difference in data service performance. However, as the average relative speed increases, the proposed scheme shows better robustness.

4 Conclusions

In this paper, the platoon-based V2V connectivity enhancing scheme for data sharing under bidirectional highway scene has been presented. Based on the analysis of platoons in V2V communication, the connectivity probability and expected link lifetime model are derived in several sub-cases first. Then, one popular dynamic cluster scheme has been used to compare with the proposed scheme in data service capacity. Simulation results indicate that the proposed scheme has shown better performance in data service, especially in the case of

higher relative speed. From another point of view, it is verified that connectivity among vehicles moving in the opposite direction can be enhanced by the effect of the proposed scheme. In the future work, the expected link lifetime model based on the proposed framework with more complex movement conditions and impact factors such as wireless channel access information, transmission capacity will be considered. In addition, more reliable routing algorithm to display more complex changes in the group movement of vehicle will be applied. It is to be noted that this research is beneficial to further routing design, content distribution, and real-time information sharing in IoVs.

Acknowledgments. This work is supported by Natural Science Foundation of China (Grant No. 61702066), Scientific and Technological Research Program of Chongqing Municipal Education Commission (Grant No. KJ1704080), and Chongqing Research Program of Basic Research and Frontier Technology (Grant No. cstc2017jcyjAX0256).

References

1. Khan, Z., Fan, P., Fang, S.: On the connectivity of vehicular ad hoc network under various mobility scenarios. *IEEE Access* 22559–22565 (2017)
2. Shelly, S., Babu, A.V.: A probabilistic model for link duration in vehicular ad hoc networks under Rayleigh fading channel conditions. In: Fifth International Conference on Advances in Computing and Communications, pp. 177–182 (2016)
3. Shelly, S., Babu, A.V.: Performance modeling of link duration in vehicular ad hoc networks under Weibull fading channel conditions. *Wirel. Pers. Commun.* **96**(4), 6047–6068 (2017)
4. Hu, M., Zhong, Z., Ni, M.: Mobility aware link lifetime analysis for vehicular networks. In: Wireless Communications and Networking Conference, pp. 1853–1858. IEEE (2015)
5. Chunxiao, L., Anran, Z., Jun, S., Meixiang, Z., Xuelong, H.: Analysis of Connectivity Probability in VANETs Considering Minimum Safety Distance. Accepted by: WCSP, Yangzhou, China, p. 2016 (2016)
6. Caixing, S., Supeng, L., Yan, Z.: Performance analysis of connectivity probability and connectivity-aware MAC protocol design for platoon-based VANETs. *IEEE Trans. Veh. Technol.* **64**(12), 5596–5609 (2015)
7. Jia, D., Lu, K., Wang, J.: On the network connectivity of platoon-based vehicular cyber-physical systems. *Transp. Res. Part C* **40**(1), 215–230 (2014)
8. Caixing, S., Supeng, L., Yan, Z.: Research of the multi-way connectivity probability for platoon-based vehicle-to-infrastructure communication network. *J. China Univ. Posts Telecommun.* **23**(1), 1–7 (2016)
9. Gongan, Q., Zhihua, B., Guoan, Z.: Connectivity analysis of passive cluster with high stability in vehicular wireless network. *J. Commun.* (2016)
10. Gongjun, Y., Rawat, D.B.: Vehicle-to-vehicle connectivity analysis for vehicular ad-hoc networks. *Ad Hoc Netw.* (2016)
11. Junhua, W., Kai, L., Ke, X., Sanghyuk, S.: Dynamic clustering and cooperative scheduling for vehicle-to-vehicle communication in bidirectional road scenarios. *IEEE Trans. Intell. Transp. Syst.* (2017)
12. Junhui, Z., Yan, C., Gong, Y., Yan, C.: Study of connectivity probability of vehicle-to-vehicle and vehicle-to-infrastructure communication systems. In: IEEE, Vehicular Technology Conference, pp. 1–4. IEEE (2016)



An Intelligent Water Regimen Monitoring System

Yaping Fan, Heng Dong^(✉), Ying Jiang, Jinqiu Pan, Shangang Fan,
and Guan Gui

Key Lab of Broadband, Wireless Communication and Sensor Network
Technology, Ministry of Education, Nanjing University of Posts and
Telecommunications, Nanjing 210003, China
dongh@njupt.edu.cn

Abstract. Intelligent water regimen monitoring system is required to design for protecting people's safety and property in real time. It is bound to unstoppable that the traditional water level detecting system using manpower is replaced by automatic, intelligent monitoring system. This paper designs an intelligent water regimen monitoring system, which contains remote monitor stations and monitoring center. The remote monitoring station contains an embedded data acquisition module in order to collect environmental data including water level and image periodically. The collected data is then sent to a server over the Internet via cellular network, which is stored on a database and processed to determine early warning of flood in the area based on historical data. The design system not only solves measuring accuracy problem, but also increases working efficiency.

Keywords: Water regimen remote monitoring · Embedded device · Real-time monitoring

1 Introduction

In China, as a large agricultural country, the construction of water conservancy entailing is not only vital to flood control project construction, water supply security, but also is important to the construction of ecological civilization and economic civilization. In December 2016, the General Office of the Central Committee of the Communist Party of China and the State Council have jointly issued a document to implement a “River chief” system nationwide, which is an institutional innovation for improving integrated water governance and guaranteeing national water security. Hence, using information technology as a support has become an inevitable trend to solve management issues of rivers. Machine vision is a newly developed and widely applied subject in recent years, which is expected to realize object tracking, measurement, and recognition by using cameras and computers instead of human beings.

When machine vision is applied to hydrological monitoring, it has developed into a novel noncontact method of water level measurement [1, 2]. Compared with the traditional networked sensor measurement method, the preparatory time is greatly shortened, and more manpower and financial resources are saved in equipment

installation and maintenance. Compared to the automatic technology of water detection in foreign countries, China is still in its infancy and lack of mature products [3]. However, with the continuous upgrading of digital hardware and continuous research of digital image processing technology, such products will surely become the most reliable and widely used measuring methods.

In this paper, we design an intelligent water regimen monitoring system and introduce the design details. Based on advanced techniques such as geographic information system, wireless communications, computer technology, and embedded technology, the proposed system is provided with function of real-time data acquisition, graphic collection, history query, and hydrographic analysis. This system is also capable to realize the functions of distant monitoring and storing of hydrologic data, as well as the function of database storage analysis after taking the place of the former paper-based recording by hydrologic observatories. The experimental results show that the system can work stably, with high accuracy, low cost, and easy to market.

2 System Structure

The intelligent monitoring system generally includes three parts: hardware and software of remote measuring station, hardware and software of the data base station, and software for the remote monitoring center. Remote measuring stations are responsible for calculating water level with algorithms using the captured continuous image from the video camera in real time. The results of computation are then sent to a server over the Internet via a cellular data network, which are stored on a database and processed to determine early warning of flood in the area based on historical data. Figure 1 shows the architectural design of the system.

3 Remote Measuring Station

The remote measuring station is responsible for collecting videos, images, and water levels, which have four blocks: camera, embedded data acquisition device, 4G router, and solar power supply equipment [4].

3.1 Water Level Measurement

The video stream collected by the camera is sent to the embedded data acquisition device for immediate processing, and the measurement results are obtained using a dictionary learning algorithm. The water level data is then transmitted to the server at a certain frequency by HTTP communication (Fig. 2).

3.2 Image Identification

We obtain the water level by using dictionary model [5]. A water gage is set up in the river to be measured and we will collect the image of the water gage near the water surface through a camera. By taking images of different scenes as training samples, we

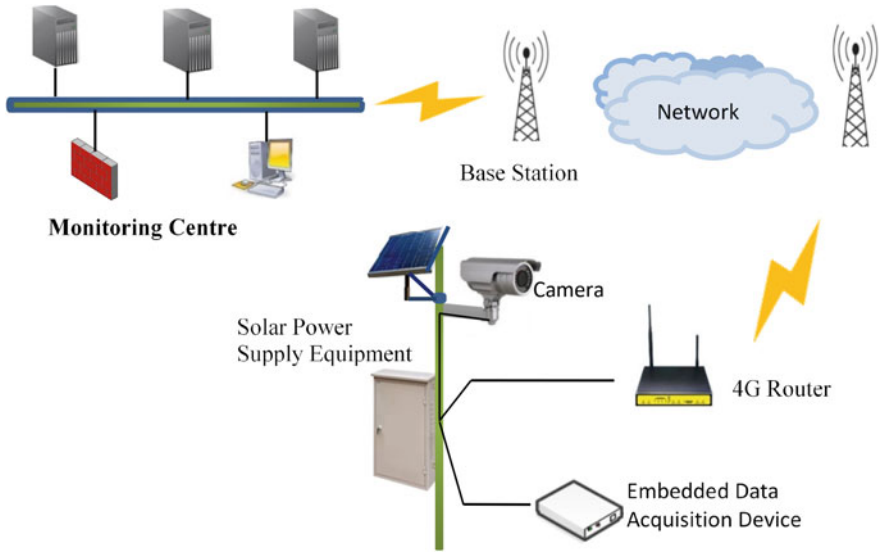


Fig. 1. Architectural design of the intelligent water regimen monitoring system.

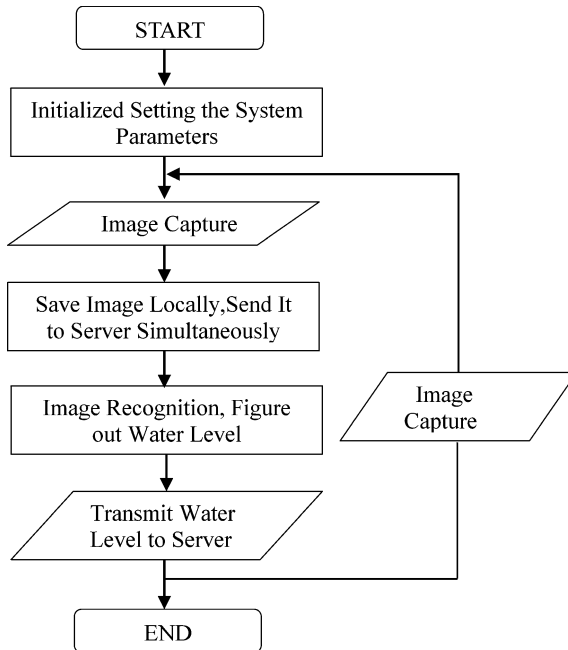


Fig. 2. Flowchart of measuring water level.

will obtain different training dictionaries. The next step is computing water level according to the dictionary structure and the learning algorithm (Fig. 3).

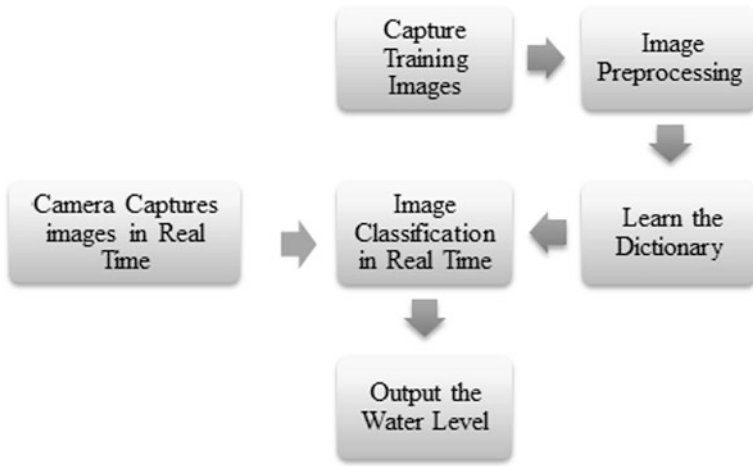


Fig. 3. Flowchart of proposed algorithm.

3.3 Power Consumption Estimation of the Water Level Monitoring Station

Since the remote measuring stations are usually far from town, the system is designed to be powered using a solar power supply equipment with a backup battery that helps the monitoring station to work at least 7 days without recharging [6].

4 Monitoring Center

Based on the digitalization management platform and computer, the monitoring center includes the database and monitoring software. The monitoring software takes responsibility of dealing with the hydrographic information from locale remote measurement terminals and handling how to remotely monitor, providing a complete monitoring interface to carry out historical data queries, display real-time data, data collection, data storage, data analysis, and alarming for non-normal status. The system would realize distributed measuring, centralized monitoring, and management. Figure 4 shows the functional architecture design of the monitoring center.

4.1 GIS

Based on GIS strong spatial data structure to interrelate kinds of hydro-geophysical exploration data and spatial data, the system would show the geographical distribution, operation status, and hydrological data of each station in real time, which provides users with an intuitive and concise way of viewing information (Fig. 5).

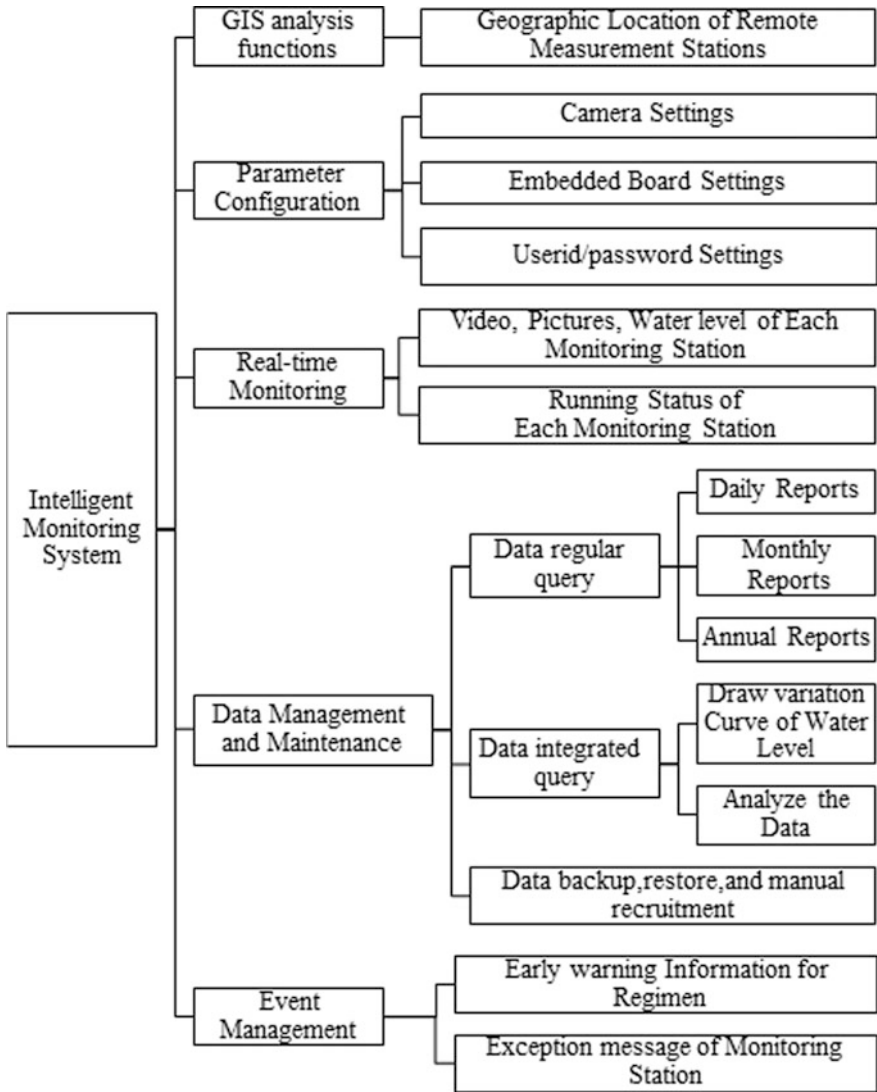


Fig. 4. Functional architecture of the monitoring center.

4.2 Parameters Configuration

Remote measuring station requires the users to first set the required parameters to the server, and then initialize the camera and embedded board. There are three ways to set needed parameters which include reading parameters from the configuration file, getting parameters via a web interface, and getting parameters via client application software.

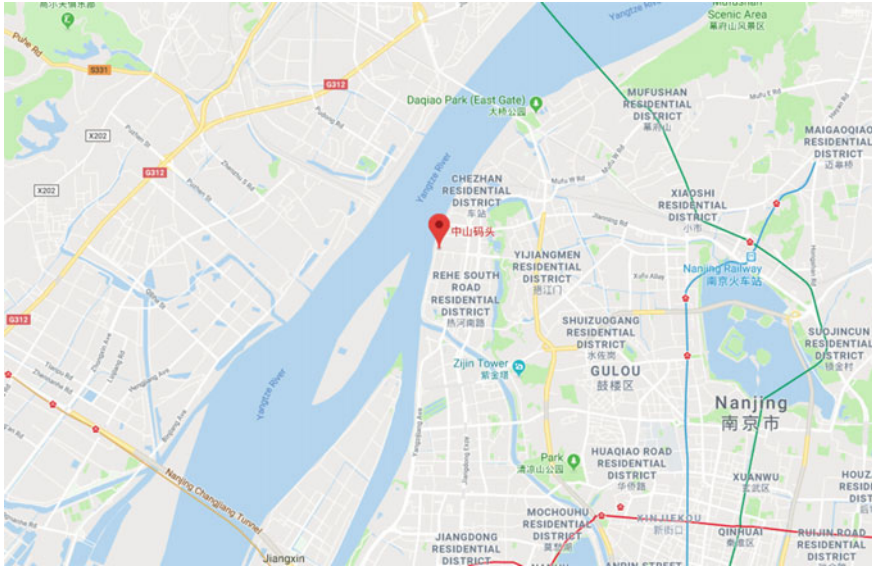


Fig. 5. Web application shows map with overlay flood information.

4.3 Data Management and Maintenance

The system supports functions of processing data and archiving data automatically in real time. The data transmitted from the telemetry stations is statistically classified according to time or geographical nodes, and periodic reports are generated according to the needs of the water level monitoring, so as to report, keep an archive, and reorganize hydrological data.

Under normal circumstances, the intelligent system records the water level data every 5 min, and the data storage validity period is 5 years. After selecting the station and time period, you can select the elements you want to compare, query, and display them as a chart.

In order to guarantee the reliability and the recoverability, the technologies of data backup online and system recovery are adopted. The backup date including measuring instruments used in every remote measurement stations, installation locations, measurement conversion algorithms and parameters, output template settings, and so on. After completing the algorithm and parameter settings of a batch of measurement stations, a system information backup is immediately performed. This backup helps to automatically recover the system in the future. If the monitoring project fails, manual observation data can be entered artificially.

4.4 Event Management

According to the processing of data information uploaded by the device, the water level of the detection station is alerted and alarmed in real time. The system displays the pictures of the alarming site and the detailed contents of the warning, so that it can release the flood or drought warning of water within the scope of influence timely.

The telemetry station has a self-check function, which can perform self-check on the data memory, CPU, clock, power supply status, battery voltage, and measurement circuit in the measurement and control device. When something goes wrong, the monitoring host can display exception information for timely maintenance.

5 Conclusion

The designed system applies low-power consumption, and having a sampling period which is usually over 5 min, the data can be uploaded to the monitoring center and backed up on the local node simultaneously. It supports real-time remote access control so as to initiate video transmission as needed, providing intuitive live image information for the flood-fighting command. Using the system, testing accuracy and time-effectiveness of hydrological monitoring data would be greatly improved. With the characteristic of running stabilization, convenient maintenance, and low cost, it is easy to expand and popularize, which can be used on the basis of existing hydrological monitoring stations. It supports application scenarios such as river channels, lakes, reservoirs, and urban culverts. By linking the water level data of various rivers, the water level of key rivers can be predicted and the floods can be alerted.

References

1. Chen, J.: Method of water level data capturing based on video image recognition. *Water Resour. Inform.* (2013)
2. Kim, J., Han, Y., Hahn, H.: Embedded implementation of image-based water-level measurement system. *IET Comput. Vis.* **5**(2), 125–133 (2011)
3. Dersingh, A.: Design and development of a flood warning system via mobile and computer networks. In: 2016 International Conference on Electronics, Information, and Communications, pp. 1–4 (2016)
4. Thekkil, T.M., Prabakaran, N.: Real-time WSN based early flood detection and control monitoring system. In: 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), pp. 1709–1713 (2017)
5. Wang, X., Gu, Y.: Cross-label suppression: a discriminative and fast dictionary learning with group regularization. *IEEE Trans. Image Process.* **26**(8), 3859–3873 (2017)
6. Visconti, P., Orlando, C., Primiceri, P.: Solar powered WSN for monitoring environment and soil parameters by specific app for mobile devices usable for early flood prediction or water savings. In: 2016 International Conference on Environment and Electrical Engineering (EEEIC), pp. 1–6 (2016)



Attitude Stabilization Control Method for Quadrotor UAV Based on ADRC

Sen Yang^{1,2}(✉), Leiping Xi¹, Guanghong Gong², and Hairui Dong¹

¹ Department of UAV Engineering, Army Engineering University, Shijiazhuang, China
568657132@qq.com

² School of Automation Science and Electrical Engineering, Beihang University, Beijing, China

Abstract. In order to solve the problems in attitude stabilization control of quadrotor UAV, an attitude stabilization controller based on Active Disturbance Rejection Control (ADRC) is proposed. First, the dynamics model of quadrotor UAV is introduced, and then the attitude stabilization controller based on ADRC is designed for quadrotor pitch, roll and yaw channels, respectively. The simulation results show that the designed controller can meet the control precision and rapidity requirements of the quadrotor.

Keywords: Attitude stabilization control · Quadrotor · ADRC

1 Introduction

The essence of PID control is to build a control strategy based on the error between the target value and the actual value, so as to eliminate the error. This advantage makes PID control to be widely used in control engineering [1, 2]. However, with the development of technology and the continuous improvement of the control accuracy, speed, and adaptability to the environment in the industry, PID control is difficult to meet the requirements. For this reason, people are discussing the solution, which can mainly be summarized as two major approaches [3]. The first one uses a combination of modern control technology methods and PID control, using modern control theory to adjust PID control parameters to improve the performance and adaptability of PID control, such as adaptive PID, neural network PID, fuzzy PID, etc. The second is to analyze the advantages and disadvantages of PID control from a more in-depth perspective and to conduct a comprehensive improvement on its shortcomings. This is the auto-disturbance rejection control technology. The auto-disturbance rejection control technology is a new and practical technology that inherits and develops the ideology of PID control, absorbs modern control theory results, and develops and utilizes nonlinear effects [4–6].

In this paper, aiming at the susceptibility to interference during the flight process of quadrotor UAVs, an attitude control method based on ADRC is proposed, which can estimate and compensate the interferences. The feasibility and effectiveness of the proposed method are illustrated by simulations.

2 Quadrotor UAV Mathematical Model

The motion of the drone is a spatial six degrees of freedom motion, including 3 degrees of freedom of linear motion and 3 degrees of freedom around the centroid of angular motion. According to Newton’s second law, the aircraft dynamics equation is [7, 8] given as follows:

$$\vec{F} = m \frac{d\vec{V}}{dt} \tag{1}$$

$$\vec{M} = \frac{d\vec{H}}{dt} \tag{2}$$

In the formula, \vec{F} is the external force acting on the quadrotor drone and m is the quality of the drone. \vec{V} is the speed of the movement of the center of gravity of the drone. \vec{M} is the external torque acting on the drone and \vec{H} is the absolute angular momentum of the drone relative to the ground coordinate system.

- (1) Linear Motion Equations Four-rotor UAVs are shown in Figs. 2 and 3. Forces on the UAV are gravity, rotor lift, and air resistance.

$$G = mg \tag{3}$$

$$F_i = \frac{1}{2} \rho C_l \omega_i^2 = k_l \omega_i^2 \tag{4}$$

$$D_i = \frac{1}{2} \rho C_d \omega_i^2 = k_d \omega_i^2 \tag{5}$$

In the formula, the gravity of the drone is the lift G of the i -th rotor, the resistance of the i -th rotor, the lift F_i coefficient of the rotor, the drag coefficient of the rotor, and the angular velocity of the i -th rotor. The lift k_l coefficient is the drag coefficient, where $i = 1, 2, 3, 4$.

The total lift \vec{F} generated by the four rotors is given below:

$$\vec{F} = R \begin{bmatrix} 0 \\ 0 \\ \sum_{i=1}^4 F_i \end{bmatrix} = \begin{bmatrix} \cos \psi \sin \theta \cos \phi + \sin \psi \sin \phi \\ \sin \psi \sin \theta \cos \phi - \sin \phi \cos \psi \\ \cos \theta \cos \phi \end{bmatrix} \sum_{i=1}^4 F_i \tag{6}$$

Substituting Eq. (6) into Eq. (1) yields the following:

$$\begin{cases} \ddot{x} = \left[(\cos \psi \sin \theta \cos \phi + \sin \psi \sin \phi) \sum_{i=1}^4 F_i - K_1 \dot{x} \right] m^{-1} \\ \ddot{y} = \left[(\sin \psi \sin \theta \cos \phi - \sin \phi \cos \psi) \sum_{i=1}^4 F_i - K_2 \dot{y} \right] m^{-1} \\ \ddot{z} = \left[(\cos \theta \cos \phi) \sum_{i=1}^4 F_i - K_3 \dot{z} \right] m^{-1} - g \end{cases} \quad (7)$$

In the formula, (x, y, z) is the centroid position of the drone, $K_i (i = 1, 2, 3)$ is the total drag coefficient, and g is the gravitational acceleration.

(2) The equation of angular motion, the relationship between Euler angle angular velocity $(\dot{\phi}, \dot{\theta}, \dot{\psi})$ and body angular velocity (p, q, r) is given below:

$$\begin{bmatrix} p \\ q \\ r \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\sin \theta \\ 0 & \cos \phi & \sin \phi \cos \theta \\ 0 & -\sin \phi & \cos \phi \cos \theta \end{bmatrix} \begin{bmatrix} \dot{\theta} \\ \dot{\phi} \\ \dot{\psi} \end{bmatrix} \quad (8)$$

By type (8), it is available as

$$\begin{bmatrix} \dot{\theta} \\ \dot{\phi} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} p + p \sin \phi \tan \theta + r \cos \phi \tan \theta \\ q \cos \phi - r \sin \phi \\ q \sin \phi \sec \theta + r \cos \phi \sec \theta \end{bmatrix} \quad (9)$$

Due to the above assumptions, the quadrotor UAV is symmetrical in shape and uniform in mass distribution, and it can be assumed that the center of mass of the aircraft is located in the center of the quadrotor UAV, and its inertia matrix I is a diagonal matrix [9] as given below:

$$I = \begin{bmatrix} I_x & 0 & 0 \\ 0 & I_y & 0 \\ 0 & 0 & I_z \end{bmatrix} \quad (10)$$

In the formula, they are the rotational inertia X, Y, Z of the four-rotor UAV.

The equation of angular motion of the quadrotor UAV can be obtained from the rigid body rotation theorem, which is as follows:

$$M = I \begin{bmatrix} \dot{p} \\ \dot{q} \\ \dot{r} \end{bmatrix} \quad (11)$$

According to the calculation method of angular momentum, the quadrotor UAV angular motion equation is given as

$$\begin{bmatrix} M_x \\ M_y \\ M_z \end{bmatrix} = \begin{bmatrix} \dot{p}I_x - \dot{r}I_{xz} + qr(I_z - I_y) - pqI_{xz} \\ \dot{q}I_y - pr(I_x - I_z) + (p^2 - r^2)I_{xz} \\ \dot{r}I_z - \dot{p}I_{xz} + pq(I_y - I_x) + qrI_{xz} \end{bmatrix} \tag{12}$$

In the formula, M_x, M_y, M_z is the component of the quadrotor UAV subjected to the external moment around the shaft.

By type (12), it is available as

$$\begin{bmatrix} \dot{p} \\ \dot{q} \\ \dot{r} \end{bmatrix} = \begin{bmatrix} [M_x + (I_x - I_z)qr]/I_x \\ [M_y + (I_z - I_x)pr]/I_y \\ [M_z + (I_x - I_y)pr]/I_z \end{bmatrix} \tag{13}$$

By finishing formula (12) and formula (13), we get

$$\begin{cases} \ddot{\phi} = [M_x - \dot{\theta}\dot{\psi}(I_z - I_y)]/I_x \\ \ddot{\theta} = [M_y - \dot{\phi}\dot{\psi}(I_x - I_z)]/I_y \\ \ddot{\psi} = [M_z - \dot{\phi}\dot{\theta}(I_y - I_x)]/I_z \end{cases} \tag{14}$$

According to the ‘‘X’’ flying method of the quadrotor UAV, the momentum theorem can be obtained from the following:

$$\begin{bmatrix} M_x \\ M_y \\ M_z \end{bmatrix} = \begin{bmatrix} l(F_1 + F_4 - F_2 - F_3) \\ l(F_1 + F_3 - F_2 - F_4) \\ \lambda(F_1 + F_2 - F_3 - F_4) \end{bmatrix} \tag{15}$$

In the formula, l is the distance from the motor shaft to the center of mass of the quadrotor UAV (arm length), and λ is the coefficient between the lift and torque generated by the rotation of the motor.

The system input for defining the quadrotor UAV is given below:

$$U = \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \end{bmatrix} = \begin{bmatrix} F_1 + F_2 + F_3 + F_4 \\ F_1 + F_4 - F_2 - F_3 \\ F_1 + F_3 - F_2 - F_4 \\ F_1 + F_2 - F_3 - F_4 \end{bmatrix} \tag{16}$$

In the formula, U_1 is the amount of vertical lift control, U_2 is the amount of roll path control, and U_3 is the amount of pitch path control, and U_4 is the yaw path control amount.

Comprehensively (14), (15), and (16) are given as

$$\begin{cases} \ddot{\phi} = U_2 I_x^{-1} \\ \ddot{\theta} = U_3 I_y^{-1} \\ \ddot{\psi} = U_4 \lambda I_z^{-1} \end{cases} \quad (17)$$

Considering that the currently set four-rotor flight environment is a small indoor or outdoor wind speed, the resistance factor $K_i (i = 1, 2, 3)$ in Eq. (7) is ignored. Thus, the following mathematical model can be obtained by combining Eqs. (7) and (17) as

$$\begin{cases} \ddot{x} = [(\cos \psi \sin \theta \cos \phi + \sin \psi \sin \phi) U_1] m^{-1} \\ \ddot{y} = [(\sin \psi \sin \theta \cos \phi - \sin \phi \cos \psi) U_1] m^{-1} \\ \ddot{z} = [(\cos \theta \cos \phi) U_1] m^{-1} - g \\ \ddot{\phi} = U_2 I_x^{-1} \\ \ddot{\theta} = U_3 I_y^{-1} \\ \ddot{\psi} = U_4 \lambda I_z^{-1} \end{cases} \quad (18)$$

3 ADRC-Based Attitude Stabilization Control for Quadrotor UAV

The ADRC technology mainly includes three parts: arranging the transition process, Extended State Observer (ESO) and Nonlinear State Error Feedback (NLSEF). Among them, the Tracking Differentiator (TD) is often used to arrange the transition process. The purpose is to reduce the initial error of the system, solve the problem that the rapidity and overshoot must be taken into consideration, and extract the differential signal of the input signal. An extended state observer can estimate the state of the system and the “total disturbance” of the internal disturbances and external disturbances in the system, and compensate for “total disturbances”. The nonlinear state error feedback control law can compensate the suppression residual and improve the controller function [10, 11].

Taking quadruple rotor pitch channel as an example to study the attitude control of drone based on ADRC. Because the transfer function of the pitch channel of the quadrotor UAV is third order, a linear ADRC control strategy is adopted for the convenience of calculation, and its structure is shown in Fig. 1.

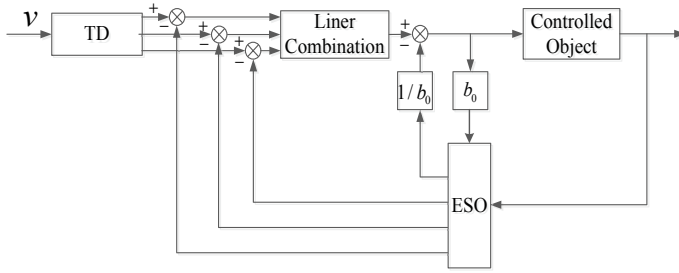


Fig. 1. Linear ADRC control structure

$$\begin{cases} e = v_1 - v \\ v_1 = v_1 + hv_2 \\ v_2 = v_2 + hv_3 \\ v_3 = v_3 + h\{-r[r(r(e) + 3(v_2)) + 3(v_3)]\} \end{cases} \quad (19)$$

The third-order extended state observer is given as

$$\begin{cases} e = z_1 - y \\ fe = \text{fal}(e, 0.5, \delta) \\ fe_1 = \text{fal}(e, 0.25, \delta) \\ fe_2 = \text{fal}(e, 0.125, \delta) \\ z_1 = z_1 + h(z_2 - \beta_{01}e) \\ z_2 = z_2 + h(z_3 - \beta_{02}fe) \\ z_3 = z_3 + h(z_4 - \beta_{03}fe_1 + bu) \\ z_4 = z_4 + h(-\beta_{04}fe_2) \end{cases} \quad (20)$$

The linear combination of status feedback is given as

$$u = \beta_1 e_1 + \beta_2 e_2 + \beta_3 e_3 - z_4 \quad (21)$$

The parameters of linear ADRC are $r = 10$, $h = 0.01$, $\beta_{01} = 24$, $\beta_{02} = 430$, $\beta_{03} = 8720$, $\beta_{04} = 2,00,000$, $\beta_1 = 10$, $\beta_2 = 200$, and $\beta_3 = 2$. Convert the transfer function of the pitch channel to the controllable standard type as given below:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 \\ \dot{x}_3 = x_4 \\ x_4 = -11.4943x_3 + 1017.8u \\ y = x_1 \end{cases} \quad (22)$$

Linear ADRC control response results are shown in Figs. 2 and 3.

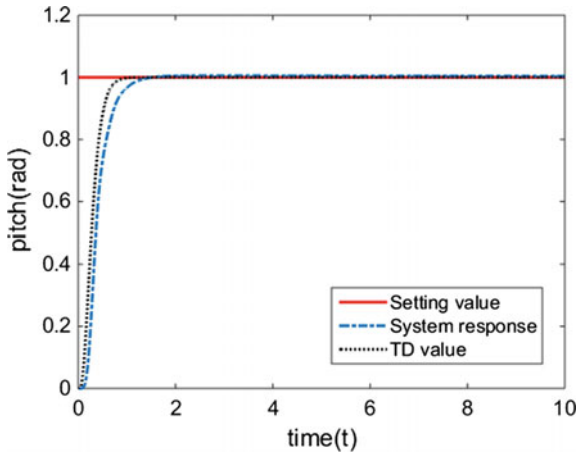


Fig. 2. ADRC control response of pitch channel

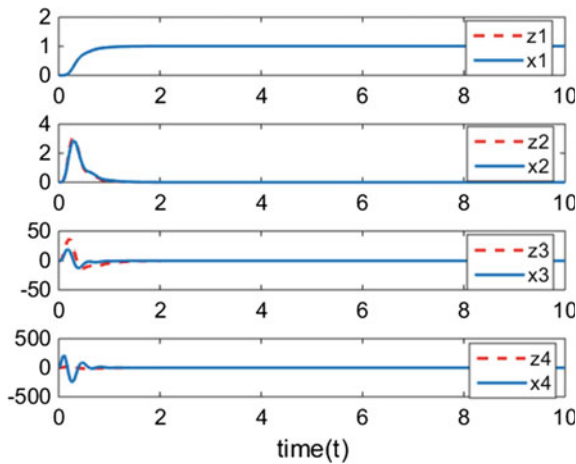


Fig. 3. Expanded state observer

It can be seen from Fig. 2 that the linear ADRC's control response to the pitch channel is fast and there is no overshoot. Expanded state observer versus system state vector is shown in Fig. 3. x_1 , x_2 , x_3 all have good estimates. There is some error in estimating x_4 expansion status.

The control effect of the linear ADRC on the roll channel and the yaw channel is shown in Figs. 4 and 5. The parameter selection is the same as that of the pitch channel.

As can be seen from Figs. 4 and 5, the linear ADRC has a good control effect on the roll channel and the yaw channel when the parameters are unchanged.

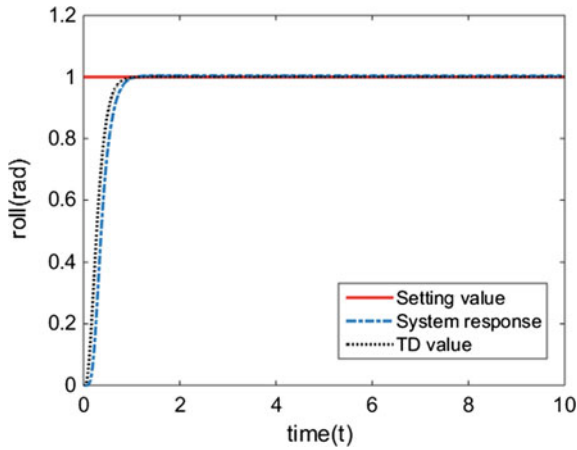


Fig. 4. ADRC control response results for roll channel

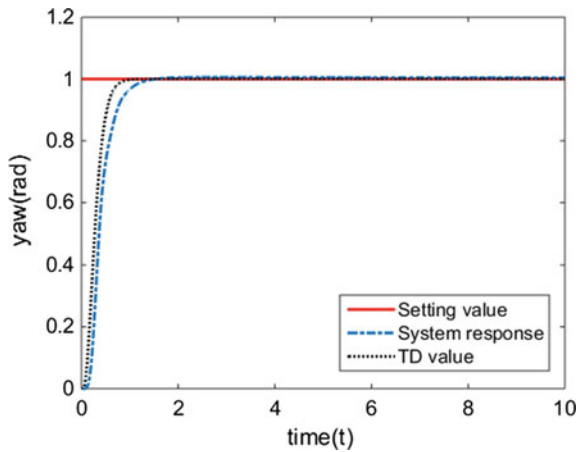


Fig. 5. ADRC control response results for yaw channel

4 Conclusion

In this paper, based on the analysis of the mathematical model of the quadrotor UAV, a four-rotor UAV attitude stabilization controller based on ADRC is designed. The simulation results show that the ADRC controller has good control effects on all three attitude channels such as pitch, roll, and yaw.

References

1. Altug, E., Ostrowski, J., Mahony, R.: Control of a quadrotor helicopter using visual feedback. In: IEEE International Conference on Robotics & Automation, pp. 72–77, May 2002
2. Mohta, K., Kumar, V., Daniilidis, K.: Vision-based control of a quadrotor for perching on lines. In: IEEE International Conference on Robotics & Automation, pp. 3130–3136 (2014)
3. Jingqing, H.: Active Disturbance Rejection Control Technique. National Defense industry Press, Beijing (2013)
4. Jing, Z., Hua, Z., Heng, L., et al.: Fuzzy-PID control for macro-quadrotor UAV. *Ordnance Ind. Autom.* **33**(6), 58–62 (2014)
5. S. Bouabdallah, A. Noth, R. Siegwart, PID vs LQ control techniques applied to an indoor micro quadrotor. In: IEEE International Conference on Intelligent Robots and Systems, pp. 2451–2456 (2004)
6. Mellinger, D., Kumar, V.: Minimum snap trajectory generation and control for quadrotors. In: IEEE International Conference on Robotics and Automation, pp. 2520–2525, May 2011
7. Jie, L., Xiao-hui, Q., Shuai-tao, H., et al.: Design and implementation of flight control system for small quad-rotor. *China Meas. Test* **02**, 90–93 (2014)
8. Bouabdallah, S., Siegwart, R.: Backstepping and sliding-mode techniques applied to an indoor micro quadrotor. In: International Conference on Robotics & Automation, Barcelona, pp. 2274–2252, April 2005
9. Hang-ke, C., Dong-sheng, Z.: Modeling and attitude control simulation for four-rotor aircraft in Hover. *Comput. Simul.* **30**(11), 41–45 (2013)
10. Kun, W., Yuanli, C.: Tracking-differentiator based on artificial neural network. *Comput. Dig. Eng.* **42**(3), 378–381 (2014)
11. Jingqing, H.: From PID technology to “active disturbance rejection control” technology. *Control Eng. China* **9**(3), 13–18 (2002)



Fuzzy Adaptive PID Control for Translational Flight of a Tail-Sitter UAV

Leiping Xi, Dizhou Zhang, and Sen Yang^(✉)

Department of UAV Engineering, Army Engineering University,
Shijiazhuang 050003, China
568657132@qq.com

Abstract. This study proposes a new fuzzy adaptive PID control strategy to enhance the performance of a tail-sitter UAV in translational flight. First of all, the dynamic model of UAV is built, and then the interference of model parameter perturbation and the disturbance in translational flight is analyzed; Finally, the fuzzy adaptive PID controller is designed to solve the problem of model uncertainty. The simulation results show that the controller has good tracking quality and robustness.

Keywords: Tail-sitter UAV · Vertical takeoff and landing · Fuzzy adaptive PID control · Flight control

1 Introduction

VTOL aircraft is a special aerial vehicle which can take off and land with zero velocity like a helicopter, and fly at high speed horizontally like fixed wing airplanes [1]. The tail-sitter UAV is a kind of superior performance VTOL aircraft due to its flexible handling and does not require complex thrust reversing mechanism. Tail-sitter UAV vertically takes off to a certain height and turn into the horizontal flight mode. When it lands, it climbs with the nose up, and then reduces the thrust and lands vertically [2]. The tail-sitter UAV converts from vertical flight to horizontal flight by synchronous rotary of the lift orientation and fuselage. There are many unmanned aerial vehicles which apply this kind of lift system, e.g., Goldeneye as shown in Fig. 1.

Tail-sitter UAV achieves the pitch, yaw, and roll movement in flight process by changing the aileron deflection and the rotating speed of rotors [3]. The two-rotor tail-sitter UAV model is shown in Fig. 2.

The paper is organized as follows: In Sect. 2, the general attitude dynamic model of a tail-sitter UAV is presented; in Sect. 3, the fuzzy PID controller is designed to solve the problem of model uncertainty in translational flight and the performance of the controller is verified; and in Sect. 4, the conclusion of the paper is given.

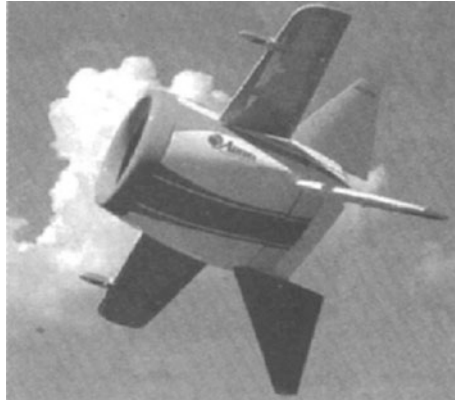


Fig. 1. Goldeneye UAV

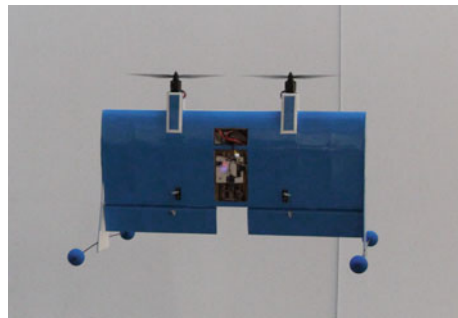


Fig. 2. Two-rotor tail-sitter UAV model

2 Vehicle Attitude Dynamic Model

Suppose that the aircraft is a rigid body and the acceleration of gravity does not change with flight altitude, and the curvature of the Earth is ignored, then take the ground coordinate system O-XYZ as inertial coordinate system. The equations of attitude motion in the body coordinate system are built as given below:

$$\dot{p} = \frac{I_z}{I_x I_z - I_{xz}^2} (M_x - (I_z - I_y)qr + I_{xz}pq) + \frac{I_{xz}}{I_x I_z - I_{xz}^2} (M_z - (I_y - I_x)pq - I_{xz}rq) \quad (1)$$

$$\dot{q} = \frac{1}{I_y} (M_y - (I_x - I_z)pr + I_{xz}(r^2 - p^2)) \quad (2)$$

$$\dot{r} = \frac{I_x}{I_x I_z - I_{xz}^2} (M_z - (I_y - I_x)pq - I_{xz}rq) + \frac{I_{xz}}{I_x I_z - I_{xz}^2} (M_x - (I_z - I_y)qr + I_{xz}pq) \quad (3)$$

where $\omega_b = [p \ q \ r]^T$ denotes the angular velocity of the aircraft in the coordinate system of the body, $M_{A,T} = [M_x \ M_y \ M_z]^T$ denotes the total torque applied to the center of gravity of the aircraft, and I_x, I_y, I_z contain the moments of inertia of the aircraft, which can be obtained by the method in [4], $I_x = 0.0122 \text{ kg m}^2$, $I_y = 0.0048 \text{ kg m}^2$, $I_z = 0.0079 \text{ kg m}^2$; I_{xz} denotes the product of inertia, because the aircraft is using symmetric airfoil, so $I_{xz} = 0$. Equations (1)–(3) can be simplified as

$$\dot{p} = \frac{(I_y - I_z)qr}{I_x} + \frac{M_x}{I_x} \quad (4)$$

$$\dot{q} = \frac{(I_z - I_x)pr}{I_y} + \frac{M_y}{I_y} \quad (5)$$

$$\dot{r} = \frac{(I_x - I_z)pq}{I_z} + \frac{M_z}{I_z} \quad (6)$$

$\Phi = [\phi \ \theta \ \psi]$ describes the vehicle orientation expressed in the classical roll, pitch, and yaw angles (Euler angles), and the relationship between Euler angles and the body angular velocity is as follows:

$$\dot{\phi} = p + (q \sin \phi + r \cos \phi) \tan \theta \quad (7)$$

$$\dot{\theta} = q \cos \phi - r \sin \phi \quad (8)$$

$$\dot{\psi} = (q \sin \phi + r \cos \phi) \sec \theta \quad (9)$$

The transition from vertical flight to horizontal flight of the tail-sitter UAV relies mainly on the large angle maneuver in the pitch direction. Thus, the major research is focused on the control law in the pitch direction during the transition.

Assuming that the angular rates of roll direction and yaw direction are zero, the kinetic equation of pitch angular motion is given by (19)

$$\ddot{\theta} = \frac{M_y}{I_y} \quad (10)$$

$M_y = \frac{1}{2} V^2 \rho S c C_m$ denotes the total torque in the pitch direction. V is the air velocity, which is the downwash flow velocity of the propeller when hovering and the synthetic speed of the airflow and downwash flow in the level flight. ρ denotes the density of the air, S is the area of wing, and c is the length of the wing chord. C_m is the pitch moment coefficient, which can be calculated by the method in [5].

$$C_m = C_{m_{\alpha}} + C_{m_x} \alpha + C_{m_{\delta_e}} \delta_e + C_{m_q} q \frac{c}{V} \quad (11)$$

Substituting (11) into (10), we get

$$\ddot{\theta} = \frac{\frac{1}{2}V^2\rho Sc}{I_y} [C_{m_z}\theta + C_{m_{\delta_e}}\delta_e + C_{m_q}\frac{c}{V}\dot{\theta}] \tag{12}$$

3 Longitudinal Attitude Control During the Transition

With the change of pitch angle and level flight speed during the transition from vertical flight to horizontal flight, the pitch moment coefficient also creates a larger change, which is nonlinear. As a result, the established dynamic model exists interference of model parameter perturbation and the disturbance is difficult to model accurately [6]. On the overall trend, it will increase the aerodynamic efficiency of ailerons. Therefore, if the controller is still designed to adapt to the vertical flight mode, aircraft will produce overshoot in the pitch direction during the transition and level flight, which will lead to aircraft out of control and crash.

The step response of PID controller for pitch motion with interference of model parameter perturbation is shown in Fig. 3. Pitch motion developed severe overshoot shocks after joining interference at 2 s. Therefore, the traditional PID controller cannot meet the needs of flight control of the vehicle during the transition.

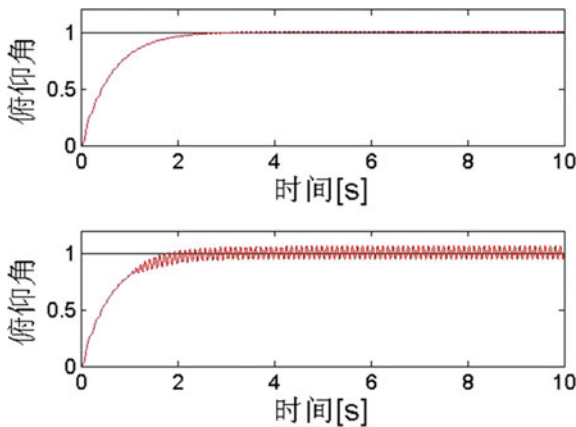


Fig. 3. The step response with interference of model parameter perturbation

Fuzzy control, which is on the basis of the operator’s experience, is independent of the accurate mathematical model of the controlled object. A fuzzy self-tuning PID controller is designed to control the longitudinal attitude of the aircraft during the transition. Fuzzy self-tuning PID controller is based on the conventional PID control. According to the error E and error change rate DE of feedback value and target value, the PID parameters are online self-tuning by using the method of fuzzy reasoning, to

meet the different requirements of the controller parameters on different running status. Thus, controlled objects can obtain good dynamic and static performance and adaptive performance. The designed fuzzy self-tuning PID controller of pitch motion is shown in Fig. 4.

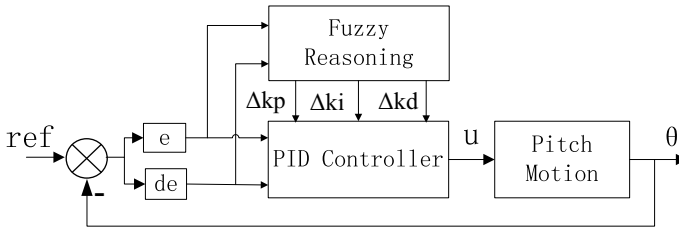


Fig. 4. Fuzzy self-tuning PID controller structure

E and DE are the inputs of the fuzzy controller, and ΔK_p , ΔK_i , ΔK_d are outputs. The domains of E, DE, ΔK_p , ΔK_i , ΔK_d are {3, 2, 1, 0, 1, 2, 3}, fuzzy sets are {NB, NM, NS, O, PS, PM, PB}.

The PID parameters setting principles are summarized as follows:

When the $|E|$ is big, K_p is taken as larger value in order to improve the response speed, K_i is smaller to avoid large overshoot, and in order to prevent excessive instantaneous of $|DE|$, K_d is taken as a smaller value.

When the $|E|$ is medium, K_p and K_i are taken as a medium value, and K_d is taken as a larger value to reduce the overshoot.

When the $|E|$ is small, K_p is smaller in order to improve the stability. K_i and K_d are taken as larger value to reduce static error.

According to the above principles, the fuzzy control rules of Δk_p , Δk_i and Δk_d are described in Table 1, 2, and 3.

Table 1. Fuzzy control rules of Δk_p

Δk_p		DE						
		NB	NM	NS	ZO	PS	PM	PB
E	NB	NM	NS	ZO	PB	PS	NS	NM
	NM	PS	PS	PS	PM	PS	NS	NS
	NS	PM	PB	PM	PB	ZO	NB	NM
	ZO	NS	ZO	PS	PM	PS	ZO	NS
	PS	NM	NB	ZO	PB	PM	PB	PM
	PM	NS	NS	PS	PM	PS	PS	PS
	PB	NM	NS	PS	PB	ZO	NS	PM

Table 2. Fuzzy control rules of Δk_i

Δk_i		DE						
		NB	NM	NS	ZO	PS	PM	PB
E	NB	NS	ZO	PM	PB	ZO	NS	NM
	NM	NB	ZO	PB	PM	PB	ZO	NB
	NS	NB	ZB	ZO	PS	PS	NM	NB
	ZO	NB	NM	ZO	ZO	ZO	NM	NB
	PS	NB	NM	PS	PS	ZO	NB	NB
	PM	NB	ZO	PB	PM	PB	ZO	NB
	PB	NM	NS	ZO	PB	PM	ZO	NS

Table 3. Fuzzy control rules of Δk_d

Δk_d		DE						
		NB	NM	NS	ZO	PS	PM	PB
E	NB	PS	NS	NS	NM	PS	PB	PB
	NM	NS	ZO	ZO	NM	PS	PS	PB
	NS	NS	PS	NS	NS	NM	NB	NS
	ZO	NB	NM	NB	NB	NB	NM	NB
	PS	NS	NB	NM	NS	NS	PS	NS
	PM	PB	PS	PS	NM	ZO	ZO	NS
	PB	PB	PB	PS	NM	NS	NS	PS

At the beginning of the simulation, the normal PID controller parameters K'_p, K'_i, K'_d are tried out. Then the parameters are online corrected by the fuzzy self-tuning PID controller, and $\Delta K_p, \Delta K_i, \Delta K_d$ are got. The parameters of the system using fuzzy self-tuning PID controller are as follows (Fig. 5):

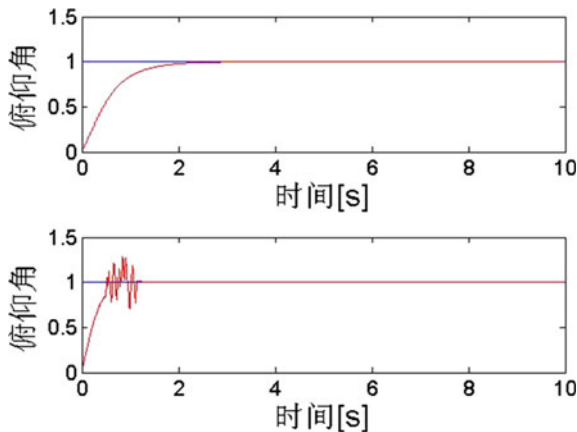


Fig. 5. The step response of pitch motion with fuzzy self-tuning PID controller

$$K_p = K'_p + \Delta K_p, K_i = K'_i + \Delta K_i, K_d = K'_d + \Delta K_d$$

The system is still able to track the step signal accurately after joining interference of model parameter perturbation. So, the pitch motion of the aircraft can overcome the interference with the designed fuzzy self-tuning PID controller which has strong robustness.

4 Conclusions and Future Works

A fuzzy self-tuning PID controller is designed to solve the problem of interference of model parameter perturbation during the transition from vertical flight to horizontal flight. In order to overcome the interference and enhance the robustness of the system, fuzzy control rules are used in tuning the parameters of PID controller. Thus, a new solution is provided for the flight control during the transition of the tail-sitter UAV.

References

1. Xi, L., Zhu, Q., Zhang, D.: Sliding mode control design based on fuzzy reaching law for yaw angle of a Tail-sitter UAV. In: 2016 22nd International Conference on Automation and Computing, Colchester, United Kingdom, vol. 09, pp. 238–243 (2016)
2. Stone, H.: Aerodynamic modelling of a wing-in-slip—stream tail-sitter UAV. In: 2002 Biennial International Powered Lift Conference and Exhibit, Williamsburg, 5–7 Nov 2002
3. Hugh Stone, R.: The T-wing tail-sitter research UAV. In: 2002 Biennial International Powered Lift Conference and Exhibit, pp. 1–10 (2002)
4. Mengzhao, L.: Moment of inertia of aircraft and its estimation. *J. Aircr. Des.* **3**, 13–20 (1997)
5. Etkin, B., Reid, L.D.: *Dynamics of Flight Stability and Control*. Wiley (1996)
6. Escareno, J., Sanchez, A., Garcia, O., Lozano, R.: Modeling and global control of the longitudinal dynamics of a coaxial convertible mini-UAV in hover mode. *J. Intell. Rob. Syst.* **54**, 261–273 (2009)



Renewable Energy-Aware IoT Data Aggregation for Fog Computing

Yusong Fu¹, Dapeng Li^{1,2(✉)}, Feng Tian¹, and Yongan Guo¹

¹ Jangsu Engineering Research Center of Communication and Network Technology,

Nanjing University of Posts and Telecommunications, Nanjing 210003, China
dapengli@njupt.edu.cn

² National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China

Abstract. This paper considers the problem of renewable energy and spectrum allocation for a fog-based IoT network where the fog node can request energy from multiple renewable power suppliers (RPSs) to serve the end devices. We consider that RPSs of different relay nodes can form coalitions. RPSs in the same coalition can better coordinate their price strategy. Then, we analyze an independent RPS's incentive to join a coalition or stay independent, and a nonindependent RPS's incentive to deviate to join other coalitions or stay in present coalition. Then, in this paper, we achieve the Nash stable coalition structure, and the corresponding energy pricing for the structure. Finally, we give simulation results, and results show that RPSs, the fog node, and spectrum owner can benefit from coalitions.

Keywords: Fog computing · M2M communication · Internet of things (IoT) · Renewable energy · Energy allocation · Coalitions formation

1 Introduction

Fog is a construction that distributes computation, communication [1], storage, and control along the cloud-to-things continuum and through the structure, these services can be closer to the end devices and users. And, Cisco introduces this new platform. It extends the cloud computing paradigm to the edge of the network [1–3]. However, fog is different from the cloud is that fog process part of services and data on fog devices instead of transmitting to cloud [4].

In [5], the authors propose a new multi-tier fog computing model and is for smart city applications. In [6], the authors summarize the opportunities and challenges of the edge computing (EC), validate the efficiency of EC, survey the edge systems widely, and then a comparative cloud system study is proposed by the authors. IoT services and applications can be run from the edge of the network as well as access points, and M2M gateways [7] with fog computing platform. Such works are committed to improve quality of service (QoS), reduce latency, and allow real-time data analysis. However, these works do not involve fog computing with renewable power suppliers (RPSs) problems.

In this paper, we explore a fog-based IoT network structure with price-sensitive end devices. There are multiple relay nodes in the network and each relay node is having multiple renewable power suppliers (RPSs). We will analyze how renewable energy allocation impacts the wireless traffic, the pricing scheme of the IoT network, and the strategy of the spectrum owner. And, in this paper, we propose a coalition structure, RPSs form coalitions, and formulate their common price strategy together. By forming coalitions, RPSs can coordinate their pricing scheme when providing energy to the network. Finally, the conclusion is that by forming RPSs coalitions, different RPSs can better coordinate their price decision which can reduce energy price, and there exists a Nash stable in coalition structure.

This paper is organized as follows. In Sect. 2, we describe a renewable fog-based IoT network used in this paper. The problem statement is described in Sect. 3. Section 4 is the analysis of the problem and a coalition structure is proposed here. We present the simulation and analysis in Sect. 5. Then, we draw conclusions in Sect. 6.

2 Network Model

Consider a network composed of a fog node, multiple relay nodes, spectrum owner, and a set μ of M MTDs attempt to access the fog node. The fog node, relay nodes, and the spectrum owner are connected to the core network through a wireless backhaul to transmit uplink or downlink data traffic. The fog node leases a bandwidth B from the spectrum owner during a certain time duration on the random access channel (RACH). The data traffic is uploaded to the fog node through multiple relay nodes.

2.1 The Data Aggregator and Spectrum Owner at the Edge

At the beginning of each gathering interval, the fog node broadcasts a beacon. When MTDs receive this beacon, they start to implement random access protocol. Clearly, end users or MTDs will transmit more data traffic if the p which is charged to end devices is lower, which will lead to more energy request from RPSs. Thus, we assume that the average number of M2M communication demands is the exponential function $Q(p) = ae^{-bp}$ in each gathering interval, where the a here is a random variable because the access demand of MTDs is uncertainty.

The fog node purchases the spectrum from the spectrum owner and provides to MTDs, in the network, the fog node, relay nodes, and the spectrum owner uses the same spectrum band and we assume that the whole price of the purchase price of spectrum is β . Each M2M device occupies a certain bandwidth B at the access stage, and a small-sized data packet is transmitted at this phase. The fog node broadcasts a beacon at the beginning of each gathering interval. When MTDs receive this beacon, they start to implement random access protocol. The whole bandwidth that the spectrum owner provides is ω . And the unit price of bandwidth is s . Thus, the profit of the spectrum owner is $Pi_s = s \cdot \omega - \beta$ (Fig. 1).

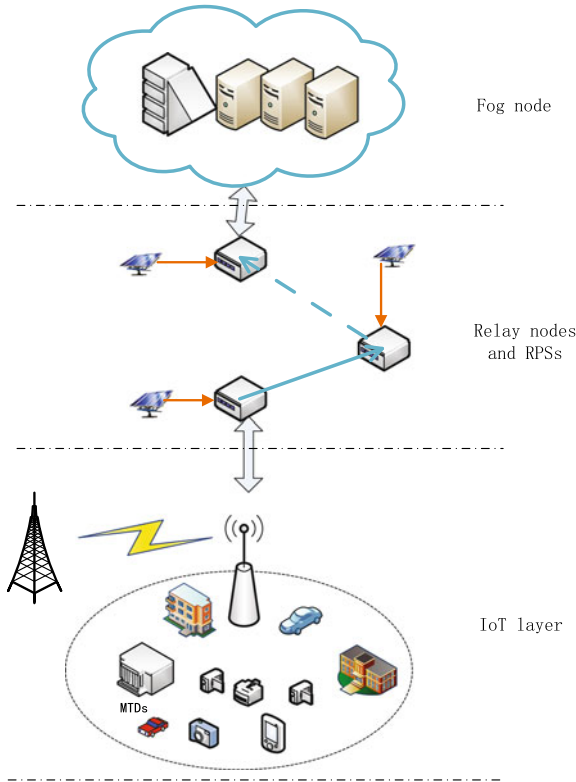


Fig. 1. Data traffic is transmitted in the network, and RPSs provide energy. In the cooperation scheme, RPSs form coalition and coordinate their energy price

We use a standard Slotted Aloha (SA) for random access in this paper, and the expected throughput of SA is $Ge^{(-G)}$, where, we use G to describe the average number of packets sent per slot. Each end device transmits an access preamble sequence to the fog node on RACH and the access sequence is randomly selected. On the same RA slot, if multiple devices transmit a different preamble sequence at the same time, there will be no collisions. We use the notation x_s to describe the number of available preamble sequences, and the notation B_s to denote the bandwidth of an SA slot. The notation t_a represents the duration of the SA slot in time domain. Note that, in an access interval, the total number of RA slots is $\frac{\omega\tau}{t_a B_s}$. Thus, $G = \frac{Q(p)t_a B_s}{\omega\tau x_s}$ (number of packets/slot) is derived. The unit of the throughput of the SA protocol is (number of successfully accessed devices)/slot. Then, in each gathering interval, the average number of successfully accessed MTDs, $T(\omega) = Q(p)e^{\frac{D(p)\kappa}{\omega}}$, where the κ here equals to $-\frac{t_a B_s}{\tau x_s}$ (Table 1).

Table 1. Table of notations

Symbol	Definition
m	The number of MTDs
B_s	The bandwidth of an SA slot
ω	The bandwidth the spectrum owner provides to the aggregator
h_i	The transmission distance of relay node i
ϵ	The bandwidth of relay nodes
R	Data transmitting rate of core network
\bar{L}_i	The path loss of fog node i
c_i	The unit cost of RPSs on relay nodes
w_i	Renewable energy price that RPS i charged the operator
p	Price charged to end users and MTDs
W	Vector of renewable energy price of RPS
B_j	Coalitions of RPSs
P_i	Transmit power of fog node i
s	The bandwidth unit price of the whole network

2.2 The Path Loss of Wireless Route of the Edge

Because the transmission follows a free path loss, according to the Friis' law, the path loss is given as $L(h) = 32.5 + 20 \log_{10}(f) + 10\gamma \log_{10}(h) + A \times h$, where the variable h is the link distance, f is the carrier frequency, A is the oxygen-and-rain attenuation coefficient, and γ is the path loss exponent, the notation depends on the geometry when transmitting. Relay nodes are independent with each other through how they price and energy that RPSs can provide. Let P_i be the transmission power of relay node i and h_i be the distance of each node. Thus, with the path loss $L(h_i)$ and Shannon theorem, we get $R = \epsilon \ln(1 + \frac{P_i L(h_i)}{n_0 \epsilon}) \Rightarrow P_i = \frac{(e^{\frac{R}{\epsilon}} - 1) n_0 \epsilon}{L(h_i)}$, where ϵ is the bandwidth of relay nodes, R is the rate of delivering one unit traffic through the network. Let $\phi(\epsilon) = \frac{(e^{\frac{R}{\epsilon}} - 1) n_0 \epsilon}{R}$, thus the energy demand $E_i = \frac{Q(p)}{R} P_i = Q(p) \frac{\phi(\epsilon)}{L(h_i)}$, which means that to deliver one unit traffic with rate R will consume energy is $\frac{\phi(\epsilon)}{L(h_i)}$. Let ω_i denote the energy price of RPS $i, i = \{1, \dots, n\}$ and we use c_i to present the cost of the RPS i .

2.3 Renewable Power Suppliers

Relay node $i = \{1, \dots, n\}$, each was powered by an RPS, denoted by RPS i , correspondingly. Such cost depends on energy efficiency and energy loss of transmission. RPSs in the network model are used to store and supply energy. In our model, we use renewable instead of conventional energy resources such as coal and petroleum. We denote the energy price is p_e and the production of RPS i by μ_i unit energy per unit time. The profits of the RPS i is $\prod_i = (w_i - c_i)Q(p)$.

2.4 The Profit of the Fog Node

In fog-based IoT networks, different relay nodes will not share their price decisions and other information with each other. According to the aforementioned network model, there is a sequential structure of decision-making. First, the fog node plays the role of the leader, determines and announces the service price p_f to optimize its expected profit. Then, the RPS i plays the role of follower, and according to the fog node price, RPSs decide the energy price p_i . And then, the fog node and the relay nodes adjust spectrum resource. The interaction among relay nodes, the fog node, and the spectrum owner can be formed as a Stackelberg game. The profit of the fog node is $\prod_f = (p - f)Q(p)$, the notation W here is the energy price charges to the fog node.

3 Problem Formulations

The aforementioned interaction among the fog node, the spectrum owner, and RPSs can be formulated as a Stackelberg game that will enable us to determine the following information: (1) What is the price that the fog node should pay each RPS and the spectrum owner at each period for transmitting the traffic and (2) How much the spectrum owner and RPSs provide spectrum and energy resource initially with the uncertain exponential traffic amount.

In this paper, we will use backward induction to solve the proposed game. By analyzing this game, we can find the equilibrium of the game. And, we begin with the unique energy price that RPSs choose in Stage I, then we propose a cooperation scheme in this paper. In the cooperation scheme, the individual RPS will form coalitions with one another freely and coordinate their price strategy in these formed coalitions. RPSs in formed coalition will offer a joint renewable price. For the process of the coalition formation, we use the concept of the Nash stable structure. The concept of Nash stable structure here means that there will be no unilateral deviations by individual RPS and RPSs in the same coalition behave as a whole with the same price strategy. For convenience, we define $N = \{1, \dots, n\}$ represents the set of relay nodes and we use $B = \{B_1, \dots, B_m\}$ to represent the coalition structure, where $\bigcup_{j=1}^m B_j = N, B_h \cap B_k = \phi$ for $h \neq k$. And, $W_{B_j} = \sum (w_i : i \notin B_j)$ stands for the alliances' wholesale price of $B_j \in B$ charge to the fog node which includes renewable energy price and other prices. Also, we use $C_{B_j} = \sum (c_i : i \in B_j)$ to express the cost of coalition $B_j \in B$ and $C = \sum_{j=0}^m C_{B_j}$ means the whole costs of all coalitions.

Pay attention to the equation $C_{B_j} = \sum (c_i : i \in B_j)$, it implies that there are no other extra costs when RPSs form coalitions. Thus, the coalition's profit function $\prod_{B_j}(W_{B_j})$ can be expressed as $\prod_{B_j}(W_{B_j}) = (W_{B_j} - C_{B_j})Q(p)$, the fog node's profit function $\prod_f = [p - \sum_{j=0}^m W_{B_j}]Q(p) = (p - W)Q(p)$. Each coalition will set their energy price strategy, then the fog node changes its wireless service price strategy p according to coalitions price strategy.

4 Analysis of Problems

When RPSs coalitions sell energy to the fog node, the coalitions formed by RPSs can better coordinate their price strategies. And, in this section, the goal is to characterize the energy price strategy of RPS coalitions. According to the result of the energy price decision, we identify the unique optimal price that fog node charges to MTDs and the price strategy of the spectrum owner. As a result, the pricing scheme obtained constructs a Stackelberg equilibrium, and the whole network will get more profit from the equilibrium.

Assumption 1. Recall that the number of successfully accessed requests $Q(p)$, we use $\eta(p)$ to represent the price elasticity of $Q(p)$, and $\eta(p) = -pQ'(p)/dp$, where $Q'(p) = dQ(p)/dp$. Assume that $\eta(p) = p/(\alpha + \beta p)$, where α and β are constants, $\alpha \geq 0$ and $\beta \leq 1$. The equation $\eta(p) = p/(\alpha + \beta p)$ requires that $\alpha + \beta p > 0$, and implies that the curvature property of $Q(p)$ is denoted by $\psi(p) = Q(p)Q''(p)/(Q'(p))^2$, where $Q''(p) = d^2Q(p)/dp^2$.

Proposition 1. *Given RPS coalition structure, in the Nash Equilibrium, all coalitions share the equal profit.*

Thus, we denote the profit of coalitions $\prod(m) = \Pi_{B_j}$. Without loss of generality, we assume that $|B_1| \leq \dots \leq |B_m|$, and one RPS can be classified into independent and nonindependent.

Proposition 2. *For the structure $\mathcal{B} = \{B_1, \dots, B_m\}$, when the demand function satisfies Assumption 1, we have following profits and strategies when the coalition structure is equilibrium.*

- (1) $W_{B_j}^* - C_{B_j} = \frac{1}{m}(W^* - C) = \frac{\alpha + \beta C}{1 - \beta m}$, when $\frac{\partial W^*}{\partial m} \geq 0$ for any $j \in \{1, \dots, m\}$.
- (2) $p^* = \frac{\alpha(1-\beta)m + \alpha + C}{(1-\beta m)(1-\beta)}$, and $\frac{\partial p^*}{\partial m} \geq 0$.
- (3) $\prod_{B_j}^* = \Pi(m) = \frac{(\alpha + \beta C)Q(p^*)}{1 - \beta m}$, when $\frac{\partial W^*}{\partial(m)} \leq 0$ for any $j \in \{1, \dots, m\}$.
- (4) $\prod_f^* = \frac{(\alpha + \beta C)}{(1-\beta)(1-\beta m)}Q(p^*)$, and $\frac{\partial \Pi_f^*}{\partial m} \leq 0$.

Here, we make an implicate assumption that $1 - \beta m \geq 0$. Actually, if $1 - \beta m < 0$, there will be no transactions between RPSs and coalitions.

From this proposition, we can see that in a given coalition structure, each coalition will gain the equal profit and the profit of a coalition is not related to the actual composition of the coalition, but only the demand function of MTDs and the number of the RPSs in network model. In order to promote the study of the proposed question 1 above, we need to recognize how RPSs allocate the coalition profit among members in the same coalition. We use Shapley value to generate allocation among different RPSs in coalitions because of its fairness. And, the Shapley value is often be used in gains allocation. According to the Shapley value, RPSs in this coalition share an equal distribution of the coalition profit in a coalition. According to the Nash stability concept, for an individual RPS, if it obtains rigorous benefits from such a deviation, then it tends to make a unilateral defection from the current coalition. We assume that for any given coalition structure, there are two feasible deviations, an RPS can deviate from the

current coalition to join another coalition, or to become independent. It means that an RPS will share more profit if it deviates the current coalition. We denote the equilibrium coalition profit in which the coalition is composed of m RPSs with $\Pi(m)$. Recall that $\Pi(m) = \frac{(\alpha + \beta C)Q(p^*)}{1 - \beta m}$.

We denote the number of RPSs in coalition B_j by $|B_j|$, and $U(m) = \frac{\Pi(m)}{\Pi(m+1)}$. Here, $U(m)$ represents the ratio of the coalition's profile before and after some RPSs defection, it indicates the stability of the coalition. If $U(m)$ is a small value, RPSs are more likely to become independent. And, we can say a given coalition is Nash stable if $U(m-1) \leq |B_2| + 1$ and $U(m) \geq |B_m|$. The former provided that there is at least one independent RPS and the latter provided that there is at least one nonindependent RPSs. For convenience, we denote $\mathcal{B}_{(1)(2)} = \{B_1, \dots, B_m\}$, where $|B_1| = 1, |B_2| = \dots = |B_m| = 2$, $\mathcal{B}_{(2)(2)} = \{B_1, \dots, B_m\}$, where $|B_1| = \dots = |B_m| = 2$.

Result. For $Q(p) = ae^{-bp}$, $\mathcal{B}_{(1)(2)}$ is uniquely stable when n is an odd and $\mathcal{B}_{(2)(2)}$ is stable when n is an even.

According to Assumption 1, the demand curvature $\Psi(p) = 1$, and it can be derived that $U(m) = e$. If in a given structure, and there are m coalitions in this coalition structure, only two RPSs in such a coalition structure are independent. When these two RPSs form a new coalition, they can share more coalition profit $\Pi(m-1)$ than they are independent, because $U(m-1) = \frac{\Pi(m-1)}{\Pi(m)} > 2$, and the coalition structure will not be stable. When a coalition has l RPSs, and $l \geq 3$, then the coalition will not be stable, RPSs would deviate the current coalition and to be independent, because of $\frac{\Pi(m)}{l} \leq \frac{\Pi(m)}{3} < \Pi(m+1)$. Therefore, we can conclude that a coalition will never be stable if three or more RPSs are in one coalition. Thus, we can get conclusion that the coalition structure $\mathcal{B}_{(1)(2)}$ is stable for an odd n , and when n is an even, $\mathcal{B}_{(2)(2)}$ is stable.

5 Simulation Results and Analysis

In this section, how the RPSs' coalitions formation and how coalitions influence profits through numerical simulations will be studied. In simulations, we set the bandwidth of the backhaul wireless link as 20 MHz, at 25×10^3 GHz frequency, and the basic data rate to 1×10^3 Mbit/s. For the path loss function $L(h)$, we set $\gamma = 2.2$ and the value of A to 8 dB/km in function $L(h)$. And, the notation $\phi(\in)$ is a fixed value.

Figure 2 shows the profit of each RPS with the different distances of different relay nodes and the profit of the fog node with different relay nodes. We can see that the RPS 3 owns more profit than other RPSs, and RPS 3 follows the minimum distance.

We can see that as coalitions increase from 2 to 22, with the bandwidth of relay nodes is 20 MHz, the profits decrease to a very low level. With different numbers of coalitions, the equilibrium profit of coalitions and the fog node how to change, respectively. We can see that with the increase of coalitions from 1 to 10, profits of the fog node and coalitions decrease to a very low value. The profit will decrease to the

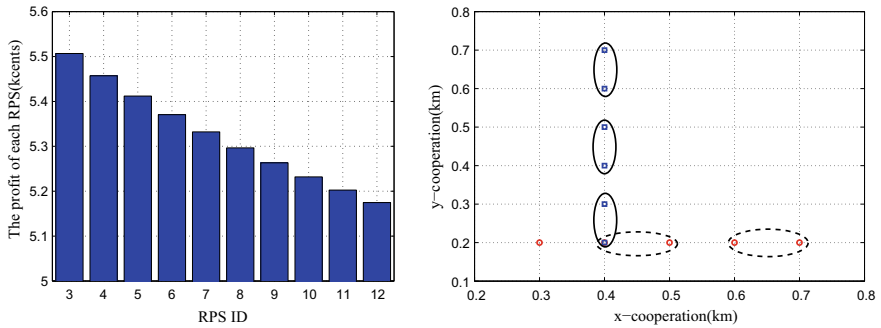


Fig. 2. The profit of each RPS under the conditions of equilibrium at different relay nodes and the formation of coalitions

lowest when coalitions are not formed. RPSs can form coalitions with each other freely and determine their energy prices. Different coalitions set their prices independently, RPSs in a coalition will set a common energy price for the renewable energy provided by the coalition.

6 Conclusion

We have conducted a comprehensive study on the issue of renewable power allocation for a fog-based network with price-sensitive end devices in this paper. We proposed a cooperation scheme where RPSs provide that renewable power suppliers tend to form coalitions, through the coalition structure, and they can better coordinate their pricing strategies. And, we proved that RPSs always prefer selling in groups in Nash equilibrium. By forming RPSs coalitions, different RPSs will better coordinate their price decision which can reduce energy price. Lower energy price means the fog node will charge end devices a lower fee and more end device will transfer data to the fog node. Furthermore, the result has also demonstrated that RPSs, the fog node, and spectrum owner can benefit from coalitions.

Acknowledgments. This work was supported in part by the NSF of Jiangsu Province under Grant BK20161518 and Grant BK20171444, in part by the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University, under Grant 2018D05, in part by the National Natural Science Foundation of China under Grant 61772287, Grant 61771252, and in part by the Open Research Fund of Jiangsu Engineering Research Center of Communication and Network Technology, NJUPT.

References

1. Mebrek, A., Merghem-Boulahia, L., Esseghir, M.: Efficient green solution for a balanced energy consumption and delay in the IoT-fog-cloud computing. In: 2017 IEEE 16th International Symposium on Network Computing and Applications (NCA), pp. 1–4 (2017)

2. Yan, S., Peng, M., Wang, W.: User access mode selection in fog computing based radio access networks. In: 2016 IEEE International Conference on Communications (ICC), pp. 1–6 (2016)
3. Shen, S., Huang, L., Zhou, H., Yu, S.: Multistage signaling game-based optimal detection strategies for suppressing malware diffusion in fog-cloud-based IoT networks. *IEEE Internet Things J.* (2018)
4. Deng, R., Lu, R., Lai, C., Luan, T.H., Liang, H.: Optimal workload allocation in fog-cloud computing toward balanced delay and power consumption
5. He, J., Wei, J., Chen, K., Tang, Z., Zhou, Y., Zhang, Y.: Multi-tier fog computing with large-scale IoT data analytics for smart cities. *IEEE Internet Things J.* **99** (2017)
6. El-Sayed, H., Sankar, S., Prasad, M., Puthal, D., Gupta, A., Mohanty, M., Lin, C.T.: Edge of things: the big picture on the integration of edge, IoT and the cloud in a distributed computing environment. *IEEE Access* **6**, 1706–1717 (2017)
7. Hou, X., Li, Y., Chen, M., Wu, D., Jin, D., Chen, S.: Vehicular fog computing: a viewpoint of vehicles as the infrastructures. *IEEE Trans. Veh. Technol.* **65**(6), 3860–3873 (2017)



Renewable Energy Powered IoT Data Traffic Aggregation for Edge Computing

Cunchao Peng¹, Dapeng Li^{1,2}(✉), Feng Tian¹, and Yongan Guo¹

¹ Jangsu Engineering Research Center of Communication and Network Technology, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

² National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China
dapengli@njupt.edu.cn

Abstract. With the development of the Internet of Things (IoT) industry and the arrival of the 5G era, edge computing is considered to be the more suitable computing technology for the IoT. In this paper, we propose an edge-computing-based M2M data aggregation wireless transmission system powered by efficient renewable energy allocation servicing for the edge devices. The pricing scheme problem is formulated as a Stackelberg game between the operator and multi-RPSs. Simulation results show how the previous pricing scheme and bandwidth of each node affect the renewable energy storage levels of each RPS and his own profit. The results also show the operator's optimal service price scheme and the equilibrium renewable energy storage level of each RPS.

Keywords: Edge computing · M2M communication · IoT · Renewable Energy Supplier (RPS) · Pricing scheme · Simultaneous decisions · Sequential decisions

1 Introduction

Edge computing extends cloud computing and plays a more and more important role in the application of the Internet of Things (IoT) such as latency-sensitive edge devices [1, 2]. Further, edge computing was proposed using resources to data processing at the edge of network instead of establishing channels for the centralized cloud storage and utilization. In addition, edge computing can save bandwidth and facilitate distributed architectures as the data traffic processed at the edge of the network and considered to be a solution to the resource-limited edge devices while offloading data traffic [3].

With the increasing demand for latency-sensitive service, the more efficient energy management in the edge computing is considered to be reliable solution. In [4, 5], the authors proposed a cooperation mechanism based on fog/edge computing from resource-constrained devices for Wireless Sensor Networks (WSN) applications which can effectively allocate the computing load among different

nodes. Kiani and Ansari [6] studied the edge-computing-based Non-Orthogonal Multiple Access (NOMA) technique to reduce the latency and promote the Quality of Service (QoS) for 5G networks. In [7], the authors outlined the random access challenges from the LTE-A perspective and analyzed the most common schemes to avoid the RACH access congestion. However, most of these efforts focus on the solutions to the RACH access congestion and latency.

In this paper, we introduce a data aggregator and multi-hops data transmitting system for alleviating the massive access and enhance the M2M communication. These relay nodes are powered by the renewable energy and we studied the optimal energy storage strategy to obtain the optimal network performance.

The rest of the paper is presented as follows. In Sect. 2, we illustrate the system model architecture and propose the problem formulation. Section 3 provides the unique renewable allocation NE solutions and analyzes the network performance. Then, we provide numerical results and analyze the performance under different impact factors in Sect. 4. Section 5 concludes the paper.

2 System Model

In this section, we first discuss each component of the network model and the various parameters. Then, some assumptions and uncertain traffic demand functional formation provided for the service chaining. An overview of this network model is shown in Fig. 1 and summed up main notations used in this paper in Table 1.

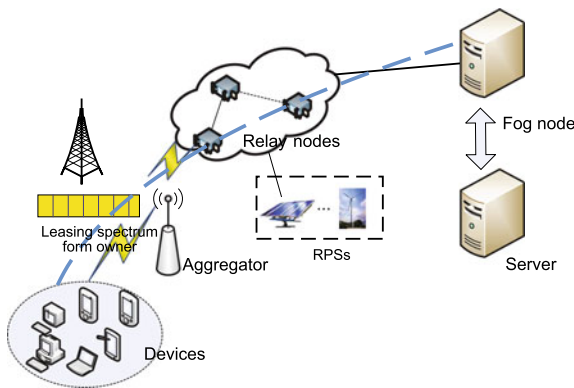


Fig. 1. The traffic transmitting to the fog node and furthermore to the corresponding server through a set of relay nodes which are powered by renewable energy suppliers.

2.1 Network Description

M2M Aggregator and Operator The aggregator is collecting the M2M data traffic from edge devices in certain area. At first, the aggregator broadcasts a beacon at the beginning of each gathering interval. Then, MTD devices begin to

implement their random access protocol as long as receiving such a beacon. Let p be the service price chosen by the operator. The average number of demands in each gathering interval is D , and depends on the price of p . We take the form of exponential function to capture price sensitivity of traffic demand $\tilde{D}(p) = \varepsilon \cdot y(p) = \varepsilon \cdot ap^{-b}$.

The operator is leasing the bandwidth W with unit price k from the cognitive operator used as RACH resource. From the access model, we know that each M2M device occupies a certain bandwidth w which has a unit price l to transmit a small-sized data traffic packet. Let $T(W)$ be the successful access number of all MTDs in each interval τ . Thus, the total on-demand band cost would be $T(W) \cdot w \cdot l$.

For random RACH access, we use Slotted Aloha (SA) scheme instead of Pure Aloha (PA) which improves channel utilization. We assume that the usable preamble sequence for each time slot has a value of x and the bandwidth of an SA slot is w . At result, we get the number of edge devices successfully accessed to the network, $T(W) = \frac{W\tau x}{w \cdot t} \cdot G \cdot e^{-G} = D(p)e^{-\frac{D(p)wt}{W\tau x}} = K \cdot D(p)e^{-\frac{D(p)}{W}}$.

For the operator, it gains r share of the revenue and remits to the RPS. In any Nash equilibrium, the traffic through the service network chaining will be identical. Let $C = \sum_{i=1}^n c'_i = \sum_{i=1}^n \frac{L(h_i)\rho\rho'}{\varphi_B} \cdot c_i$ be the each-unit energy storage cost and on-demand bandwidth cost of $c_0 = T(W) \cdot w \cdot k$ through the service chaining network. Thus, we can express the operator's profit function as $\Pi_o(r) = -c_0q + rpE[\min\{q, \tilde{D}(p)\}]$.

Renewable Energy Suppliers and Assumptions Each node may have m RPSs to compete with each other to be the unique energy resource. Finally, there will be only one RPS i to be the renewable energy supplier of relay node i . The winner of the competition must provide lower price p_i of energy and have lower energy storage cost of c_i .

In our model, we assume the relay nodes of the chaining are all LOS so as the operator can acquire the path loss value along the wireless communication links. From Friis' law, we can describe the path loss in dB as the expression: $L(h) = 32.5 + 20 \log_{10}(f) + 10\gamma \log_{10}(h) + A \times h$, where h stands for link distance, f is the carrier frequency, and γ is the path loss exponent. A is the oxygen-and-rain attenuation coefficient.

Consider the path loss model, the later RPS should provide much more renewable energy to guarantee the data traffic collected from end users transport to core network. Let P_i be the transmission power and h_i be the hop distance of node θ_i , and according to wireless network transmission model, we have $R = \frac{(e^{\frac{R}{B}} - 1)n_0B}{L(h_i)}$, where B is the bandwidth used for M2M data transmission.

2.2 Problem Formation

The strategy of the operator is the pricing scheme p and the strategy of each RPS is energy storage level and the price charged the operator. For the lower service price, the RPS may not store enough energy to guarantee the QoS of end

Table 1. Table of notations

Symbol	Definition
h_i	The transmission distance of relay node θ_i
P_i	Transmit power of relay node θ_i
R	Data transmitting rate across the service chaining
W	The bandwidth of aggregator leasing from cognitive operator
B	The bandwidth of MTDs for transmitting data traffic
\bar{L}_i	The path loss of relay node θ_i
ρ	The energy storage efficiency parameter
ρ'	The energy transmission efficiency parameter
p	Operator's service price for transmitting data traffic

users which affects the operator's profit. And finally, for this proposed game will reach the point where a unique Pareto-optimal NE between the operator and the RPSs. Specifically, the proposed game can be described as follows.

The operator will identify service price p to maximize its profit and the equilibrium energy allocation tends to be unique. The two-stage Stackelberg game will help us to determine the following information. What the price the operator should pay to each RPS for forwarding the uncertain traffic over the service chaining, considering the timely service and his own profit; What strategy the RPS should choose his renewable energy price under the service price scheme p by the operator; What the price the operator should choose about the on-demand bandwidth through the chaining, considering the relationship between the traffic demand and the service price scheme p .

To identify the unique Pareto-optimal NE from the proposed game, we will use backward induction, which begins with the followers game (Stage II) and goes back to the interaction between the operator and the RPS (Stage I), as shown in Fig. 2. Then, using the result of Stage II considering QoS cost, we then find the optimal service price scheme for the operator in Stage I. Therefore, we will get the pricing scheme p and the relevant energy price of RPS which is constituting a Stackelberg equilibrium.

3 Analysis and Performance of the RPSs' Price Problem

In this section, we will discuss the strategies of energy price of RPSs, given the service pricing scheme p by the operator. Then, identify the unique Pareto-optimal NE between the operator and each RPS.

Define $z \equiv q/y(p)$, for the service chaining, the final demand traffic depends on the chosen value of p instead of each individual p_i . Thus, we can get the total expected profit of the service chaining network for offloading traffic $\Pi_c(p, q) = -(C + c_0)q + pE[\min\{q, \tilde{D}(p)\}]$. Obviously, the optimal service price and the

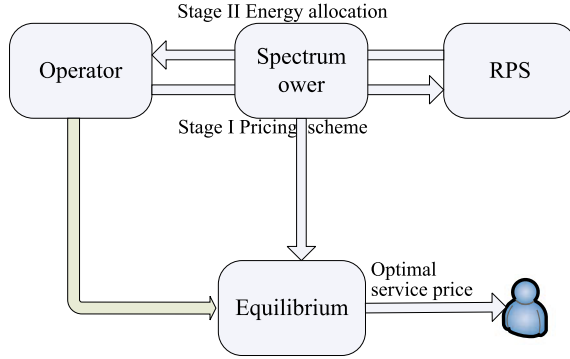


Fig. 2. The solution of optimal service price between operator and RPSs.

equilibrium factor z satisfy the following two first-order conditions as $p_c^* = b(C + c_0) \cdot z_c^*/((b - 1)[z_c^* - \Lambda(z_c^*)])$ and $F(z_c^*) = ((b - 1)\Lambda(z_c^*) + z_c^*)/(bz_c^*)$.

3.1 Simultaneous Energy Price Decisions of RPSs

The RPSs sell its energy to relay node for transmitting data traffic and signing the contract about the revenue sharing with the operator. Under such a contract, each RPS is deciding his own energy price p_i . Define $P_{-i} \equiv \sum_{j=1}^n p_j - p_i = p - p_i$. Let $\Pi_i(p_i | P_{-i}, z)$ denote the RPS's expected profit function, we have $\Pi_i(p_i | P_{-i}, z) = y(p_i + P_{-i})\{(1 - r)p_i[z - \Lambda(z)] - c_i z\}$.

For any prices chosen by all other RPSs, RPS i 's profit function $\Pi_i(p_i | P_{-i}, z)$ is quasi-concave in his price p_i and has the unique maximizer. Then, we have the following theorem.

Theorem 1. For any given factor z , if $b > n$, the n RPS' pricing game has a unique Nash equilibrium that is given by $p_i^*(z) = \frac{z}{z - \Lambda(z)} \cdot \frac{1}{(b - n)(1 - r)} [(b - n)c_i + C]$, $i = 1, \dots, n$.

In equilibrium, the total price for a set of n RPSs can be calculated as $p^*(z) = \sum_{i=1}^n p_i^*(z) = bCz/((z - \Lambda(z))(b - n)(1 - r))$. If $d[xh(x)]/dx = h(x) + xdh(x)/dx > 0$, then each RPS i 's profit function $\Pi_i(p_i^* | P_{-i}^*, z)$, $i = 1, \dots, n$ is quasi-concave in z and reaches its maximum at $z = z_c^*$ -the centralized factor.

The two-stage procedure works well for solving the price gaming problem between the operator and RPSs. Substituting $z = z_c^*$ and $p = p^*(r)$ into operator's profit function, we can rewrite it as $\Pi_o(r) = \frac{\alpha(b - n)^{b-1}[z_c^* - \Lambda(z_c^*)]^b}{[b(1 - \alpha)]^b(C + c_0)^{b-1}(z_c^*)^{b-1}} \cdot g(r)$, where $\alpha = \frac{c_0}{C + c_0}$ and $g(r) = \{(1 - \alpha)b + (b - n)\alpha\}r - (b - n)\alpha\}(1 - r)^{b-1}$.

Theorem 2. Theorem 2. For any $b > n$, the operator's profit function $\Pi_o(r)$ is quasi-concave in r , and has the unique maximizer $r^* = \frac{(b - n - 1)\alpha + 1}{b - n\alpha}$.

As expected, if the operator has a bigger share α of the revenue, it will gain more profit. However, when there are more RPS involved in the service chaining, the more share $(1 - r)$ of the revenue will leave to RPSs. For the service elasticity

b, its effect is more complicated on r^* . An increase in b will result in lower traffic demand and optimal energy service price.

3.2 Sequential Energy Price Decisions of RPSs

For given energy price choose by RPS 1 through $i-1$, RPS i choosing its own energy price p_i . We define $P_{\{1,i-1\}} \equiv \sum_{j=1}^{i-1} p_j$ and $P_{\{i+1,n\}}(p_i) \equiv \sum_{j=i+1}^n p_j$. Thus, we can write the pricing problem as the following dynamic programming: $\max \Pi(p_i | P_{\{1,i-1\}}, z) = y[P_{\{1,i-1\}} + p_i + P_{\{i+1,n\}}] \cdot \{(1-r)p_i[z - A(z)] - c_i z\}$. If the price of elastic parameter $b > 1$, then for any given factor z , the n RPSs' sequential pricing game has a unique equilibrium satisfying the first-order condition and we can get the n RPSs' total price $P^*(z) = \sum_{i=1}^n p_i^*(z) = (b/(b-1))^n \cdot Cz / ((1-r)[z - A(z)])$.

Theorem 3. *The operator's profit function $\Pi_o(r)$ is quasi-concave in r , and has the unique maximizer $r^* = \frac{(1-\alpha)b^{n-1} + \alpha(b-1)^n}{(1-\alpha)b^n + \alpha(b-1)^n}$.*

The operator will be obtained optimally profit as a big share of r^ of the revenue and also has a big share of α of the total channel cost or larger numbers of n RPSs. The outcome is similar to simultaneous energy prices's decision of RPSs.*

3.3 Simultaneous Versus Sequential Channel Performance

Here, we provided some important features for analyzing the network performances of two settings in decentralized solutions as the following proposition.

Proposition 1. *(a) In the simultaneous decisions of each RPS, all RPSs have same amount of profit, even through each RPS incurs different renewable energy costs and storage levels. That is, we have $\frac{\Pi_{i+1,I}}{\Pi_{i,I}} = 1, i = 1, \dots, n$. (b) In the sequential decisions of each RPS, RPS $i+1$ earns $b/(b-1)$ times of RPS i 's profit. That is $\frac{\Pi_{i+1,I,I}}{\Pi_{i,I}} = b/(b-1) > 1, i = 1, \dots, n-1$.*

We can conclude from this that the operator's profit in sequential decisions is always higher than that of simultaneous decisions for which lower price may have more devices trying to access the network and will generate more traffic through the service chaining. Further, the part (b) implies that earlier deciding of renewable price will put the RPS in the disadvantageous position which will lead to earning of less profit compared to the later one.

4 Simulation Results and Analysis

In this section, we will study how the RPSs allocate its energy storage and deciding energy price based on the scheme service price by the operator. We also examine the optimal r share of the revenue chosen by the operator at the situation of simultaneous and sequential renewable energy price decisions of RPSs.

4.1 Parameter Setting

We assume the service chaining contains eight hops and the distance between relay nodes is set to be a random value of 100–900 m. The renewable energy storage cost of each RPS has a random value between 1 and 10 cents/kWh and the energy storing and transmitting efficiencies have a value of $\rho = 0.89$, $\rho' = 0.86$, respectively. For the path loss model through the data transmission link, we set $\gamma = 2.3$ and the value of A is assumed to be 12 dB/km. Further, the price-sensitive exponent b of traffic demand has a value of 11. Also, we set the carrier frequency f at 25 GHz and basic data transmitting rate is set to 100 Mbps.

4.2 Performance Analysis

We first study the equilibrium of renewable energy price and the energy storage level of each RPS. In Fig. 3, we show the equilibrium renewable energy of each RPS. Figure 4 shows that the profit of operator performances better off with the number of hops increasing, and however, the operator earns more profit under the sequential decisions than simultaneous settings at the same number of hops.

Tables 2 and 3 imply that the performance of the two channels with a bandwidth of 150 MHz and 100 MHz. Compared with Table 3, Table 2 shows that in equilibrium a larger bandwidth will lead to a higher service price and then the

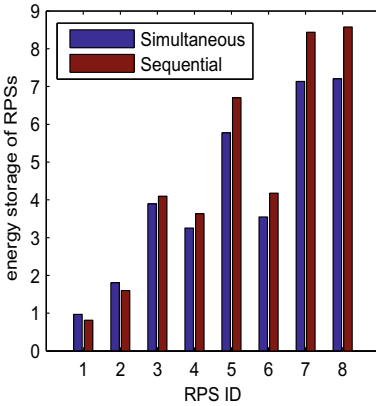


Fig. 3. The equilibrium renewable energy storage level of each RPS.

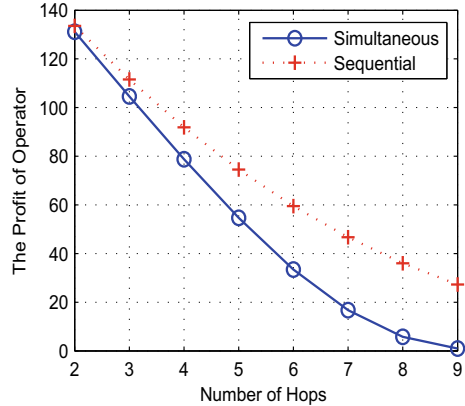


Fig. 4. The profit of operator under multi-hops.

Table 2. The profit of operator with a bandwidth of 150 MHz

Strategy	Price	Traffic (bits)	Profit
Simultaneous	14.6910	4845.3	6.4520
Sequential	14.6860	4863.4	38.0547

Table 3. The profit of operator with a bandwidth of 100 MHz

Strategy	Price	Traffic (bits)	Profit
Simultaneous	11.8380	5209.7	5.7976
Sequential	11.7726	5537.2	36.0213

data traffic will be cut off for the traffic demand model. However, the operator earns more profit when the bandwidth is increased. These two tables have further proved that the strategy of sequential settings performs better than the simultaneous decisions.

5 Conclusion

In this paper, we have studied edge-computing-based data traffic aggregation network service chaining with multi-hops which is powered by the independent RPSs. In the studied model, we use backward induction to specify the unique Pareto-optimal equilibrium storage strategies of the RPSs under the price scheme p chosen by the operator. Further, we studied the strategies of RPSs under simultaneous and sequential settings and analysis of the performance in decentralized solution. Simulation results reveal that in equilibrium the sequential decisions have better performance and earn more profit for the operator. We also observe that the equilibrium profit of the operator decreases with the increase of the hops and increases with the bandwidth through the chaining.

Acknowledgments. This work was supported in part by the NSF of Jiangsu Province under Grant BK20161518 and Grant BK20171444, in part by the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University, under Grant 2018D05, in part by the National Natural Science Foundation of China under Grant 61772287, Grant 61771252, and in part by the Open Research Fund of Jiangsu Engineering Research Center of Communication and Network Technology, NJUPT.

References

1. Cisco White Paper: Global Mobile Data Traffic Forecast Update, 2016–2021, Cisco Visual Networking Index, pp. 1–5, Feb 2017
2. Dastjerdi, A., Buyya, R.: Fog Computing: Helping the Internet of Things Realize its Potential. IEEE Computer Society, pp. 112–116 (2016)
3. Chen, X., Shi, Q., Yang, L., Xu, J.: Thriftyedge: resource-efficient edge computing for intelligent IoT applications. IEEE Netw. 61–64, Jan 2018
4. Chang, Y., Chen, S., Wang, T., Lee, Y.: Fog computing node system software architecture and potential applications for NB-IoT industry. In: IEEE International Computer Symposium, pp. 727–730 (2016)
5. Mao, Y., Zhang, J., Song, S.H., et al.: Power-delay tradeoff in multi-user mobile-edge computing systems. In: IEEE Global Communications Conference, pp. 1–6, Dec 2016

6. Kiani, A., Ansari, N.: Edge computing aware NOMA for 5G networks. *IEEE Internet Things J.* 1–7 (2017)
7. Fawal, A., Mansour, A., Roy, F., Jeune, D., Hamie, A.: RACH overload congestion mechanism for M2M communication in LTE-a: issues and approaches. In: *International Symposium on Networks*, pp. 1–6, May 2017



Design and Implementation of Sensory Data Collection and Storage Based on Hadoop Platform

Zhen Bai¹(✉), Shaohua Cui², and Chenglin Zhao¹

¹ Key Laboratory of Universal Wireless Communications, MOE, Beijing University of Posts and Telecommunications, Xitucheng Road 10, Beijing, China
baizhenbupt@163.com

² China Petroleum Technology & Development Corporation, Beijing, China

Abstract. At present, modern manufacturing and management concepts such as digitization, networking, and intelligence are widely used in the industry. Industry automation and information have been unprecedentedly improved, and therefore the entire life of industrial production link involves massive amounts of data, and the status monitoring data of industrial machine have large, multiple source, heterogeneous, and complex data characteristics. What is more, the traditional processing methods and tools could not meet the requirements for massive data, and may miss the best time to repair machine. So, to resolve the challenges that the industrial sensory big data faces, this paper proposes the sensory data collection and storage based on Hadoop platform.

Keywords: Hadoop platform · Sensory data · Data collection · Data storage

1 Introduction

According to the statistics of McKinsey, an internationally renowned consulting company, the amount of data stored in the manufacturing industry far exceeds the total amount of data in other industries [1]. In fact, during industrial production process, the machinery and equipment are in continuous operation, especially sensor networks and smart devices that continuously monitor equipment and collect data. As time goes on, the data size, which is collected by sensors, has reached hundreds of terabyte (TB), and even petabyte (PB) [2–4]. So, before data mining, we need to deal with the problem of how to efficiently collect and store large amount of industrial semi-structured data.

The traditional sensory data management and analysis platforms are all based on data warehouse technology, but data warehouse technology has some defects that maybe could not meet the requirements of large amounts of sensory data [5, 6]. In order to deal this situation, in this paper, we use the real-time sensory data storage solution based on Hadoop developed by the author as an example to summarize and propose a solution for collecting and storing big data for semi-structured data source.

2 Data Collection and Storage

2.1 Real-Time Sensory Data

The sensory data used in this article comes from the following sensors: temperature sensor, oil pressure sensor, pressure sensor, acceleration sensor, and three-phase electric sensor formed five collection types. Real-time sensor data is in JSON format, and uploaded every few seconds. It mainly contains several significant values, such as the time sensor uploads data, the value sensor collects real-time status of sensor, and so on.

2.2 Data Collection

The data collected by sensors is semi-structured data. For later data processing and analysis, we need to extract and integrate widely heterogeneous data and the results are stored according to uniform standards, so that we can easily extract useful information from it and present the results in an appropriate way. So, collection of these data is to process semi-structured data, so that they can meet the quantifiable standards of storage data, and require that these processes are scheduled automatically.

To process semi-structured data, the combination of most programming languages and databases can be completed very well. The question is which communication protocol we choose. There are many ways for sensor to access server, such as TCP, UDP, MQTT, and HTTP. We also notice that the real-time sensory data is in JSON format, so we choose HTTP as communication protocol, this is because it can reduce development costs and improve reliability. We can place real-time sensory data into entity body, and when server receives the request, it continues to parse JSON and deal with data.

2.3 Data Storage

In reality, there are multiple collection stations at the production site, and the data related to production monitoring generated during data collection every day is aggregated into data oceans. The existing database technology is difficult to efficiently store and manage large amounts of historical data. So, we propose a multi-table architecture based on MYSQL combined with HDFS.

In order to achieve multiple table architecture, we use sharding technology. It means we put data into two or more separate tables based on certain condition. That is split by record, different records can be saved separately, each sub-table contains the same number of columns, but fewer rows of data. In this article, we split table based on the following two conditions. First, we split sensory data according to sensory type, and this avoids storing all types of sensory data on one table. Second, the other condition is based on time. For example, a table containing billions of records can be divided into several tables according to month, and each small table contains data representing a specific period of time. When you need to query the data of a specific year or month, you only need to operate on the corresponding table.

But with the continuous increase of data, relational databases have gradually exposed some deficiencies in the face of large amounts of historical data storage and access issues. Therefore, this paper uses the multi-table method of traditional database to realize the real-time data storage, real-time data operation, and query display. In addition, cumulative sensory data is transferred to HDFS for historical preservation using data migration technology. This not only solves the problem of storing large amounts of data, but also enables deep analysis and processing of historical data through the use of big data technology. It maximizes the use of sensory data.

2.4 Architecture

According to the above three subsections, we propose an overall framework for semi-structured sensory data solution. As shown in Fig. 1, the solution is that data collection and storage is writing semi-structured data into MYSQL in a unified data format, then regularly migrating data to HDFS for big data analysis. The communication of components is mainly through HTTP protocol.

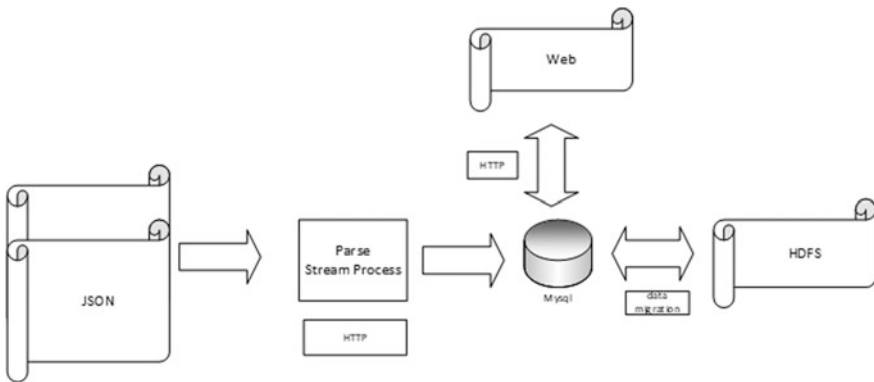


Fig. 1. The architecture of data collection and storage

3 Hadoop Platform

As described above, we know that the big data has some significant characteristics, such as volume, variety, velocity, and veracity [7]. Due to the limits of the traditional platform, it could not reach the requirement of big data. But through Hadoop platform, it makes sure that the platform can run in a complex environment, and can provide a safer, more reliable, and more effective platform [8]. The Hadoop platform’s architecture is just shown in Fig. 2. As shown in Fig. 2, the Hadoop platform is divided into five parts as given below:

- Data source: It is responsible for the source of the platform data, as described in the previous section, the data mainly comes from the migration of historical data in the database. And, the real-time data is mainly used to analyze the status of equipment.

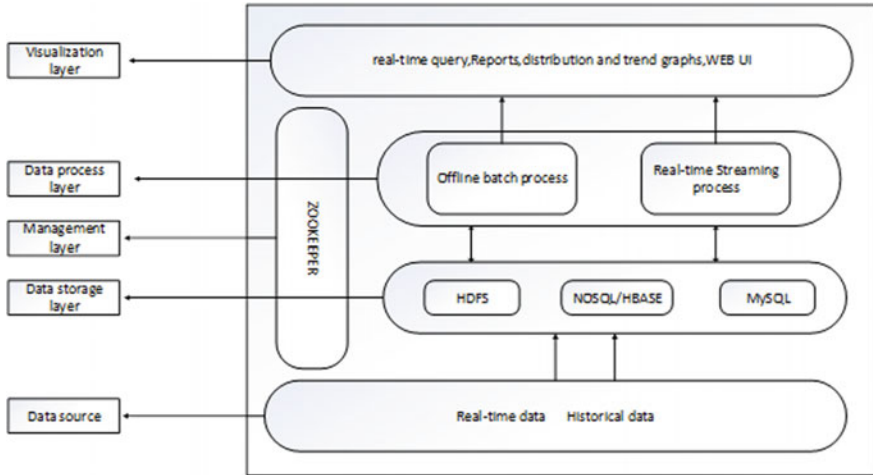


Fig. 2. The architecture of Hadoop platform

- **Data storage layer:** It is responsible for historical storage and real-time data storage, historical data is used for big data analysis, so it is stored in HDFS historical data. In addition, it adopts the master–slave structure, using three servers to build Hadoop cluster, one for master server and the other two for slave server, so that it could improve the efficiency of data storing in each node.
- **Data process layer:** This layer is mainly used for data analysis, including off-line data batch processing and real-time data stream processing. Off-line data batch processing is core part of system, and is mainly responsible for machine learning and data mining. It mainly uses MapReduce and SPARK to implement batch processing. Stream processing is mainly used for the analysis of real-time data.
- **Management layer:** It mainly uses Zookeeper as a distributed coordination tool to coordinate services and manage Hadoop cluster.
- **Visualization layer:** It is responsible for periodically storing the results of the analysis in a database, and receiving the requests submitted by users, and then providing an intuitive way to display data.

4 Conclusion

This paper has primarily focused on the sensory data collection and storage, so we design and implement the sensory data collection and storage solution based on Hadoop platform. First, it provides an interface, which is used for collection of real-time data from sensor and formatting semi-structured data. Second, since the system not only needs real-time query, but also needs to analyze and process historical data, so we propose a multi-table architecture based on MYSQL combined with HDFS, which uses traditional multi-table method of database to realize real-time operation and query display and uses data migration technology to implement big data analysis on HDFS.

In the end, the system uses ECharts to realize the visual analysis and provides a friendly web interface for users. Meanwhile, the system has some significant advantages such as mass storage, fast retrieval and query, reliable data management service, and high-performance processing. In conclusion, it provides an effective solution to collect and store real-time sensory data.

Acknowledgements. This work was funded by the National Intelligent Manufacturing Project.

References

1. Liu, F.C., Shen, F.: Building a research data science platform from industrial machines. In: IEEE Conference (2015)
2. Manwal, M., Gupta, A.: Big Data and Hadoop A Technological Survey, Emerging Trends in Computing and Communication Technologies (ICETCCT), February (2018)
3. Lyu, Y., Fan, X., Liu, K.: An optimized strategy for small files storing and accessing in HDFS. In: IEEE International Conference on Computational Science and Engineering (CSE), vol. 1, pp. 611–614 (2017)
4. Sarnovsky, M., Bajus, D.: Building environment analysis based on clustering methods from sensory data on top of the Hadoop platform. In: IEEE Conference (2017)
5. Xie, J.: Construction for the city taxi trajectory data analysis system by Hadoop platform. In: IEEE Conference (2017)
6. Zhonghua, M.: Seismic data attribute extraction based on Hadoop platform. In: IEEE Conference (2017)
7. Longley, P.A., Goodchild, M.F., Maguire, D.J., Rhind, D.W.: Geographic Information Systems and Science. Wiley, New York (2001)
8. Mazumdar, S., Dhar, S.: Hadoop as big data operating system—the emerging approach for managing challenges of enterprise big data platform. In: IEEE Conference (2015)
9. Ding, X., Tian, B.: A scheme of structured data compression and query on Hadoop platform. In: IEEE Conference (2015)
10. Yan, H., Song, G.: Based on the Hadoop cloud computing experiment platform. J. Shenyang Norm. Univ. (Natural Sciences) **1**, 85–89 (2013)



Design of a Data Acquisition and Transmission System for Smart Factory Based on NB-IoT

Ruijian Zhang¹(✉), Shaohua Cui², and Chenglin Zhao¹

¹ Key Laboratory of Universal Wireless Communications, MOE, Beijing University of Posts and Telecommunications, No. 10 Xitucheng Road, Beijing, China

zhangruijian1994@foxmail.com

² China Petroleum Technology & Development Corporation, Beijing, China

Abstract. In the industrial network, the trend of combining different communication systems has been observed in the past years and it will continue in the future. Nowadays, with the rise of Wireless Sensor Network (WSN) and Low-Power Wide-Area Network (LPWAN), wireless technology is gradually infiltrating into the field of industrial control. This paper proposes an industrial data acquisition and transmission system using ATMEGA328P microcontroller based on Narrowband Internet of Things (NB-IoT) technology. System realizes the integration of heterogeneous networks and seamless acquisition and transmission of data, and enables remote real-time monitor to the sensor nodes distributed in the smart factory.

Keywords: Industrial network · WSN · LPWAN · NB-IoT · Smart factory

1 Introduction

At present, industrial equipment mainly uses cable communication. The traditional communication methods have the disadvantages of being difficult to assemble, poor anti-interference performance, and the communication efficiency is low [1]. Besides, the cables are exposed to harsh industrial environment, so they require regular maintenance and replacement. With the rapid development of wireless sensor network [2], wireless sensors have broad application prospects in industrial automation, daily life, environmental monitoring, military, medical care, and other fields [3, 4]. WSN has advantages in flexibility, convenience, and stability. Sensors acquire and send the information through wireless communication technologies, such as Wi-Fi, Bluetooth, and ZigBee [5]. Low-power wide-area network has advantages in wide network coverage range and low power consumption of the terminal, so it is befitting for large-scale deployment of Internet of Things applications. The overall structure of system is shown in Fig. 1.

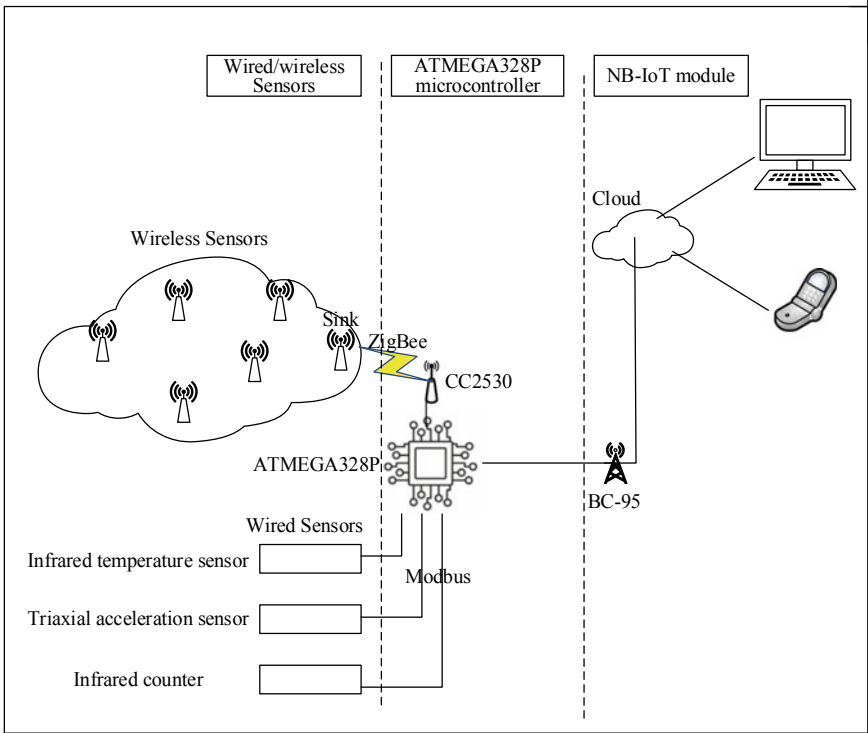


Fig. 1. Structure of system

The wireless sensors mainly monitor the temperature and humidity of smart factory, and send data through ZigBee [6] technology. The wired sensors collect information during the production process, including the temperature of machine, the vibration of machine, and the production quantity. The wired sensors send data following Modbus fieldbus protocol. Microcontroller reads and encapsulates all the data and transmits it to the server through NB-IoT module.

The Modbus fieldbus protocol was designed and developed by Modicon, Schneider’s subsidiary, in 1979. It is an application-layer bus standard that is widely used in the field of industrial automation [7].

ZigBee is a wireless communication technology based on the IEEE 802.15.4 standard [8]. ZigBee technology is widely used in many fields including smart home, building automation, industrial automation, and automatic metering.

NB-IoT is an emerging technology of the Internet of Things that supports low-power device cellular data connections over wide-area networks, also known as low-power wide-area networks [9, 10]. NB-IoT supports efficient connection of devices with long standby time and high network connection requirements [11].

2 System Design

2.1 Hardware Design

The system is designed to monitor data in the production line as shown in Fig. 2; the infrared temperature sensor MLX90614 measures the temperature of the slide rails of the hydraulic punching machine, the triaxial acceleration sensor WT61C is used to measure the vibration of the punching machine, and the infrared counter QT50CM counts the production quantity.

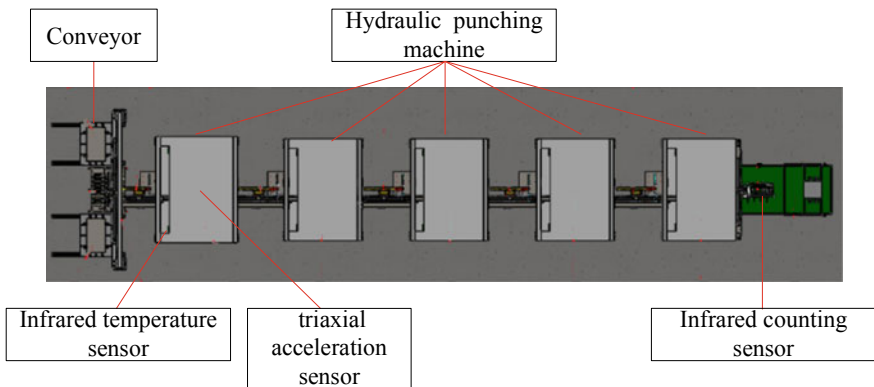


Fig. 2. Production line of the smart factory

The CC2530 is a true System-on-Chip (SoC) solution for IEEE 802.15.4 with the frequency of 2.4 GHz and a capacity of 250 Kbps. It enables building of robust network nodes with very low total Bill-of-Material (BOM) costs. These applications can easily be built on top of the IEEE 802.15.4 based standard protocols or on top of the proprietary SimpliciTI network protocol. We use DHT11 sensor to detect ambient temperature and humidity, and microcontroller communicates with CC2530 through serial port.

ATMEGA328P is a microcontroller with an open-source electronic prototype platform that is flexible and easy to use. The model of NB-IoT module is LTE BC95-B5, which is produced by Quectel company.

2.2 Software Design

CC2530 makes up communication network following protocol Z-Stack provided by TI. The end devices collect ambient temperature and humidity data and send it to the sink node. Then, MCU reads the data in the sink node regularly through the serial port. Simultaneously, the MCU stores the data collected by the wired sensor into the buffer. MCU transfers data into an array of string type following a certain order. Then, MCU converts the array to hexadecimal format and sends it to BC95 through serial port.

BC95 starts with a self-test, it uses AT commands to detect current status such as signal strength, status of SIM card, and network registration status. Then, BC95 sends the data to the server. In this paper, the server IP is 47.93.98.105:22.

BC95 automatically converts the data to ASCII format and sends it to the server following UDP protocol. The server reads UDP package and converts the data back to decimal format, and then saves it in a log file. Administrators can remotely read the contents of log files on server to remotely monitor plant data.

3 Experiment and Verification

First, waiting for CC2530 modules to complete networking and the NB-IoT to complete self-testing. Then, the system starts acquiring and transmitting data. The MCU stores and sorts the data collected by sensors, and prints the results through serial port. We can monitor this process locally on computer through serial monitor, as shown below.

Simultaneously, the NB-IoT module sends data to the server. As shown in Fig. 3, server records the date, time, and resource IP address of every data. The initial data is automatically converted to ASCII format by BC95, and the server saves all the details into a log file (Fig. 4).

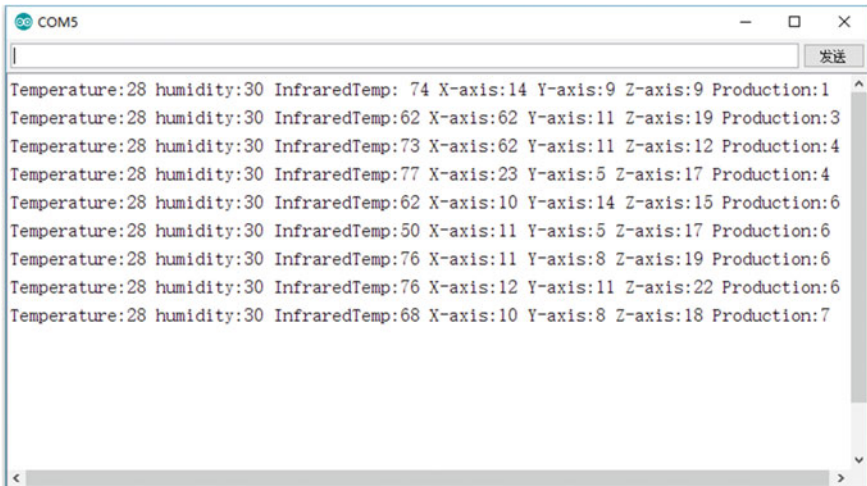


Fig. 3. MCU prints data through the serial monitor

```
2018-05-22 15:46:12 Udp Receive Data: ^\^^J^N ^A from 117.61.17.72
28 30 74 14 9 9 1

2018-05-22 15:47:00 Udp Receive Data: ^\^^>>^K^S^C from 117.61.17.72
28 30 62 62 11 19 3

2018-05-22 15:47:48 Udp Receive Data: ^\^^I>^K^L^D from 117.61.17.72
28 30 73 62 11 12 4

2018-05-22 15:48:35 Udp Receive Data: ^\^^M^W^E^Q^D from 117.61.17.72
28 30 77 23 5 17 4

2018-05-22 15:49:23 Udp Receive Data: ^\^^>^N^0^F from 117.61.17.72
28 30 62 10 14 15 6

2018-05-22 15:50:10 Udp Receive Data: ^\^^2^K^E^Q^F from 117.61.17.72
28 30 50 11 5 17 6

2018-05-22 15:50:58 Udp Receive Data: ^\^^L^K^H^S^F from 117.61.17.72
28 30 76 11 8 19 6

2018-05-22 15:51:45 Udp Receive Data: ^\^^L^L^K^V^F from 117.61.17.72
28 30 76 12 11 22 6

2018-05-22 15:52:33 Udp Receive Data: ^\^^D^H^R^G from 117.61.17.72
28 30 68 10 8 18 7
```

Fig. 4. Data received by the server

4 Conclusion

This paper introduces an industrial wired/wireless hybrid data acquisition system using ATMEGA328P microcontroller based on ZigBee and NB-IoT technology. Microcontroller works as the gateway of the system, which integrates the heterogeneous networks. System achieves seamless acquisition and transmission of data, and enables remote real-time monitor to the sensor nodes distributed in the smart factory.

Acknowledgements. This work is supported by Project of Intelligent Manufacturing Integrated Standardization and New Model Application 2017ZNZX03.

References

1. Lee, K.C., Lee, S.: Performance evaluation of switched Ethernet for real-time industrial communications. *Comput. Stand. Interfaces* **24**(5), 411–423 (2002)
2. Willig, A.: Recent and emerging topics in wireless industrial communications: a selection. *IEEE Trans. Industr. Inf.* **4**(2), 102–124 (2008)
3. Khemapech, I., Duncan, I., Miller, A.: A survey of wireless sensor networks technology. In: 6th Annual Postgraduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting, vol. 13 (2005, June)
4. Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E.: Wireless sensor networks: a survey. *Comput. Netw.* **38**(4), 393–422 (2002)

5. Pothuganti, K., Chitneni, A.: A comparative study of wireless protocols: bluetooth, UWB, ZigBee, and Wi-Fi. *Adv. Electron. Electr. Eng.* **4**(6), 655–662 (2014)
6. Baronti, P., Pillai, P., Chook, V.W., Chessa, S., Gotta, A., Hu, Y.F.: Wireless sensor networks: a survey on the state of the art and the 802.15. 4 and ZigBee standards. *Comput. Commun.* **30**(7), 1655–1695 (2007)
7. Yanfei, L., Cheng, W., Chengbo, Y., Xiaojun, Q.: Research on Zigbee wireless sensors network based on modbus protocol. In: *IFITA'09. International Forum on Information Technology and Applications*, 2009, vol. 1, pp. 487–490. IEEE (2009, May)
8. Galkin, P.: Interaction model design of ZigBee-gateway between wireless sensor network and industrial network. In: *Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)*, 2017 4th International, pp. 501–504. IEEE (2017, October)
9. Sinha, R.S., Wei, Y., Hwang, S.H.: A survey on LPWA technology: LoRa and NB-IoT. *ICT Express* **3**(1), 14–21 (2017)
10. Mahmoud, M.S., Mohamad, A.A.: A study of efficient power consumption wireless communication techniques/modules for internet of things (IoT) applications. *Adv. Internet Things* **6**(02), 19 (2016)
11. Petrenko, A.S., Petrenko, S.A., Makoveichuk, K.A., Chetyrbok, P.V.: The IIoT/IoT device control model based on narrow-band IoT (NB-IoT). In: *Young Researchers in Electrical and Electronic Engineering (EIConRus)*, 2018 IEEE Conference of Russian, pp. 950–953. IEEE (2018, January)



Design of Laboratory Monitoring Management System Based on Internet of Things

Nannan Chong^(✉)

Tianjin University Renai College, Tianjin, China
chongnannan@163.com

Abstract. This paper designs an indoor location and environment information testing laboratory monitoring management system. The system includes four parts: a development board based on STM32F103RET6 provided by China Mobile, W5100 Internet module, the Wi-Fi-RFID-WSN sensing part, data server, and Android client. Users can locate and monitor assets such as apparatus equipment of labs, monitor environmental information, make it convenient for both self-help studying or research and management, etc. This system has low power consumption, good scalability, and client interface simple operation, which can automatically identify where things with RFID tags, real-time query laboratory environment information, through the client application user can set up the environment parameters threshold, and start the transfinite alarm. Moreover, testing results of the effectiveness of this system are proposed in the end.

Keywords: Indoor localization · Laboratory monitoring management system · Internet of things

1 Introduction

As the promotion of the Internet of Things technology and the development of the intelligent hardware, the demand for intelligent laboratory management system is growing with each passing day. A large amount of researchers is working on designing intelligent monitoring management system applicable to various scenarios. The authors in [1] designed an integrated module of ultrasonic medical laboratory management system, in which each module can interact with relational databases, be organized by the Internet, realize the business scheduling, data record, referral diagnosis, information security, privacy of configuration, image management, and other functions. A laboratory consumable inventory management system is designed in [2], which includes middleware, processor, and database module, and also can update inventory information and consumable quantity. In [3], a laboratory information management system is designed for mining and metal industry, the system is easy to access and manage through the models of user trip graph, requirement specification list, database, and system interface. Some scholars design to realize the integrated management information system for environmental monitoring lab, which can standard management resources such as implement sample, acquisition task review, sample selection, sample scanning, data analysis, coding rules, quality control, and so forth [4]. In [5], authors

promote a combining online learning and laboratory information management system suitable for water hydraulic, the system uses the model B/S (browser/server) structure to control, which can improve the efficiency of teaching and learning. [6] promoted an information management system of computer lab by which the management of experiment operation is optimized. An application to the information management system of the environmental monitoring laboratory is designed in [7] that authors focus on monitoring the information of environmental flow parameters. From the current situation, there is a lack of positioning monitoring and management laboratory system, so we focus on this application technology gap and carry out related research.

With the rapid development of indoor positioning technology, the location of the service under the indoor environment can be used in the laboratory, library, supermarket navigation, elderly care, and the Internet of things, in the scene of which accurate, economic positioning is particularly important. To meet the demand of university laboratory teaching management, this paper promotes a laboratory management system based on STM32F103RET6 (Laboratory Monitoring Management System (LMMS)), suitable for Electrical, Chemical, Biology, and other professional laboratory or library for information management and equipment monitoring, which can realize precise localization of teaching instruments which is based on Wireless Fidelity (Wi-Fi), Radio Frequency Identification (RFID), and Wireless Sensor Networks (WSN). This system consists of STM32F103RET6 provided by China Mobile, W5100 Internet module, the Wi-Fi-RFID-WSN sensing part, data server and Android client, apparatus and equipment, achieve convenient self-help study and research, the real-time environmental information monitoring, and security of personnel and asset. Generally speaking, the system can provide convenience for teachers, students, and administrators.

The remainder of this paper is organized as follows. In Sect. 2, we will introduce all parts of LMMS and details of main technologies. Testing and analysis of both hardware and Android client will be discussed in Sect. 3. In Sect. 4, we will provide the conclusion and plans of next steps.

2 LMMS Structure

The positioning technology of LMMS system is based on Wi-Fi-RFID-WSN which is more suitable than the other indoor positioning technology for the laboratory regulation scenarios. The most significant character of LMMS is that the fully integrated RFID reader and WSN beacon are in a separate node, and sensor network of more than one node is responsible for collecting information from RFID and sensor and then uploading in LMMS server. Besides, LMMS can be easily configured in the traditional experimental teaching environment and can increase the availability without improving the network structure complexity of the system. All measured items in LMMS system (such as equipment information tools, etc.) are attached with RFID Tag, and important rooms are equipped with Wi-Fi module RFID read-write device that can identify the room and all items with WSN beacon. The location and status information of each item will be tracked by the RFID system according to the coverage of the router in every floor and ensure that all RFID readers are covered by Wi-Fi signals. As a WSN, the

experimental indoor configuration sensor group can collect environmental information such as temperature, humidity, light intensity, PM2.5, and other environmental information. The information data are uploaded via the RFID gateway and WSN gateway by Wi-Fi to the cloud or data server. Administrators can get all items with RFID positioning through the PC or Android client (precise to the floor of the classroom location), view real-time laboratory environmental information and configure security threshold, and realize the monitoring and management of laboratory safety efficiently. Main structure of LMMS is shown in Fig. 1.

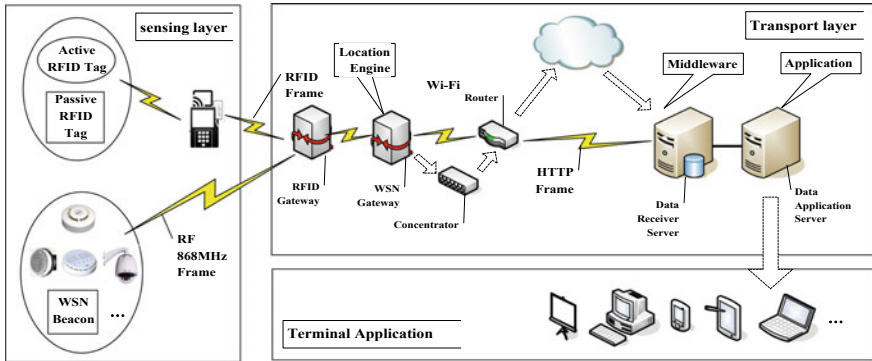


Fig. 1. Structure of LMMS

2.1 STM32F103RET6 and W5100 Internet

LMMS system is designed based on STM32F103RET6 which uses ARM®Cortex™-M3 MCU and is provided by China Mobile enhanced development board, which relies on complete TCP/IP protocol processing functions of W5100, and connected Wi-Fi module with RFID Reader module. The core connection pins between them are shown in Table 1.

Table 1. The connection pins between STM32F103RET6, Wi-Fi, and ATR03 RFID reader

Module	Source	Ground	Core pins connection			
STM32F103RET6	3.3 V	GND	USART2-TX	USART2-RX	USART3-TX	USART3-RX
Wi-Fi			J4-RX	J5-TX		
RFID reader					RX	TX

2.2 Wi-Fi

Based on the 802.11ac standard, Wi-Fi can be used to locate the location of the mobile device with Wi-Fi communication module through receiving the wireless network signal strength (RSSI) on the site. In practice, the positioning accuracy of large indoor

space can be improved by increasing the number of Wi-Fi hotspots, which can be used for complex indoor scenes and hot spots. However, since Wi-Fi positioning accuracy can theoretically reach 3–5 m, the actual accuracy is about 10 m [8], so we use Wi-Fi to make initial rough positioning, and if the objects need to be precisely located, the RFID module can be used for it. We use the fingerprint algorithm based on RSSI [9, 10]. There are many advantages of the fingerprinting approach [11, 12], including the fact that no special hardware is required on the user mobile station (MS) side [13, 14]. Wi-Fi location follows two steps which are below:

Offline training: Divide the indoor area into different parts and make it as a digital positioning network, establish a sampling point (spacing of 3 m), sample the locate point by Wi-Fi receiver, record the point position and obtain the RSSI and access point (AP) address, and then process the sampling data (filtering, average, etc.).

Online positioning: The user can hold the mobile device, move in the location region, get the real-time RSSI and AP address, and then upload the information to the server for matching. Finally, one can get the estimated position by using KNN algorithm.

We denote $A_n (n = 1 \dots N)$ as the n-th AP, so the fingerprint of the i-th reference point (RP) can be defined as

$$R_i = \begin{bmatrix} P_{A_1}(T_1) & P_{A_2}(T_1) & \dots & P_{A_N}(T_1) \\ P_{A_1}(T_2) & P_{A_2}(T_2) & \dots & P_{A_N}(T_2) \\ \vdots & \vdots & \vdots & \vdots \\ P_{A_1}(T_M) & P_{A_2}(T_M) & \dots & P_{A_N}(T_M) \end{bmatrix} \tag{1}$$

T is the strength of received signals, and P can be defined as

$$P_{A_n}(T_m) = \frac{C_{T_m}}{N_i} \tag{2}$$

N_i denotes the number of the training sequence and C_{T_m} means the number of T_m in RP training sequence. So the total fingerprint amount can be expressed as

$$D = [R_1, R_2, \dots, R_w] \tag{3}$$

w is the amount of RP in covered area.

In order to improve the operation speed, the signal strength distribution is divided into P pieces. The i-th RP fingerprint could be expressed as

$$R_i = \begin{bmatrix} P_{A_1}(B_1) & P_{A_2}(B_1) & \dots & P_{A_N}(B_1) \\ P_{A_1}(B_2) & P_{A_2}(B_2) & \dots & P_{A_N}(B_2) \\ \vdots & \vdots & \vdots & \vdots \\ P_{A_1}(B_M) & P_{A_2}(B_M) & \dots & P_{A_N}(B_M) \end{bmatrix} \tag{4}$$

The distribution of the B_k RSS of the i -th RP AP is given as

$$P_{A_n}(B_m) = \frac{C}{N_i} \tag{5}$$

where C is the sample number of B_k signal strength.

2.3 RFID-WSN Integration

LMMS system innovation is to integrate RFID reader and WSN. There are four main models of fusion technology: sensors-tag model, the WSN-tag model, WSN-reader model, and WSN-RFID model which are shown in Fig. 2. In LMMS, we use WSN +RFID model, as the hardware is independent, and data analysis and system collaboration are carried out on the software level.

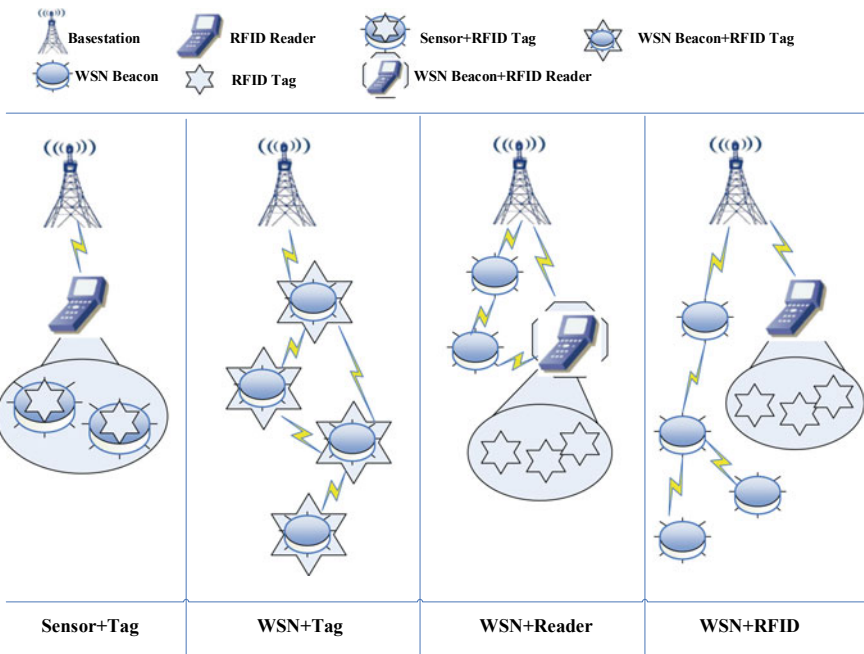


Fig. 2. Mode of integration of RFID and WSN

We designed the RFID gateway to identify the personnel and property within the scope of the recognizer and obtain the data report to WSN. As shown in Fig. 3, the main structures are given as follows:

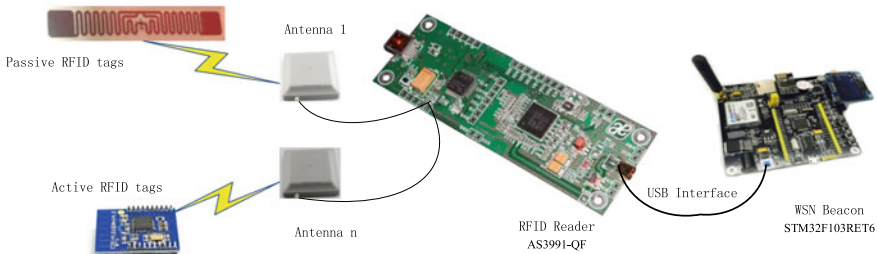


Fig. 3. RFID gateway structure

Antenna: RFID antenna produces electromagnetic fields to provide drivers for passive WSN beacon recognition, and collect the label data.

UHF RFID front end: It converts binary digital information into RF signal to transmission. To deal with the diversity of labels, the Keonn AdvanReader-150 identifies the EPC codes of the tag, and if the important part is encoded as 0, it is the label of the EPC Gen2 process. If an important part is encoded as 1, a sensor is found on the label, and the reader enters the access stage after identification and retrieves the information of the sensor.

Microcomputer: It is responsible for management and pretreatment of the received information. The Python application is developed in the Linux system to execute the query protocol, and the RFID data is transmitted through the serial port to the WSN beacon node. Python applications continue to ask whether RFID readers have identified new tags. If identified, the information is sent to the WSN beacon node and confirmed; if the message is not recognized or timed out, the message is sent again. Finally, RFID data and various sensor information are uploaded to WSN beacon nodes through WSN.

Beacon node: It is used to connect to a microcomputer by USB, and responsible for transferring the collection information from the WSN beacon to WSN. The Waspnote 868 MHz beacon node is similar to other beacon nodes in the LMMS system, which can accept, process, and resend data from the sensor.

Power supply: Power supply or batteries.

RFID gateway of LMMS improves the laboratory management and efficiency by testing instruments and equipment and environmental information. Because RFID and WSN packets are integrated in the gateway, all RFID monitoring is transparent to WSN. RFID gateway improves the system performance. Each tag can be used to track the equipment, and the cost of WSN beacon is only a few cents/euro, and users can choose the passive RFID monitor inventory without batteries.

2.4 Data Service

Upload data to the data receiving server in Linux system through WSN gateway and Wi-Fi connection. The control program files of all modules in LMMS are stored in the server root directory, and SQL server database storage information is collected by terminal. Create dedicated VLAN on existing wireless devices. Use WPA personal protocol for secured communication, and each wireless network device uses a 256-bit

key encryption, and the HTTP protocol will upload the data encapsulated in JSON format to the client server. In addition, to ensure that the entire network is synchronized, the WSN gateway periodically takes a timestamp from the HTTP server and passes it to the rest of the WSN beacon nodes.

2.5 Android Client

Android client is comprised of Android Studio development, responsible for the WSN gateway which receives data in the database, and displays it in the user interface, and it also can display the different floors of the user and the location of the assets and more other information. In addition, the platform also manages incoming alerts (for example, the instrument does not exceed the default threshold in the specified area or environment parameter), and the pre-established protocol is started. Users with permission can configure most of the system parameters on the client side (such as registering new users or adding assets, setting up alarm thresholds, defining the RFID region, and so on).

In conclusion, the function structure of LMMS system is shown in Fig. 4.

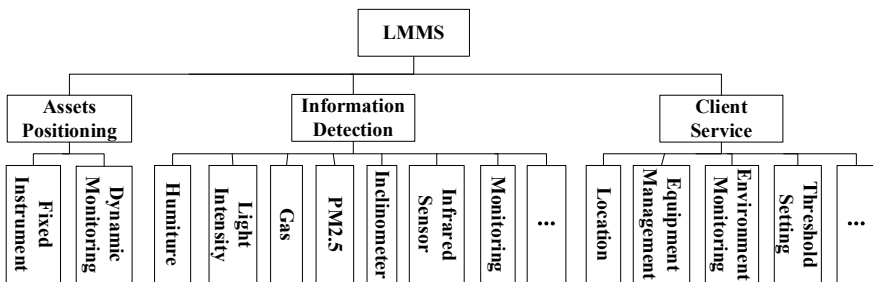


Fig. 4. Function structure of LMMS

3 Testing

3.1 Location Testing

To test system positioning effect, floor of placing a gateway and three reference node, tag location test four groups, all sign on the floor in the floor plan, and test results as shown in Fig. 5. The system positioning error is within 2 m, and the average error is about 1.3 m, which is sufficient to find the required equipment or teaching devices. The results are accurate enough for the laboratory or library to verify the feasibility of the system.

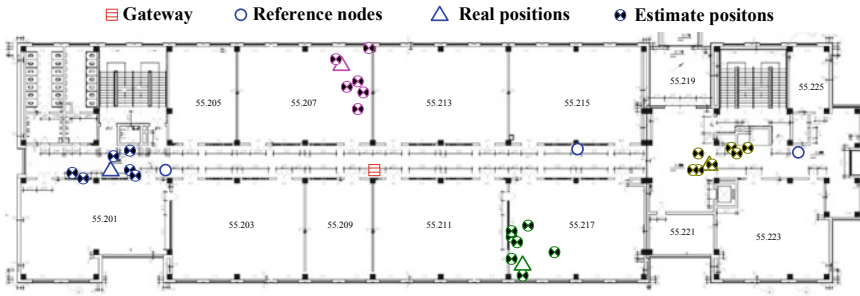


Fig. 5. Location testing

3.2 Android Client Testing

Open the LMMS app under Android system, set the serial port terminal, and configure the serial port to COM1. After the client application starts, it will automatically realize communication through Wi-Fi and collect information of temperature, humidity, lighting, and monitoring video. As shown in Fig. 6, things equipped with RFID tags can be located, environmental information and system parameter configuration can be achieved by clicking the buttons of temperature, humidity, illumination, gas, PM2.5, dip angle, infrared, and monitoring. One can add and configure new sensors through the app.



Fig. 6. Android client application interface

4 Conclusion

Based on the development platform of China mobile Internet of Things STM32F103RET6, the technology of Wi-Fi, RFID, and WSN is integrated and designed for the monitoring and management system of university laboratories.

This LMMS system has low power consumption, good scalability, by which client interface is simply operated. What is more, it can automatically identify where things with RFID tags, real-time query laboratory environment information, set up the environment parameters threshold, and start the transfinite alarm. The test system is basically the same as expected, all modules work normally, and the system is extendable. However, LMMS is still improving, and in the next step we will improve the network construction, optimize routing algorithm, and sensor integration.

Acknowledgements. The author wishes to thank the reviewers for the careful review and valuable suggestions. This work was supported by the Teaching Reform Project (2017-1-18) fund by Tianjin University Renai College. Besides, we wish to thank Yajuan Li of logistics department for the foundation structure of our Laboratory building for future tests.

References

1. Seward, J.B., Khandheria, B.K., Koerner, J.A., Wellik, T.J., Hansen, W.H., Gilman, G.: Ultrasound laboratory information management system and method. US, US 8417536 B2 (2013)
2. Barnes, K.M., Birolini, L., Williams, W.B., Khanal, D.: System and method for inventory sharing in a laboratory management system. US20180032954 (2018)
3. Laboratory information management system module for integration into manufacturing operations management system. In: 2010 IEEE International Conference on Industrial Technology, 14–17 March 2010. 978-1-4244-5697-0 (2018)
4. Zheng, W., Xu, Y.: Construction and application of integrated laboratory information management system. Shanxi Electronic Technology (2018)
5. Wang, H., Xu, H., Li, Q., Fu, Y.: Php-based collaborative education and management system for water hydraulic laboratory. *Comput. Appl. Eng. Educ.* **26**(2) (2018)
6. Zhang, Q., Li, P.: Design of laboratory management information system based on cloud computing. *China Comput. Commun.* (2018)
7. Liang, Q.: Laboratory information management system (lims) application research. *Environ. Sci. Manag.* (2018)
8. Hardyanto, W., Purwinarko, A., Adhi, M.A.: End-user satisfaction analysis on library management system unnes using technology acceptance model towards national standard of integrated library. *J. Phys. Conf. Ser.* **983**(1), 12006 (2018)
9. Bahl, P.: An in-building rf-based user location and tracking system. In: Proceedings of IEEE INFOCOM, vol. 2, pp. 775–784 (2000)
10. Youssef, M.A., Agrawala, A., Shankar, A.U.: WLAN location determination via clustering and probability distributions. In: IEEE International Conference on Pervasive Computing and Communications, pp. 143–150 (2003)
11. Kohoutek, T.K., Mautz, R., Wegner, J.D.: Fusion of building information and range imaging for autonomous location estimation in indoor environments. *Sensors* **13**(2), 2430–2446 (2013)
12. Guerrero, L.A., Vasquez, F., Ochoa, S.F.: An indoor navigation system for the visually impaired. *Sensors* **12**(6), 8236–8258 (2012)

13. Kaemarungsi, K., Krishnamurthy, P.: Properties of indoor received signal strength for WLAN location fingerprinting. In: International Conference on Mobile and Ubiquitous Systems: NETWORKING and Services, 2004. MOBIQUITOUS, pp. 14–23. IEEE (2004)
14. Kaemarungsi, K.: Distribution of WLAN received signal strength indication for indoor location determination. In: International Symposium on Wireless Pervasive Computing, pp. 6-pp. IEEE (2006)



Online Taxi-Hailing Platform Using Blockchain Technology

Chuang Ma^{1(✉)}, Dou Hu¹, Xinyi Wang¹, and Wenlong Liu^{1,2(✉)}

¹ School of Information and Communication Engineering, Dalian University of Technology, Dalian, China

machuang@mail.dlut.edu.cn, liuwl@dlut.edu.cn

² Dalian University of Technology, Innovation Park A528, Dalian, China

Abstract. With the rapid development of the Internet, more and more network technologies are used for the taxi service. At present, a number of popular online taxi-hailing platforms have emerged in the taxi market in China, which have played a significant role for the convenience of hailing taxi. The appearance of the traditional online taxi-hailing platforms has provided new ideas for resolving previous conflicts, and they also have some defects in management and profit model. To address these problems, we propose a scheme for the online taxi-hailing platform based on the blockchain technology, aiming at improving the security and fairness of traditional online taxi-hailing platform.

Keywords: Online taxi-hailing platform · Blockchain · Decentralization

1 Introduction

We focus the concentrations on providing a new framework for the traditional online taxi-hailing platform and solving the problems that may appear in it. In particular, we hope to improve the security, fairness, efficiency, and flexibility of the platform.

The “online taxi-hailing” is a relatively new concept emerging in recent years. In this mode, the customer’s demand is mainly achieved through the operation on the mobile phone, and the scene of stopping the taxi directly on the roadside will no longer exist. Online taxi-hailing platform is a smartphone application that serves as a platform for passengers to call cars and for drivers to accept orders. In this platform, passengers can easily release taxi information through mobile phones and can also communicate directly with the driver. It can greatly improve the service efficiency.

However, in the traditional online taxi-hailing platform, many problems are gradually exposed. The traditional online taxi-hailing platform uses the central agency to manage the entire operation system. During this period, several operators are involved [1]. The workload is very large and the algorithm has a large amount of time complexity and space complexity. Under such operating models, unreasonable third-party prices, market monopolies [2], and poor data security have gradually emerged [3].

Blockchain is a new technology that has emerged in recent years [4]. It is a kind of chained data structure in which data blocks are connected by time series, and is cryptographically ensured that cannot be modified. The blockchain has the features of

low computational complexity and high security [5, 6], also for removing the characteristics of complicating The new architecture proposed in this paper is based on blockchain technology. On the basis of the original, this paper decentralizes the online taxi-hailing platform to ensure the validity and security of the data. This scheme enables customers to take a taxi in a safer situation. Finally, we propose a scheme on online taxi-hailing platform based on blockchain technology.

This paper is organized as follows: Sect. 2 first introduces the general architecture and problems of the traditional online taxi-hailing platform. Section 3 begins to describe the application of blockchain technology on this platform. Section 4 describes the overall architecture of the design. The last section summarizes the previous content.

2 Traditional Online Taxi-Hailing Platform Architecture

The traditional online taxi-hailing platform modes are basically consistent, we take didi dache, a popular taxi software in Chinese market, for instance. The system architecture of didi dache can be mainly divided into four layers: data layer, service layer, access layer, and user application layer. Its architecture is shown in Fig. 1.

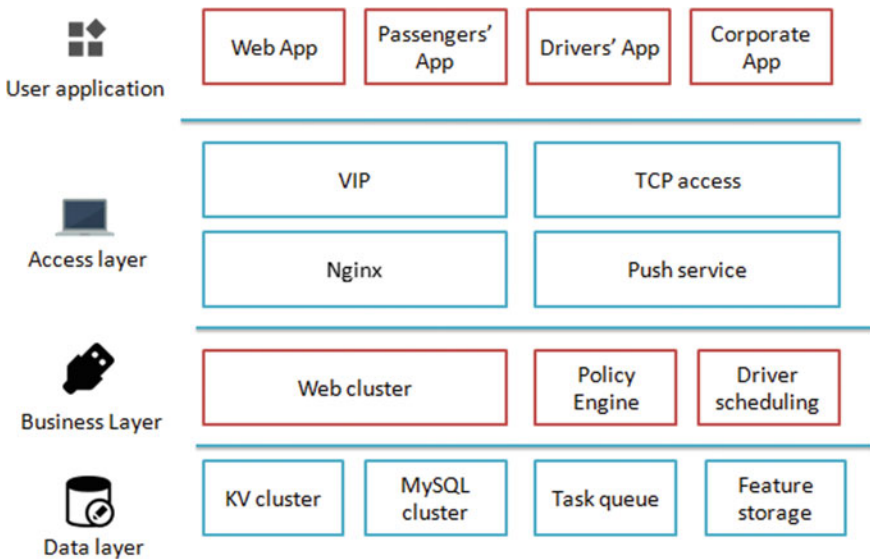


Fig. 1. The traditional system in online taxi-hailing platform

From Fig. 1, the data layer contains all the data information such as location information of vehicles and passengers, transaction data, and transaction locations. The business layer is mainly used for the arrangements and dispatch of drivers and passengers, which are listed below:

1. The driver reports the latitude and longitude every few seconds and the system stores them in a database based on distributed file storage called MongoDB.
2. When a passenger submits an order, the system selects a nearby driver through MongoDB.
3. The system pushes the order to the driver through the long connection service.
4. The driver accepts orders and then the service begins.

The access layer is mainly connected to the servers, and important protocols are all at this layer. The top of the system is the user application interface, and as a front-end device, it provides downloadable app for users and companies.

The online taxi-hailing platform mode has completely changed the operation way of taxis, but as time goes by, some of the problems in the software are gradually exposed. The main problems are as follows:

1. The intermediary agency to which the software belongs controls the price, resulting in the phenomenon of unreasonable prices and high taxi expenses.
2. Information leakage and tampering. A large amount of transaction information and user's personal data are stored in the online taxi-hailing platforms. The transaction information is easily tampered with, and the user's personal data may also be stolen after the central organization is attacked.
3. In the presence of an intermediary agency, an uneven distribution of benefits may occur. The efficiency of system information transmission gradually decreases.

Because of the above defects, we propose an improved design.

3 Application of Blockchain Technology in the Online Taxi-Hailing Platform

The concept of blockchain was first proposed by Satoshi Nakamoto in 2008.

In the blockchain connection list [7], each block contains a timestamp, a link to the previous block, and the stored data of all the timestamp. The data in each block is cryptographically encrypted and cannot be tampered with, so it ensures the security of the blockchain [8].

Under the background of the online taxi-hailing platform, the advantages of the blockchain technology that we add in the online taxi-hailing platform are as follows:

1. In the block corresponding to each timestamp, there are all taxi locations where the current timestamp exists and the taxi transaction information is stored in chronological order in a blockchain after another, forming unique data structure. For example, the location of taxi passengers and transaction information such as taxi order information are all encrypted by numerous cryptographic algorithms that use hashing algorithms as their mainstream. In blockchain technology, better security comes from the correct use of cryptographic algorithms. Using only one cryptographic algorithm is difficult to effectively ensure the security of the entire system or even an application function. Effective combination of multiple algorithms is required. Asymmetric cryptography can be used for identity verification, hashing

algorithms can be used for irreversibility, symmetric encryption can be used for large amount of data encryption, directional encryption can be used for digital envelopes, etc., so that any data that has already been recorded cannot be tampered with, which guarantees the security of the technology. As shown in Fig. 2.

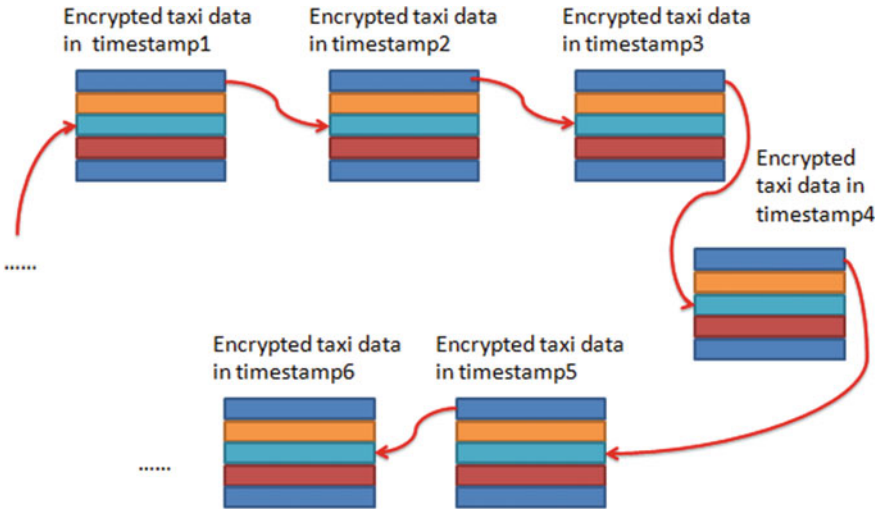


Fig. 2. The blockchain in online taxi-hailing platform

2. Each block records a unique timestamp information and the location of the previous block. These blocks are linked in chronological order to form a blockchain in the online taxi-hailing platform. The blockchain contains all the data under each timestamp. The location information can be reached through the orders in each block corresponding to each timestamp. The actual time spent in each taxi trip and the corresponding distance can be calculated, so as to accurately calculate the price of each time. This guarantees fairness of costs.
3. Decentralization is one of the important features of blockchain technology. The “centralization” in the blockchain is a kind of distributed data storage which can protect the security of information and things. Although the traditional centralization is combined with information security, there is potential risk in this form. But in the blockchain technology, the things become different. First, because of the open and transparent distributed storage characteristics of blockchains, the nodes become more reliable and trustworthy. Second, the blockchain has the characteristics that it cannot be modified and can be traced, that is, the information of the new platform will not be modified and the security can be fully guaranteed. Finally, the blockchain technology removes redundant centralization and decentralizes information to both parties of the transaction, making information communications more efficient. The form also lowers the operating expense, fasters response times, and greatly reduces the time for both parties to interact.

4 Blockchain-Based Online Taxi-Hailing Platform

As described in Sect. 2, the online taxi-hailing platform consists of a data layer, a business layer, an access layer, and a user application layer [9]. For a blockchain system, it is generally composed of a data layer, a network layer, a consensus layer, an incentive layer, a contract layer, and an application layer.

The data layer encapsulates the underlying data blocks, associated data encryption and time stamping techniques. The network layer includes data transmission mechanisms and data verification mechanisms. The consensus layer encapsulates various consensus algorithms for network nodes. The incentive layer integrates economic factors into the blockchain technology system, mainly including the issuance and distribution mechanisms of economic incentives. The contract layer mainly encapsulates various types of scripts, algorithms, and smart contracts, and it is the basis of the blockchain’s programmable features. The application layer encapsulates various application scenarios and cases of blockchain. In the absence of a central organization, information is encrypted and stored in each block. Simply completing each transaction through smart contracts and algorithms is the biggest highlight of our innovation.

It can use JAVA language to achieve a blockchain design. First use the crypto-js library to create a block, and then connect the block, and define the hash of the last block. The block can be obtained through a simple function. Then using peer-to-peer (P2P) or proof-of-work (POW) [10] network to optimize, so that we can build a blockchain system and apply this technology to the online taxi-hailing platform.

For the online taxi-hailing platform, after integrating the above two structures, our improved design plan is shown in Fig. 3.

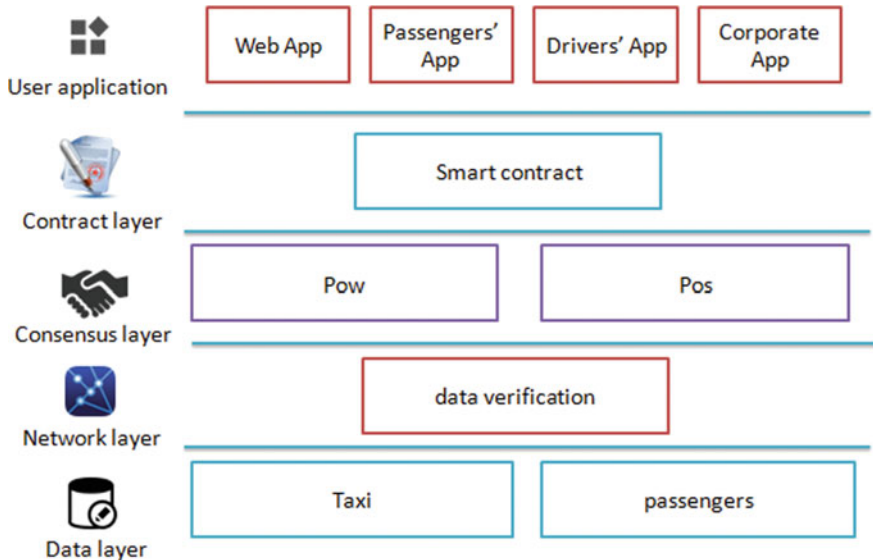


Fig. 3. The online taxi-hailing app platform

In the new scheme, we still divided the system into five layers: data layer, network layer, consensus layer, incentive layer, contract layer, and application layer. The content is modified as follows:

1. The data layer records all data and the current timestamp is required by the blockchain, including the current location of the taxi, the current location of the passenger, all transaction information and location information of the block corresponding to the previous timestamp. The data is asymmetrically encrypted on the original basis to ensure the security of all data.
2. The network layer is mainly responsible for the verification of data to ensure the accuracy of the data. When calculating the distance, consider the multiple types of road segment, and we can appropriately select an appropriate distance calculation method. Such as using Manhattan distance in the city as given below:

$$d_1 = |x_1 - x_2| + |y_1 - y_2| \quad (1)$$

Using Euclidean distances in remote areas or on a flat principle, we get

$$d_2 = \sqrt{(x_1^2 - x_2^2) + (y_1^2 - y_2^2)} \quad (2)$$

3. The consensus layer is the Proof of Stake (POS) and Proof of Work (POW). POS is also called the equity certificate and is similar to the property stored in the bank. POW assigns you the corresponding interest according to the quantity and time of your digital currency. Most virtual currencies, such as Bitcoin, Litecoin, etc., are virtual currencies based on the POW model. With these two types of certification mechanisms, the completion of the work can be guaranteed and the effectiveness of blockchain technology can be reflected.
4. At the contract level, smart contracts allow for trusted transactions without third parties. These transactions are traceable and irreversible. All users on the blockchain can see blockchain-based smart contracts.
5. The application layer is oriented to passengers, drivers, and enterprises. The purpose of this layer is to design a complete platform and to facilitate the public to download and use it on their mobile phones.

This paper has completed an improved program based on the blockchain technology for the online taxi-hailing platform.

5 Conclusions

This paper put forward a new improvement plan based on the traditional taxi-hailing platform. In this paper, we apply blockchain technology to the traditional way to remove the original tedious third-party central organization and to protect all transaction data through encryption algorithms. We can record all data from the start of the technology to the current time and use the most objective algorithm to calculate the price of each time so that the security and fairness of the system can be greatly improved.

References

1. Cohen, B.D., Kietzmann, J.: Ride on! mobility business models for the sharing economy. *Organ. Environ.* **27**(3), 279–296 (2014)
2. Economides, N., White, L.J.: Networks and compatibility: implications for antitrust. *Eur. Econ. Rev.* **38**(3–4), 651–662 (1994)
3. Govender, K.: A theoretical overview of public transport service quality: a focus on bus and mini-bus taxi service in South Africa. *J. Soc. Sci. (COES&RJ-JSS)* **3**, 301–316 (2014)
4. Swan, M.: *Blockchain: Blueprint for a New Economy*. O'Reilly Media, Inc (2015)
5. Swan, M.: Blockchain thinking: the brain as a decentralized autonomous corporation [commentary]. *IEEE Technol. Soc. Mag.* **34**(4), 41–52 (2015)
6. Godsiff, P.: *Bitcoin: Bubble or Blockchain. Agent and Multi-agent Systems: Technologies and Applications*. Springer (2015)
7. Cárdenas, Z.J.P., Avellaneda, M.A.V.: Bitcoin como alternativa transversal de intercambio monetario en la economía digital. **6**(1), 106 (2015)
8. Liu, W., Zhu, S.S., Mundie, T., Krieger, U.: Advanced block-chain architecture for e-health systems. In: *IEEE, International Conference on E-Health Networking, Applications and Services*, pp. 1–6. IEEE (2018)
9. Yuan, Y., Wang, F.-Y.: Blockchain: the state of the art and future trends. **42**, 481–494 (2016). <https://doi.org/10.16383/j.aas.2016.c160158>
10. Yang, Z., Miao, Y., Chen, Z.Y., Tang, C.B., Chen, X.: Zero-determinant strategy for the algorithm optimize of Blockchain PoW consensus, pp. 1441–1446 (2017)



The Intangible Culture Heritage “New Ecology” Under Emotional Demand of Interactive Experience

Dongna Cai^(✉), Yuning Li, Zhi Li, and Yongjian Huai^(✉)

Beijing Forestry University, Beijing 100083, China
caidongna@bjfu.edu.cn, 404620662@qq.com,
huaiyj@163.com

Abstract. The traditional Chinese flower arrangement as a national intangible cultural heritage has high historical value and artistic value, with the increasing demand of user’s emotional needs for interactive experience, its popularity, and interaction are relatively limited. In the information age of the rapid development of new media technology, using digital technology to protect and inherit this excellent intangible cultural traditional art has far-reaching significance and it has high theoretical value for developing the history and innovation of traditional flower arranging art. Based on the augmented reality technology of intelligent mobile terminal, the article integrates the digital resources of flower arranging art from the perspective of satisfying the user’s emotional needs, restore the art of flower arranging classic works, and enhance user interaction experience. Build an operable digital virtual flower arranging application system that transcends time and space, the combination of traditional flower arranging art and new media communication platform can form a three-dimensional digital “new ecosystem” for the dissemination and protection of intangible cultural heritage.

Keywords: Intangible culture heritage · Interactive experience · Emotional needs · New digital ecology

1 Introduction

In June 2008, Chinese traditional flower arrangement as an important part of Chinese traditional culture was included in the national list of intangible cultural heritage (ICH). The traditional flower arrangement has the characteristics of strong ornamental and short retention time [1]. Its defects lie in poor interaction, weak sense of participation, and cannot be preserved for a long time. China is the origin country of flower arranging but the traditional flower arranging is almost cut off [2]. Facing the plight of traditional flower arranging and combining the characteristics of flower arranging that is imminent to create and preserve the ICH. In recent years, attempts to digitally innovate and preserve material cultural heritage or ICH are increasing, scientific and technological innovation promotes cultural heritage in a new form in the public view and many traditional cultures reappearance vitality through digitalization such as Dunhuang Research Institute in order to protect the world cultural heritage. Dunhuang grottoes uses various technical means to digitally reproduce Dunhuang culture [3]. In 2009,

Wang Xiao-Fen et al. discussed how to digitally preserve the paper cut of ICH [4]. In 2016, the new mode of cultural dissemination of rice paper was discussed based on augmented reality technology [5]. In the same year, Li Long-xing, Yi Dong-cheng discussed the necessity of the realization of the digital museum and put forward the important steps to establish the digital museum [6]. In 2017, Wang De-fang and Zhang Xiao-juan listed several digital media technologies that played a role in the protection of ICH in Qinghai [7]. Hayun Kim et al. have studied the application of mobile augmented reality technology in information modeling and user research in cultural heritage websites [8]. Francesco Piccialli and Angelo Chianese reveal the trends and challenges between cultural heritage and new technologies, they point out that new technologies can shorten the time and space distance between human and cultural heritage [9]. Bernard K. Means put forward 3D documents and records management in the process of cultural heritage protection in 2017 [10]. In 2018, Zois Koukopoulos et al. joined the platform of cultural heritage management communication and multimedia participation, besides analyzing the challenges to create a multimedia cultural heritage platform [11]. In summary, the above literature research and discussion of the combination of the material cultural heritage or ICH with the new technology shows the protection and innovation in recent years.

At present, the way of cultural inheritance of Chinese traditional flower arranging is still limited to traditional classroom teaching, paper teaching materials, and association activities. It is seriously divorced from the current trend of digital technology, in particular, few digital studies on the knowledge propaganda, skill inheritance, and market promotion of flower arranging. Therefore, the use of digital inheritance, protection, and development of traditional Chinese flower arranging is needed of foreign cultural exchange and it is also an indispensable part of the traditional Chinese flower arranging for the development of national culture. In the digital technology, the development of virtual reality technology, such as holographic imaging technology and augmented reality technology (AR) develop fastest and popular. This paper will discuss the realization process and achievement of the integration of traditional flower arranging and digitalization from the methods and materials, digital transformation, art reproduction, entertainment science, and so on. Discusses the planning process and technical realization of the combination of flower arranging and digital science and technology media in the virtual reality mobile phone application form, and uses the new media platform and new technology to carry out a complete system of information digital popularization of science, art 3D reproduction and enhancement of interactive experience for the flower arrangement of ICH. The purpose is to enhance the user's interactive emotional experience on flower arranging, improve the immersion and entertainment of the application, and realize the popularization and persistent preservation of multidimensional ICH.

2 The Overall Design of the Application System

In order to enhance the interactive experience of the ICH and meet the user's emotional needs, the author team developed the augmented reality mobile application of “Fingertip Beauty”. The overall process of application is divided into functional

module design, application development, visual style design, and AR technology implementation. The overall function of APP is mainly divided into three functional modules: popularization of science interaction module, augmented reality module, flower arranging replacement game module and achieve fluency complete operation on Android platform. The popular science interactive module realizes two-dimensional graphics pages show and operable 3D Flower model. The augmented reality module is to realize the multidimensional display and artistic popularization of the flower arranging art. And the flower arranging replacement game module is to realize the application and experience of flower arrangement through game entertainment. The three modules plan to form a gradual experience effect through the in-depth understanding of the cultural information of the two modules, the replacement game module has increased participation experience under the viewing experience, and has enhanced the interactive demand and application. The visual design of the application system adopts the traditional Chinese elaborate-style painting with a fresh and elegant color design, and selects 12 representative flowers according to the four seasons, integrates the 2D and 3D digitized resources of the corresponding classic works of flower arrangement, such as graphic information, three-dimensional model, scene rendering, enhancement, and so on. Some resources and operation platforms are integrated and are innovative to realize the new display of traditional flower arrangement in the form of new media interaction.

2.1 Module Planning

Augmented reality mobile applications based on Android's operating system, unity interaction platform, and AR effect enhance the interaction between users and products. The system module is divided into the following three main modules, namely, popularization of science interaction module (Fig. 1①), augmented reality module (Fig. 1②), and flower arranging replacement game module (Fig. 1③).

Each module can be explained in detail as entering the main interface (logo interface and loading interface) to guide the user into the application, the whole application starts from the selection month module into the details page, and the user can choose to enter the ①, ②, or ③, and these three modules carry the emphasis of application. The popular science module adopts the Chinese traditional vertical layout format that has each page embedded with the 3D model of the corresponding flower. It can be rotated and displayed independently and we can click on the magnifying view, and this part realizes the stereoscopic multidimensional flower science detail display. This link is not limited to traditional graphics and text-based two-dimensional science and technology and the aim is to understand the plant structure and enhance the sense of operation. After users scan the flower two-dimensional graphics card with the smart mobile phone camera, the corresponding 3D flower modeling can be displayed on the mobile phone screen while AR images can be rotated and zoomed by fingers at 360°. At the same time, the scene animation with the corresponding season, including butterfly flying, deciduous, rain and snow particles, flower model accompanied situation rendering, and strives to restore the user's more realistic artistic perception and cultural experience. This part makes use of augmented reality technology to bring users into deeper interactive emotional experience and further strengthen cultural science

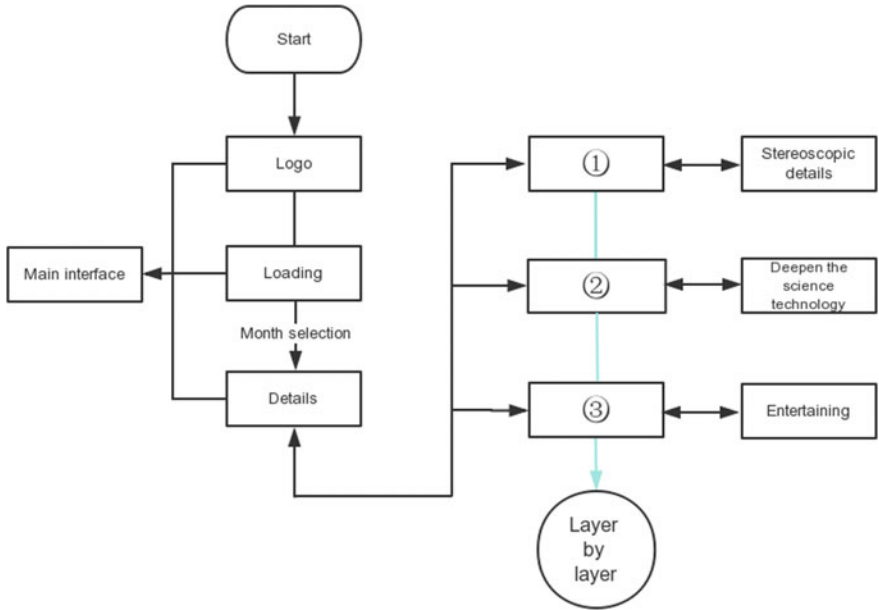


Fig. 1. Module flowchart


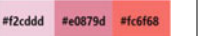

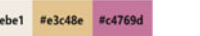
richness. The flower replacement game module provides three kinds of four-season flowers 3D material and is equipped with different pots, branches, and leaves 3D material to match the different flower arrangements. This module will entertain the art and culture of flower arrangement that makes it an interactive game and the user can form the maximum interaction with the art of inserting flower in this link, and embody the spread and application after science popularization. The relationship between the three modules is step by step and the realization of need level presents a pagoda form, which conforms to the multidimensional interpretation of Maslow’s theory of hierarchy of needs [12].

2.2 Data Integration

After module planning, the layout of flower arrangement work is integrated with 2D and 3D material data, which is divided into four seasonal plates in spring, summer, autumn, and winter. In the four seasons, three representative flowers and flower arranging works in China are selected in 12 categories. The flowers in spring are peony, magnolia, and lily. In summer, flowers are lotus, pomegranate flower, and rose. In autumn, are chrysanthemum, osmanthus, and orchids. In winter, the flowers are plum blossom, narcissus, and cyclamen. This season and flower collocation ideas continue to the core three main modules. The flower color matching of 2D design and 3D model basically follows the flower color of real life, and therefore extends the color to the interface color matching of the entire application system. Meantime, in order to show the expression and cultural connotation of the flower arrangement more

comprehensively, the traditional flower arranging art also embodies the combination of landscape and flower in the application system, making a series of dynamic materials to render the artistic conception of the flower arrangement (Table 1).

Table 1. Flower arrangement material

Season	Spring	Summer	Autumn	Winter
Flower species	Peony, magnolia, lilies	Lotus, pomegranate flower, rose	Chrysanthemum, osmanthus, orchid	Plum blossom, narcissus, cyclamen
Color				
Animated material	Butterfly flying	The breeze, the rain, and snow particles	Falling leaves fall	Snowflake

In the popular science interactive module, three kinds of representative flowers are selected for popular science introduction each season. In this module, each flower has three corresponding flower illustrations and three-dimensional details. In the AR identification module, there is a complete set of augmented reality flower arrangement restoration works in each season (Table 2). In the flower arranging replacement game module, there are 12 flowers in each season, which are divided into three flower pots, three flower branches, three flowers, and three leaves, the material can be freely matched, and there are 324 kinds of collocation for four-season flower arranging (Table 2 and Fig. 2).

Table 2. Digital flower data table

	Spring	Summer	Autumn	Winter	Total
Flower 3D model element	12	12	12	12	48
Standard dressing combination	3	3	3	3	12
Popular science page	3	3	3	3	12
Flower interface	3	3	3	3	12
Replacement game ways	81	81	81	81	324
AR recovery works	1	1	1	1	4

3 Digital Transformation and Implementation

After module planning and data integration, the digital transformation of flower arrangements needs to be implemented through three-dimensional models and technical means. The main program system architecture of this mobile phone application is listed below. Operating system is Windows 10 64-bit version 1703, running system is Android 5.1 “Lollipop” (API Level 22) development tool Unity Personal 5.3.3 (Windows) MonoDevelop Vuforia API. The hardware environment of the mobile

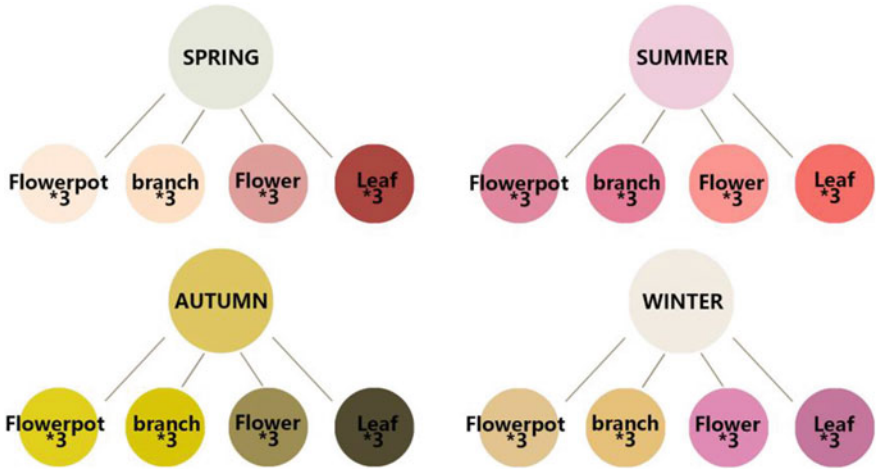


Fig. 2. Overview of the floral interface material

phone application is listed below. The CPU is Intel(R) Core(TM) I7-4500U CPU @ 1.80 GHz 2.40 GHz, 8 GB memory, graphics cards are Intel(R) HD Graphics Family and NVIDIA GeForce GT 720M, network card requires 2×2 11bgn Wireless LAN M.2 Adapter, and screen resolution is 1080p.

3.1 3D Model Implementation

All kinds of floral three-dimensional materials required for modular planning of this application are modeled with 3DMAX, the texture is created with Adobe Photoshop CS6, Adobe Illustrator CS6, and other software. While modeling, the 3D flower model must keep the curve and smoothness of the real flower. The 3D model repair tool was used to repair and polish the flower model. The characteristic of each traditional Chinese flower arrangement was sought to be kept, to preserve the charm of the ICH maximally. Part of the flower arrangement restoration is shown in (Fig. 3). To enhance the efficiency of the post-production process, the following model naming conventions have been formulated: (1) Camel nomenclature. (2) Prefixes and suffixes are required. The prefix is used to distinguish the category of the model, for which *dz* means the single plant model, and *zh* means the flower arrangement combination model. The prefix and file name are connected by a dash. The suffix is a four-digit editing date. (3) Each part of the flower arranging combination is grouped. The group name consists of three parts. Active/inactive tag to indicate whether the packet is a replaceable component, sequence number to indicate which type of component this group is, and the name of the group named after camel nomenclature.



Fig. 3. Original lotus flower arranging and 3D model

3.2 Program Implementation

The technical difficulties are mainly distributed in three aspects: the interaction in popular science module, the implement of augmented reality module, and the interaction in flower rearrangement module. The specific system architecture is shown in (Fig. 4). In the popular science module, three-dimensional flower arrangement

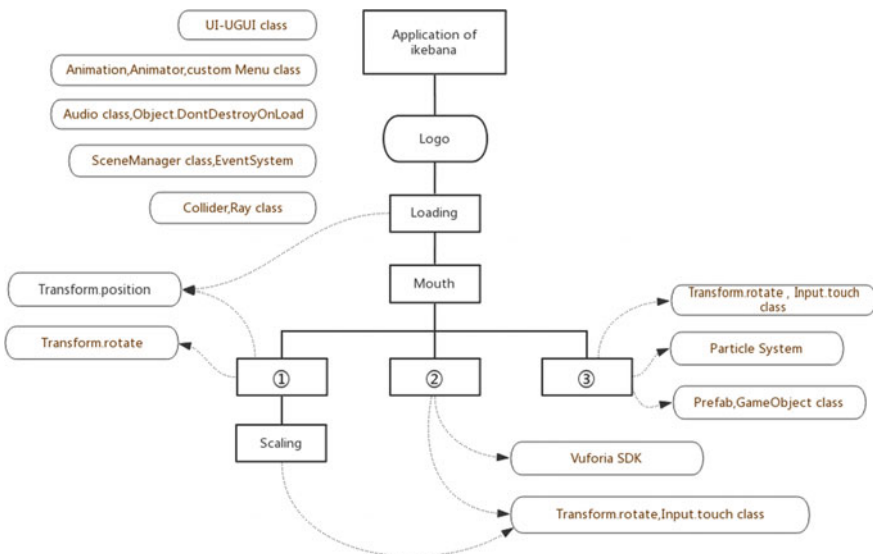


Fig. 4. Augmented reality mobile application system architecture

model is embedded in the 2D user interface. The interaction includes the 3D model rotation function and the up-down slide function, both of which are realized by Unity engine. The functions below are written in pseudocode.

Three-dimensional model rotation function:

```
void Update () {
    define the default speed;
    if (number of touch <= 0) {
        model spins at default speed;
    }
    if (number of touch = 1) {
        define the touch;

        get the position delta of the touch;

        define the new speed by the position delta;

        model spins at the new speed;
    }
}
```

Up-down slide function (2)

```
void Update() {
    define the bound of camera space;

    define the default speed for camera;
    if (touch detected) {
        define the touch;

        define the initial position of the camera;

        get the y-axis position delta of the touch;

        calculate the transform position of the camera;

        if (new position is inside the bound)
            camera moves to new position;
        else camera moves to the edge of the bound;
    }
}
```

The technical difficulty in the realization of the virtual reality part is whether to connect the real world with the virtual world seamlessly, to enhance the immersive and interaction of user. In terms of implementation, the Vuforia SDK is used to perform 3D model-to-virtual reality conversion, which makes user to feel the integration of the real world and the virtual world. In the AR part, *Transform.rotate* and *Input.touch* functions are used to enhance the experience of interaction. To activate Vuforia, we need to import the Vuforia SDK file package to Unity and enter the serial number, then the AR function provided by Vuforia is available. To prepare the target library, we need to log

in personal account on the official website, enter the library page, add a new database or select an existing database, upload the picture of the flowers, then recognition target identifies the feature point is automatically generated. When all the pictures are uploaded, select Download Database (All) to get the unity package. Import the package into Unity, then the target library is ready. To Link identification target and model, drag 3D flower model in Unity scene. In this scene, create a Vuforia image target (Vuforia-Image), and then adjust the properties of the target image in the Inspector. The 3D model should be moved onto the recognition target, and the direction and size of the model should be adjusted. Set the 3D model as a sub-item for the target to complete the link between them. To test, use Unity Remote to connect the tester or export the project to the tester. When the tester camera captures the target image, the corresponding model should be automatically displayed.

The technical difficulty of flower rearrangement is to make smooth user experience while operating the 3D model of the flower arranging material. The simulation of the flower arrangement needs to be like the action of the operation in real life. The process of rearranging is to let user to master the art of flower arrangement and aesthetic sense.

The implementation of the flower rearrangement module is given as follows:

```

define new class Menu{                                     (1)

    public bool IsOpen{

        return to "IsOpen" function in Animator;

    };

    public void Awake();

    public void Update(){

        if (current canvas' Animator state is open)

            make the canvas interactable;

        else make the canvas not interactable;

    }

}

define new class MenuManager to manage menu{

    initiate a menu as CurrentMenu;

    public void Start(){
        ShowMenu(CurrentMenu);
    }

    public void ShowMenu(Menu menu){
        if (there is already a menu in the field)
            CurrentMenu.IsOpen = false;
    }
}

```

- (2) Implementation of the user interface scrolling: A visual scrolling is implemented by setting the movement of the camera. The length of the scroll is determined by a global variable *length*.
- (3) Switching between different scenes: Set *EventSystem* in scene to listen for commands, set button to send scene-switching commands, and implement the switching through *SceneManager* class. Different commands link to different scenes. Attach the command list script on the *EventSystem* to implement the script.
- (4) Implementation of the flower rearranging: Set *EventSystem* to listen for command and button to send rearrange command. Set the arrangeable part to a prefab, set the *GameObject* class *object* as global variable, and then match the arrangeable part to each of the prefab. The display status of prefab can be changed through *SetActive* function.

4 Results and Discuss

After a series of planning and implementation, the result is the Android application “Fingertip Beauty”. During the research process, two-dimensional illustration, user interface, and three-dimensional model library of four-season flower arrangements were established (Fig. 5). The four-season flower arrangement AR function was implemented (Fig. 6) to enable new technologies to be integrated in traditional flower arrangements as well as to achieve freedom of flower arrangement. The rearrangement system enhances the fun of flower arrangement. The “Fingertip Beauty” application integrates the above functions on a new media platform, greatly enhances the digital operability of the flower arrangement art, from 2D to 3D. The improvement of the



Fig. 5. Three-dimensional library of flower arranging materials

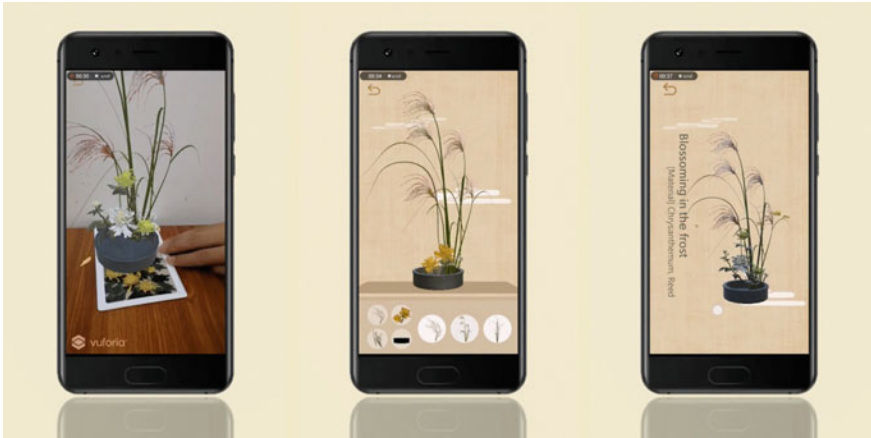


Fig. 6. Flower rearrangement module

immersive and interactive experience makes an enhanced user emotional experience. The digital form of the protection of ICH shows variety of operation and display forms of traditional cultural, a new and long-lasting preservation of the flower arrangement is established. “Fingertip Beauty” achieved a series of research goals. In terms of design, it applied antique color style and elegant design style in the art as well as finishing the combination of digital technology and art. In terms of operation, the application has innovated the traditional flower arrangement, and three modules gradually deepen the user’s interactive emotional experience and realize the digital operability of flower arrangement. In terms of technology, the system’s procedures are tightly integrated, autonomously enabling the display of augmented reality illustration models, and enabling a smooth and natural gameplay experience. At the same time, this application acquired the People’s Republic of China software copyright (registration number: 2018SR209700) and won the third prize in the second international VR/AR Jincheng award in 2017.

Although the application has achieved many goals, there are shortcomings as well, such as the promotion issue. The internal test for professional users has a good response, but the application itself has not officially put on the market as a mature product in the application store. As a prototype, the landing of this product has yet to be inspected and implemented. In addition, the flower rearrangement module does not have a sharing function. The digital flower arrangement works can only be viewed by user but not be shared to the public. Sharing the user’s floral arrangement in the future is going to be a meaningful research.

5 Conclusion

This paper starts with the emotional needs of the interactive experience, digitizes the ICH to realize a virtual reality flower arrangement application. In the process of implementation, a set of operational procedures for planning modules, data integration, and technology implementation of virtual reality mobile phone applications are planned to ICH heritage flower arrangements with new media platforms, augmented reality technologies and interactive games, and to increase the multidimensional dimension of traditional flower arrangements. The operability and interactive participation make the art of flower arrangement surpass the limitation of time and space, from physical display to digital exhibition, from static viewing to the innovation of interactive experience. Looking into the practical application of the future, “Fingertip Beauty” and its similar products can be used in teaching, to enhance the interactivity of knowledge as well as improve the flower arrangement teaching to make it more life-oriented and civilian. It can be used in exhibitions to display ICH and can be promoted based on numbers to promote popularization of ICH. It can also be used for holographic display. In the context of interactive technologies and cultural heritage, it provides a practical implementation plan for the digitization of other cultural heritage and ICH, which is conducive to the popularization of intangible cultural transmission and enhancement of information technology, and the promotion of intangible cultural heritage. The digital inheritance and the expansion of intangible cultural preservation channels to create a new digital ecological structure of ICH.

Acknowledgements. The author would like to thank Beijing Forestry University for support. This research is supported by the Funded projects: Fundamental Research Funds for the Central Universities (2016JX08).

References

1. Tan, C.J.: Choose non material cultural heritage, not to be set up. Chinese Flower Newspaper, P.1 (2008, September 02)
2. Wang, L.Y.: The brief development history of flower arrangement in China. Chin. Landsc. Arch. **2006**(11), P.44–48(2006)
3. Fan, J.X.: Application of digital technology in protection and exhibition of Dunhuang Grottoes. Dun Huang Res. (06), P.1–3(2009)
4. Wang, X., Liu, Y., Zhang, W.: Research on modelling digital paper-cut preservation. Int. J. Autom. Comput. **6**(4), 356–363 (2009)
5. Zhi, Z.: A new mode of cultural dissemination of Xuan paper based on augmented reality technology—taking the construction of cultural system of augmented reality Xuan paper as an example. Today’s Massmedia **01**, 85–90 (2016)
6. Li, L.X., Yi, D.C.: On the establishment and dissemination of the digital museum of intangible cultural heritage. West Leather **22**, 197–198 (2016)

7. Wang, D.F., Zhang, X.J.: Application of digital media technology in the protection of intangible cultural heritage in Qinghai. *J. Qinghai Norm. Univ. (Natural Science Edition)* (04), 6–10 (2017)
8. Kim, H., et al.: Ontology-based mobile augmented reality in cultural heritage sites: information modeling and user study. *Multimed. Tools Appl.* **76**(24), 26001–26029 (2017)
9. Piccialli, F., Chianese, A.: Cultural heritage and new technologies: trends and challenges. *Pers. Ubiquit. Comput.* **21**(2), 187–189 (2017)
10. Means, B.K.: 3D recording, documentation and management of cultural heritage. *Hist. Archaeol.* **51**(4), 582–583 (2017)
11. Phutane, P., et al.: Preliminary trial of augmented reality performed on a laparoscopic left hepatectomy. *Surg. Endosc.* **32**(1), 514–515 (2018)
12. Hu, J.X.: The multidimensional interpretation of Maslow’s theory of hierarchy. *Philos. Res.* **08**, 104–108 (2015)



Research on Optimization Model of Storage Capacity Based on the Consortium Blockchain

Xiaotian Wei, Jiahua Chen, and Zhihuai Li^(✉)

Information College, Dalian Maritime University, Dalian 116000, China
{weixt,v8990447,qhlee}@dlmu.edu.cn

Abstract. To solve the problem of increasing nodes data in blockchain, we propose an optimization model of storage capacity based on the consortium blockchain (OMSCCB). We have divided blockchain network into groups by the characteristics of partial decentralization of the consortium blockchain in the OMSCCB. In each group, the full node is selected according to the hardware performance and the accumulated reliability evaluation of the node. The full node is responsible for storing the blockchain data, and other ordinary nodes segment storage of the blockchain data according to the consistent hashing. And then, we have designed a replacement strategy of full node to ensure that the OMSCCB can resume normal operation when the network is attacked. The experimental results show that the storage cost of the system can be reduced effectively, and the reliability of the system can be guaranteed at the same time.

Keywords: OMSCCB · Full node · Segment storage · Consistent hashing · Consortium blockchain

1 Introduction

Blockchain [1] has recently attracted wide attention from all walks of life. They are actively exploring blockchain technology and developing the application of blockchain [2–4]. But there are still problems in the storage capacity of blockchain. Taking Bitcoin as an example, the storage capacity of a single node increased from 137 to 166 GB from May to November 2017. As the storage capacity of the node increases rapidly, the requirement of the node joining the network becomes higher and higher. Therefore, the optimization of the blockchain storage capacity has become a problem needed to be solved urgently.

In view of the optimization of blockchain capacity, in the literature [5], the author proposed a data replica allocation strategy to reduce node storage capacity, but this method results in most block data being stored on a few nodes. In the literature [6], the author proposed mini-blockchain, which introduced the concept of account tree. Nodes only store the nearest new block and account tree,

which reduces the storage cost, but it is difficult to verify the old transaction information. In the literature [7], the author proposed that lightning network can reduce the storage of blockchain data by off-chain transaction, but the payment channel of lightning network will lead to a certain degree of centralization.

In view of the increasing number of data in blockchain nodes, a capacity storage optimization model based on consortium blockchain is proposed. The main contributions of this paper are as follows:

- (1) The OMSCCB divides the nodes into two types: the full node and the ordinary node. The full node is selected according to the node hardware performance and the reliability evaluation of the node service, and the complete blockchain data is stored. The ordinary node segments the storage blockchain data according to the consistent hashing [8] to optimize the storage capacity of the blockchain. The transaction query and verification services in the group are handed over to the full node to reduce the performance burden of the ordinary node.
- (2) Based on consistent hashing, virtual node [9] is introduced into the OMSCCB to dynamically allocate block data according to node performance to achieve load balancing among nodes.
- (3) The OMSCCB considers the situation that the node is attacked and cannot work properly, so we propose a full node replacement strategy to ensure the security and stability of the OMSCCB.

2 The Design of the OMSCCB

The design of the OMSCCB mainly includes selection strategy of the full node, segment storage of the blockchain data, transaction verification, and replacement of the full node.

2.1 Selection Strategy of the Full Node

The blockchain network is divided into multiple groups in the OMSCCB. Each group selects a full node to undertake the corresponding services within the group, thus releasing the hardware performance burden of other nodes. The hardware performance, storage capacity, and reliability of the node should be considered synthetically. The appropriate node is selected as the full node, which is responsible for the query service within the group.

Calculation of the Node Performance. The OMSCCB proposed is that the full node is responsible for storing the complete blockchain data, so the storage capacity will be counted into the comprehensive performance of each node. The ability of node to process and respond to business requests is mainly affected by computing units, network bandwidth, and other factors; the performance of node is determined by the storage capacity of hard disk and CPU parameters, disk IO rate, memory, and network bandwidth; and the performance of each

node is calculated by weighted summation of hardware parameters. The formula for node performance is shown in Formula 1.

$$P(Node_i) = k_1 * S_i + k_2 * n * C_i + k_3 * D_i + k_4 * M_i + k_5 * W_i$$

$$\sum_{i=1}^5 k_i = 1 \tag{1}$$

In the above formula, $P(Node_i)$ is the performance value of $Node_i$; S_i is the storage capacity of hard disk; C_i is the frequency of CPU processing; n is the number of CPU kernels; D_i is the number of disk IO rates; M_i is the memory size; W_i is the network bandwidth; k_i is the weights corresponding to the node hardware. In the actual scenario, k_i need to set a reasonable weight value which is set based on the type of service provided by the $Node_i$. In the OMSCCB, the main service provided by nodes is to store data and query data, which requires a lot of storage capacity, computing power, and data reading and writing. Therefore, k_i is set to (0.3, 0.3, 0.3, 0.05, 0.05).

Calculation of the Node Reliability. The OMSCCB calculates the reliability of node based on the evaluation feedback and time factors. After each interaction process, the nodes in the network will evaluate the resource providing nodes, which will be recorded as E_{ij} . It represents the degree of satisfaction given by $Node_i$ to $Node_j$ after $Node_j$ provided service for $Node_i$. Its value is in the range of [0, 1]. The evaluation of the interaction between nodes determines the degree of influence on the final description by referring to the time of the action. The closer the interaction occurs to the current time, the deeper the impact is. The formula for quantifying the influence of time characteristics is shown in Formula 2. The calculation formula of the reliability of each node is shown in Formula 3.

$$T(E_{ij}) = e^{-|t_{now} - t_m|} \tag{2}$$

$$R(Node_j) = \sum_{i=0}^n (T(E_{ij}) * E_{ij}) / \sum_{i=0}^n T(E_{ij}) \tag{3}$$

$T(E_{ij})$ is the time attenuation degree of the interactive evaluation; t_{now} is the current time; t_m is the action evaluation time. $R(Node_j)$ is a reliable value of $Node_j$.

Calculation of the Node Comprehensive Evaluation. The selection of the full node needs to take into account the hardware performance and the reliability of the node. It is necessary to refer to $P(Node_i)$ and $R(Node_j)$ in the process of selecting the full node. When calculating the comprehensive evaluation of the node, the optimal performance node $P_{max}(Node)$ in the group is introduced as the reference, to balance the parameters.

$$H(Node_i) = P(Node_i) / P_{max}(Node) + R(Node_i) \tag{4}$$

$H(Node_i)$ is Comprehensive capability Evaluation of $Node_i$. According to the rank of the group, the node with the highest $H(Node_i)$ value is selected as the full node of the group.

2.2 Segment Storage of the Blockchain Data

Blockchain data is segmented storage [10] according to the consistent hashing. The hash value of each block in the blockchain is mapped to a hash ring with a range of $[0, 2^{256}-1]$. Each node is mapped to the hash ring by SHA-256 algorithm [11] according to the IP address and port number. Each node stores the segmented block data of the corresponding region. Blockchain data block mapping relationship is shown in Fig. 1. When the node joins or exits, it is also mapped to the hash loop, and the data of the corresponding region is migrated to the node or the subsequent node to realize the dynamic expansion requirements of the node.

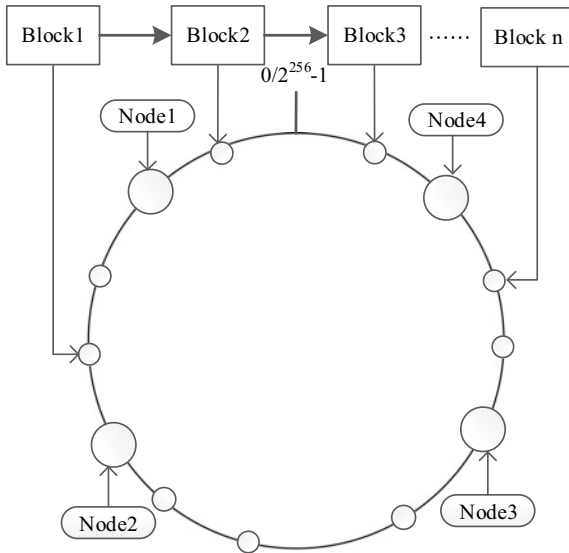


Fig. 1. Blockchain block mapping

Segment Storage. Because the hardware performance of different nodes is different, and the hash result is random, it is possible that the performance of the node does not match the storage capacity. The idea of virtual node is introduced to maintain load balance among different performance nodes in the OMSCCB. Suppose there are $N + 1$ nodes in each group, one node is selected as a full node from each group according to the full node selection strategy. The full nodes stores all the blockchain data, and the other N nodes store the blockchain data segmented. The consistent hashing ring is divided into the number of virtual nodes set for the group by M regions of the same size and M is the total number of virtual nodes set for this group. Due to the difference in performance between nodes, M to actual node number N cannot be determined. The number of virtual nodes should be allocated according to their performance.

The total performance of each node in the group can be obtained according to Formula 1. The sum of the total performance of the ordinary node in the group can be obtained according to Formula 5.

$$P_{total} = \sum_{i=0}^n P(Node_i) \tag{5}$$

P_{total} is sum of the total performance of the ordinary node in the group. According to the allocation scheme of virtual nodes, the allocation of the number of virtual nodes per node is known and calculated by Formula 6.

$$V(Node_i) = M * (P(Node_i)/P_{total}) \tag{6}$$

$V(Node_i)$ is the number of virtual nodes allocated by node i ; M is the total number of virtual nodes; $P(Node_i)/P_{total}$ is the proportion of node i in the group.

Addition of the New Node. For example, when the new node is added to the group, the performance value of the new node is calculated according to Formula 1. The virtual node assignment is recalculated according to Formula 6, and the virtual node mapping table is updated. The performance of the node is matched with the size of the stored data, and the load balance of the data is realized. The change of the virtual node mapping is shown in Fig. 2.

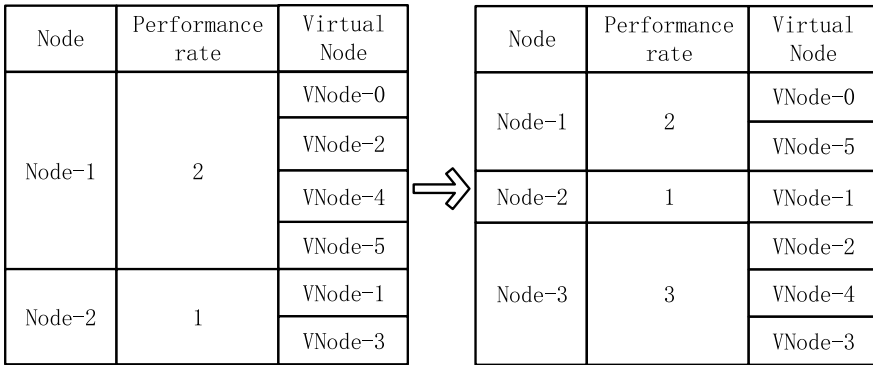


Fig. 2. Virtual node allocation

2.3 Transaction Verification

In the blockchain, each transaction can be verified by the information of the query chain. In the model, each group is responsible for the transaction query service in the group by the full node with high reliability. The ordinary node acquires the verification result by interacting with the full node.

The full node selected by this group may be malicious in disguise. When verifying the transaction information, ordinary nodes not only query the full node of the group but also randomly select several full node verifies from the full node list of the blockchain network. The verification result is decided according to the result set returned by all the full node. The specific steps are shown in Fig. 3.

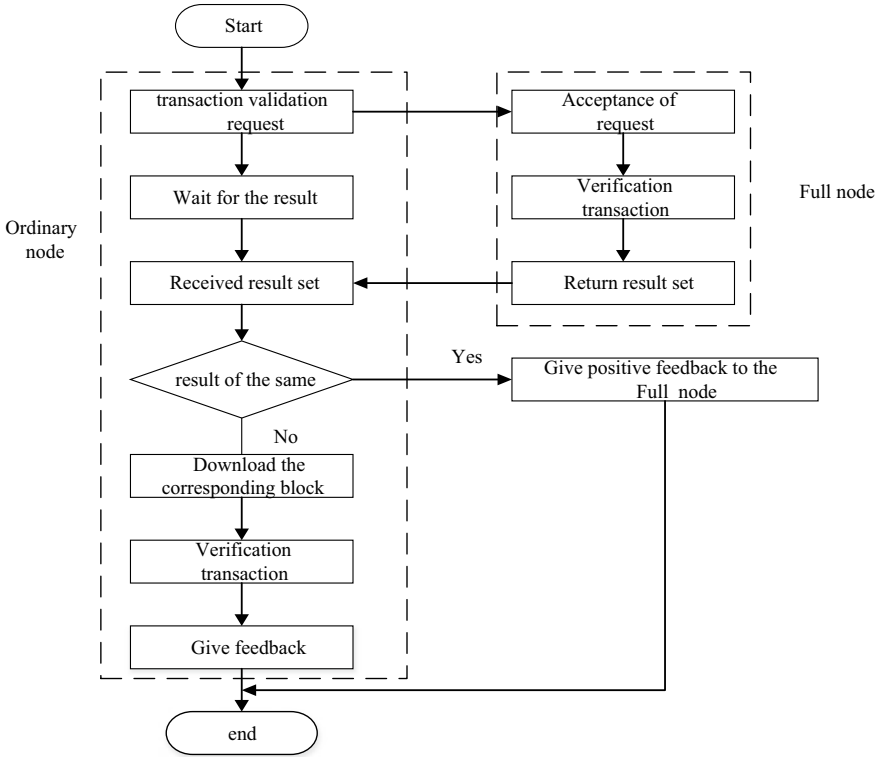


Fig. 3. Transaction verification

2.4 Replacement Strategy of the Full Node

The full node is the node with the best comprehensive performance in the group, and the reliability is high. It is undeniable that there is the possibility of being attacked or downtime. When the full node stops working, the transaction query and other services in the group cannot be carried out normally; therefore, it is necessary to reselect the full node to maintain the stability of the system. The process of replacing the full node is shown in Fig. 4.

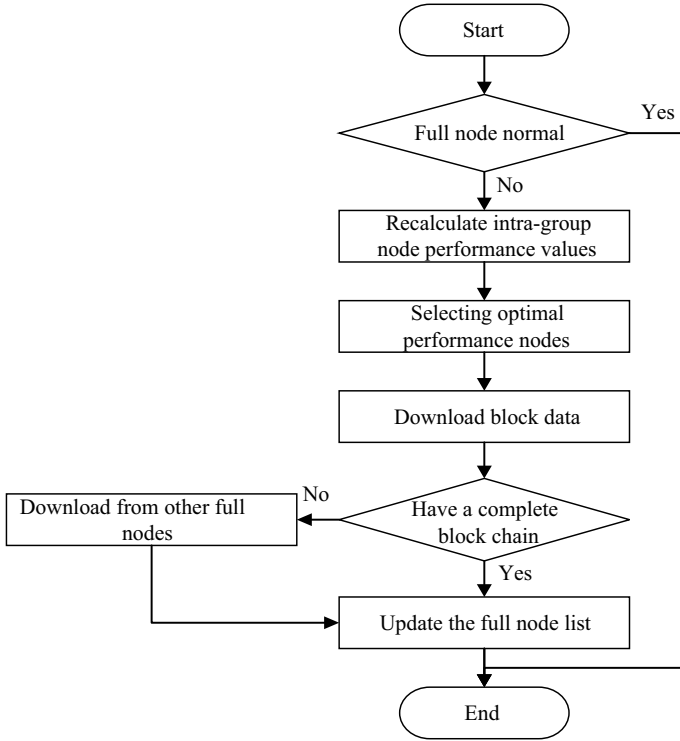


Fig. 4. Replacement strategy of the full node

3 Experimental Results

The hardware environment of the experiment is eight servers with different performance. The OMSCCB is partially modified on Hyperledger Fabric1.0 [12] storage model. The storage rules based on consistent hashing are added to simulate the experiment. In the experiment, the full node is order node in Hyperledger Fabric, participating in the consensus of blockchain. It generates new blocks for transaction sorting and is also a peer for storing block data. Other software environments are Docker 17.03.1 Go 1.6.2 and java 1.8.015. Storage overhead, transaction verification performance, and model reliability on the OMSCCB are tested.

3.1 Test of the Storage Overhead

In the experiment, the storage overhead of the OMSCCB is evaluated in units of the group. First, the size of the storage is counted when the number of blocks in the OMSCCB is 1000/2000/3000/4000/5000. Then, in the case of a specific number of blocks, we have counted and compared the size of storage capacity in

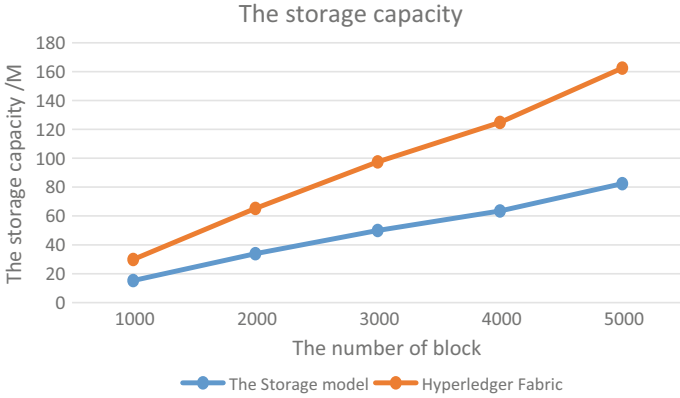


Fig. 5. The comparison of storage capacity

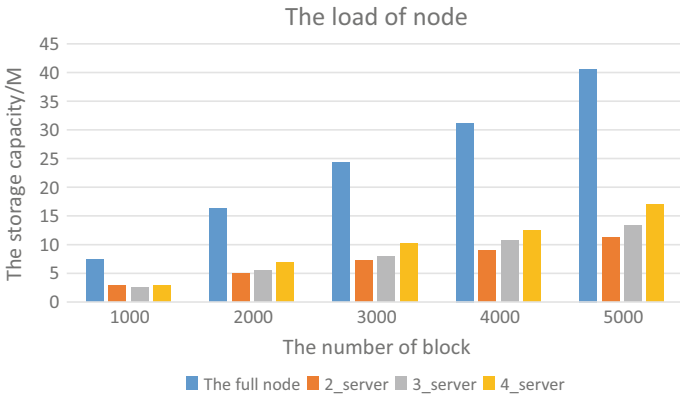


Fig. 6. The load of nodes

the Hyperledger Fabric model and the OMSCCB proposed in this paper. The results are compared as shown in Fig. 5.

For the node load test of the OMSCCB, we selected a group of nodes from Server 2 to Server 5. In this group, Server 5 is the full node in the group, and the other three Servers are the ordinary nodes. The storage capacity of each node block of Server 2 to Server 5 is calculated when the number of blocks is 1000/2000/3000/4000/5000, and the result is shown in Fig. 6.

From the above two experiments, it can be seen that compared with Hyperledger Fabric. The OMSCCB can effectively reduce the storage capacity of blockchain in the same scale network. The increase of storage overhead is relatively smooth. The OMSCCB also takes into account the hardware constraints of ordinary nodes and distributes blockchain data according to the performance of each node. The load balance between nodes is realized and the hardware limitation of adding blockchain network nodes is reduced.

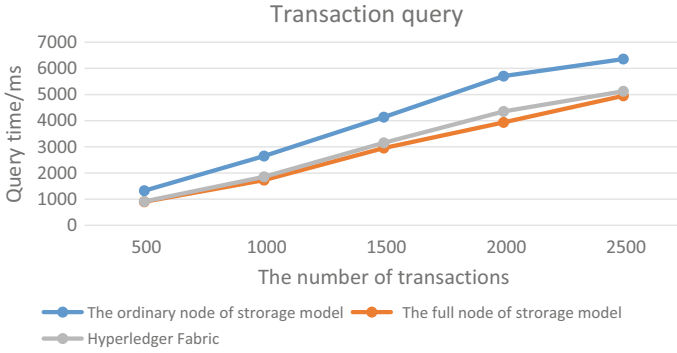


Fig. 7. Query transaction

3.2 Test of the Transaction Verification

In order to test the transaction verification of the OMSCCB, the experiment took the response time of the node to the same-level transaction query verification request as the reference. In the experiment, Server 2 is set up to query 500/1000/1500/2000/2500 transactions. We count the time them take to query in the ordinary node of the OMSCCB, full node of the OMSCCB, and Hyperledger Fabric query. The query results are shown in Fig. 7.

Figure 7 shows that the ordinary node in the OMSCCB takes a little more time than the Hyperledger fabric in handling query transaction requests of the same order of magnitude. While the full node and the Hyperledger Fabric node in the OMSCCB both store the blockchain information, the transaction query time is basically the same. In order to reduce the overhead of the storage capacity of the OMSCCB, the ordinary nodes only store block data segmentarily. The transaction verification needs the help of the full node in the group, so the transaction query time of the ordinary node transaction is a little longer than that of the Hyperledger Fabric.

3.3 Test of the OMSCCB Reliability

In the experiment, Server 5 was artificially shut down to simulate the full node to stop service because of network attack. During this period, Server 2 always sends the transaction verification request to the full node in the group to test the full node service condition. Finally, calculate the transaction verification efficiency to reflect the reliability of the model according to the number of transaction verifications in the log data int unit time. The statistics of transaction verification efficiency is shown in Fig. 8. The block storage capacity of Server 2 to Server 5 in 0 Time and 40 Time is shown in Fig. 9.

In the experiment, when the full node fails at 5 Time the transaction verification request of the ordinary node cannot be completed directly through the full node. The corresponding block should be downloaded to verify actively according to the result. The check period of the full node is set to 30 s, and it was

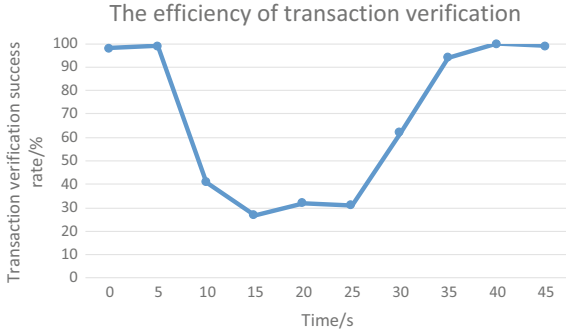


Fig. 8. The efficiency of transaction verification

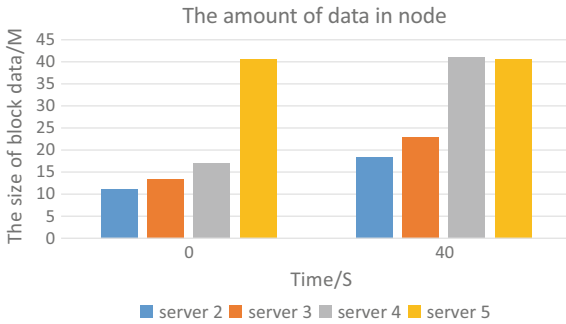


Fig. 9. The amount of data in node

found at about the Time of 25. The full node replacement strategy was started, and the new node replaced the faulty full node, and undertook the function of transaction verification. Therefore, the OMSCCB has a certain ability to resist the attack. The reliability of the OMSCCB can be guaranteed by selecting new nodes instead of fault nodes to maintain the normal operation of the system.

4 Conclusion

The storage capacity of blockchain node data is reduced significantly and the storage overhead of ordinary nodes is reduced in the OMSCCB. Meanwhile, the block data stored by ordinary nodes can be dynamically allocated according to their hardware performance thus balancing the storage load between nodes.

Because the OMSCCB has the full node replacement mechanism, it can resume normal operation in a short time in the face of an attack, and it has a certain anti-attack ability. In the future, we will further improve the reliability evaluation of nodes and optimize the segmentary scheme of blockchain to reduce the storage space under the premise of ensuring the data security and transaction efficiency.

References

1. Nakamoto, S.: Bitcoin: a peer-to-peer electronic cash system. *J. Gen. Philos. Sci.* (2008)
2. Azaria, A., Ekblaw, A., Vieira, T., et al.: MedRec: using blockchain for medical data access and permission management. In: *International Conference on Open and Big Data*. IEEE (2016)
3. Xu, T.: Research on the Development and significance of “Blockchain+” education. *J. Distance Educ.* **35**(02), 19–28 (2017)
4. Conoscenti, M., Vetr, A., Martin, J.C.D.: Blockchain for the Internet of Things: a systematic literature review. In: *Computer Systems and Applications*. IEEE (2017)
5. Jia, D., Xin, J., Wang, W., et al.: Storage capacity scalable model for blockchain (in chinese). *J. Front. Comput. Sci. Technol.* (2017)
6. Bruce, J.D.: Purely P2P crypto-currency with finite mini-blockchain (2013)
7. The Bitcoin Lightning Network. <http://lightning.network>
8. Karger, D., Lehman, E., Leighton, T., et al.: Consistent hashing and random trees: distributed caching protocols for relieving hot spots on the World Wide Web. In: *Proceedings of the Twenty-ninth Annual ACM Symposium on Theory of Computing*, pp. 654–663. ACM (1997)
9. Ziyang, B.A., Jun, W.U., Yan, M.A.: The optimization for consistent hash based on virtual node (in chinese). *Comput. Eng. Softw.* (2014)
10. Pei, X., Wang, Y., Ma, X., Xu, F.: A decentralized redundancy generation scheme for codes with locality in distributed storage systems. *Concurr. Comput. Pract. Exp.* (2017)
11. Courtois, N.T., Grajek, M., Naik, R.: *Optimizing SHA256 in bitcoin mining*. Springer, Berlin, Heidelberg (2014)
12. Cachin, C.: Architecture of the Hyperledger blockchain fabric. In: *Workshop on Distributed Cryptocurrencies and Consensus Ledgers* (2016)



Blockchain Storage Analysis and Optimization of Bitcoin Miner Node

Junying Gao, Bo Li, and Zhihuai Li^(✉)

Information College, Dalian Maritime University, Dalian 116000, China
{junying, qhlee}@dlmu.edu.cn, libo9281@163.com

Abstract. Aiming at the storage problem of blockchain, a space storage optimization scheme based on the InterPlanetary File System (IPFS) is designed. The scheme mainly achieves the purpose of space optimization by summarizing the blocks regularly, excluding the expired Transaction Output (TXO) that has no effect on verification. The Unspent Transaction Output (UTXO) will be saved as a file on the IPFS network during optimization. In this way, the miner node can quickly get the file and start mining. The ultimate goal is to optimize the storage space for the node and to increase the synchronization efficiency of the newly added miner node. The simulation experiment of the proposed optimization scheme is carried out. The experimental results show that the scheme can summarize the historical transaction data regularly, eliminate the Spent Transaction Output (STXO) which has no effect on the verification transaction, so as to effectively reduce the node storage space.

Keywords: Blockchain · Bitcoin · Storage optimization · Node classification

1 Introduction

Blockchain technology [1] has attracted much wider attention. Meanwhile, the number of users and trading volume of Bitcoin [2] has increased explosively. By November 2017, the Bitcoin blockchain space occupied more than 166G [3], as can be seen from the growth rate of the blockchain space. At present, space occupied by the Bitcoin blockchain has already become a problem making people anxious. At present, the storage space optimization for blockchain includes mini blockchain [4], segmented storage [5] and so on. By analyzing the problems of these schemes, based on the classic Bitcoin distributed architecture, this paper aims to design a block-space optimization scheme for miner nodes by file storage and forwarding function of the InterPlanetary File System and the idea of mini blockchain scheme. This key technique has been discussed and realized.

The research contents includes the following.

There are two kinds of node roles in the point-to-point network. The full node is to verify the transaction. The miner node is to generate blocks. Because

it uses less transaction information that has been spent, this part of transaction provides the possibility of spatial optimization.

Set the summary blocks and summarize the historical blocks periodically, extract and organize the unspent transaction output contained in, and clean up the useless transaction information for verification.

Calculate the efficiency of block optimization, extract effective UTXO information and store it in the interstellar file system IPFS network.

2 Blockchain Storage Optimization Scheme for Bitcoin Miner Node

2.1 Analysis of Work Flow of Miner Node

After the block is synchronized, the miner node has the qualification to participate in the mining. Miner node collects transaction information from the P2P [6] network, and finds out the data in the local block according to the UTXO referenced in the transaction, and determines whether the transaction is legal. Legitimate transaction data will be collected in the trading pool, waiting for the node to write to the latest block, so every legitimate transaction corresponds to a usable output from the previous transaction. It forms a chain of transactions. The structure is shown in Fig. 1.

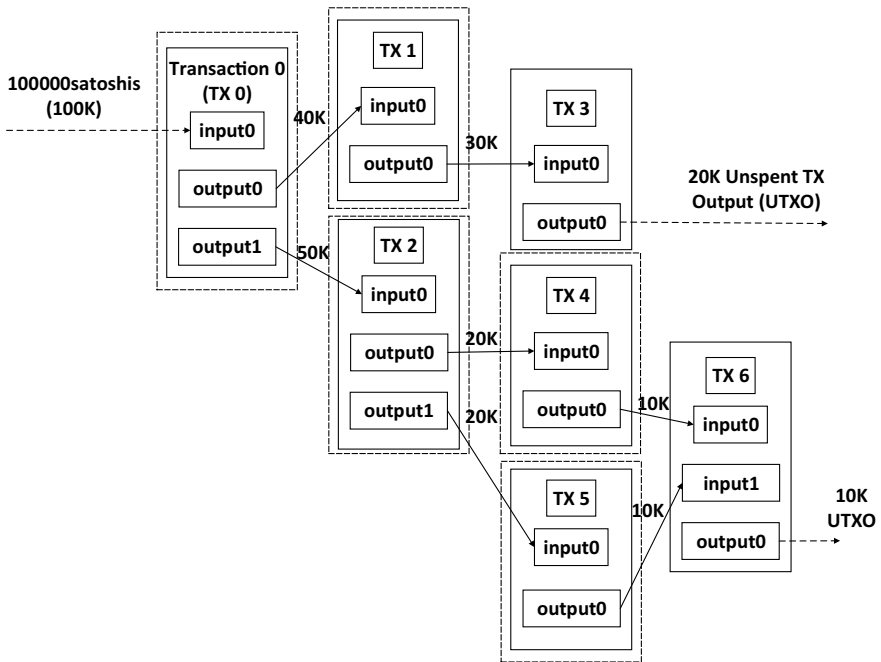


Fig. 1. Example of Bitcoin transaction structure

In the blockchain, such as 0, 1, 2, 4, 5, all the TXO has been quoted transaction information is not a small number. In essence, the miner node does not need this information in the process of verifying the transaction and calculating the block header hash, so it is possible to optimize the storage space of blockchain.

2.2 Design of Spatial Optimization Scheme

Overall Design. Based on the Bitcoin-distributed architecture [7], this paper designs a mechanism that can periodically summarize the historical blocks, extract and sort out the UTXO [8] contained therein, and clean up the useless transaction information. The overall structure of the program is shown in Fig. 2.

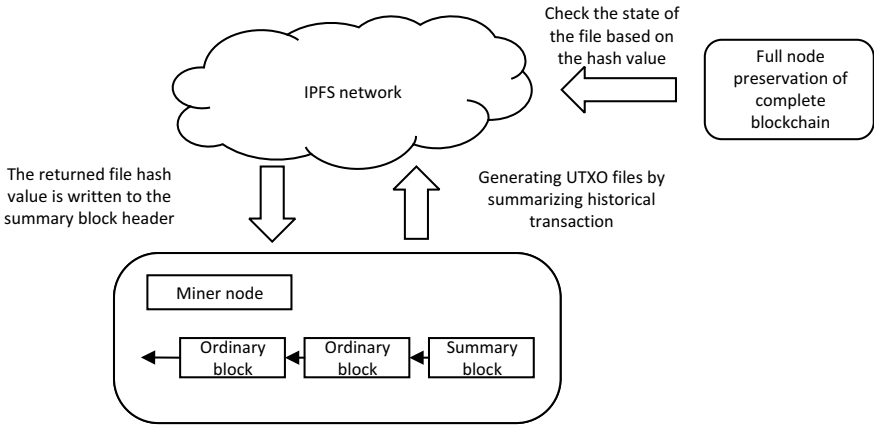


Fig. 2. The overall structure of the program

Figure 2 shows that the scheme adds the summary block. There is a certain span between the summing blocks. This periodic summary block assumes the task of regularly cleaning out the block information.

Summary blocks contain the ability to collate the currently available TXO. Miner nodes collect TXO write files, stored in the IPFS network. The hash value used to obtain the file is written to the block header of the summary block, and the work flow of the mining and broadcast summary blocks is the same as that of the ordinary block.

When other miner nodes receive the summing block, they need to download and verify the correctness of the UTXO files from the IPFS network. If the validation passes, the miner node continues to verify the transaction information contained in the summary block. The subsequent work flow is in line with the ordinary block validation process.

When the summary block is accepted by most nodes in the network and written to the local blockchain, it means that the TXO from the Genesis Block to the latest block has been finished. The node can choose to delete the block body

in the local saved blockchain according to its own situation. The UTXO data required by the miner node in the subsequent verification transaction has been saved in the IPFS network and downloaded to the local area when the summary block is verified. Therefore, the deletion of historical block has little effect on the miner node. The miner node can still verify the transaction information according to the UTXO file.

After the block body is deleted, the blockchain can still be connected to each other by the hash in the block header, forming a proof chain. Instead of completely removing all information from the historical block, this scheme chooses to retain the chain of proof and for the nodes, and it still retains the latest thousands of blocks. The latest blocks retained can be linked to the chain of proof to extend the Genesis Block, greatly increasing the cost of the attacker secretly forging a fake blockchain.

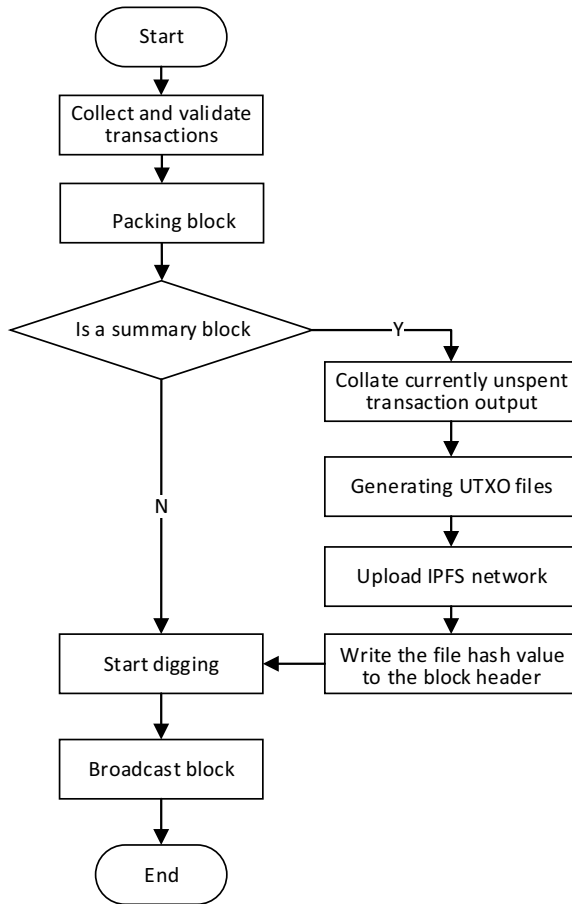


Fig. 3. The generation process of the block

Generation and Validation of Summary Block. Combined with the design idea of difficulty adjustment of Bitcoin, this scheme sets the same span value for summary block. The exact process for generating a block is shown in Fig. 3.

In the process of mining, it is found that the block with header hash is currently being calculated as a summary block, and the miner node first traverses the blockchain of the area, starting from the last summary block to the end of the latest block.

The specific optimization efficiency of blocks can be reflected by the formula and the formula is as follows.

$$E = \frac{\sum_i^n S_i}{\sum_i^n O_i} \quad (1)$$

E represents the optimization efficiency of the first n blocks; O_i represents the space occupied by the first block in the blockchain; S_i represents the space occupied by the UTXO transaction in the first n block. The denominator in the formula represents the sum of the space occupied by the first n blocks in the optimization. Molecules represent the sum of space occupied by the reserved transaction information after optimization. The division of the two is the optimal spatial efficiency.

The objective of the above-optimized operation is to extract all valid UTXO data from the Genesis Block to the current summary block. By calling the interface provided by IPFS, these data are stored into the network in the form of file. The miner node obtains a hash value pointing to the UTXO file, which can be downloaded from the IPFS network to the corresponding UTXO file. The miner node writes the hash value opposite the file to the header of the summary block. The summary block head records the hash field of the UTXO file, and ensures the authority of the UTXO file by using blockchain non-tampering.

In the mining process of calculating the block header hash, when the miner node gets a reasonable hash through a large number of calculations, they broadcast the summary block to the P2P network. The results are also announced for verification by other miner nodes. The process of the node verifying blocks is shown in Fig. 4.

UTXO File Storage. IPFS [9] is a P2P distributed hypermedia distribution protocol, whose main feature is content-based addressing. When obtaining a file, users no longer need to consider the location of the target server, name and storage path of the target files. For the above reasons, file storage and reading module of this project are completed by IPFS.

To generate UTXO, first, the node reads the UTXO file corresponding to the preceding summary block into memory, and then begins to read the transaction records in the block from the next block of the block. In the process of reading transaction records, constantly retrieving UTXO data in memory. When all the output of the retained transaction in memory has been referenced by the transaction in the subsequent block, the existence of the transaction has no effect on the subsequent process, and when it is deleted, the newly read transaction is

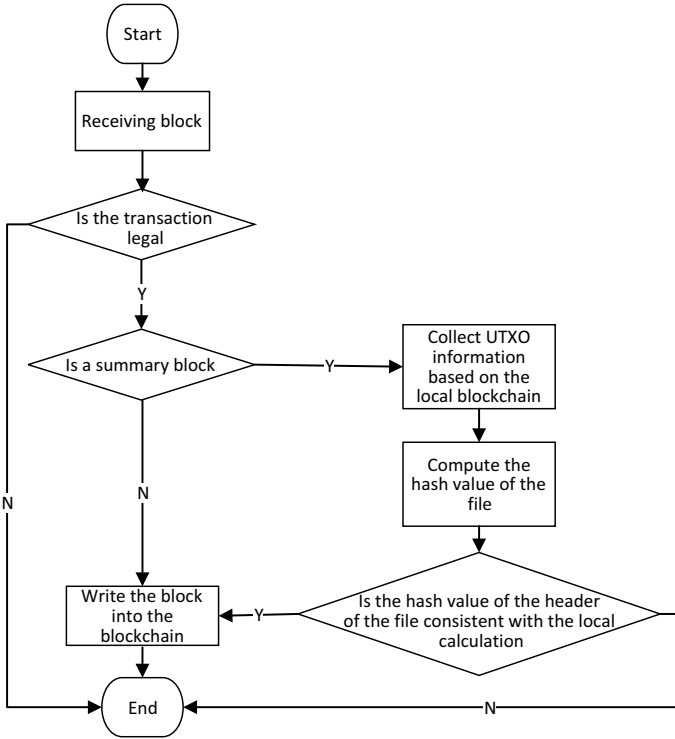


Fig. 4. The verification process of the block

added to the memory, and the subsequent process repeats the above steps, until the latest block is read. After viewing the block, the node writes the UTXO data collated in memory to the file. Next, the node goes to call the application development interface provided by IPFS to store the collated UTXO file. When the file is saved to the IPFS network, a unique hash value based on the file content is obtained.

Other nodes can use the hash value to initiate inquiring requests from the local distributed hash tables to the neighbor nodes, and keep approaching the target nodes that have the target files, and finally get the UTXO file information.

Chain of Proof. When the consensus is reached in the network, the miner node can choose to start the cleanup of the blockchain historical transaction. The main content of the cleanup is the historical transaction information. Because the available TXO information has been compiled into the UTXO file during the process of generating the summary block, historical transactions with more block confirmation can be cleaned up. The structure after cleaning is shown in Fig. 5.

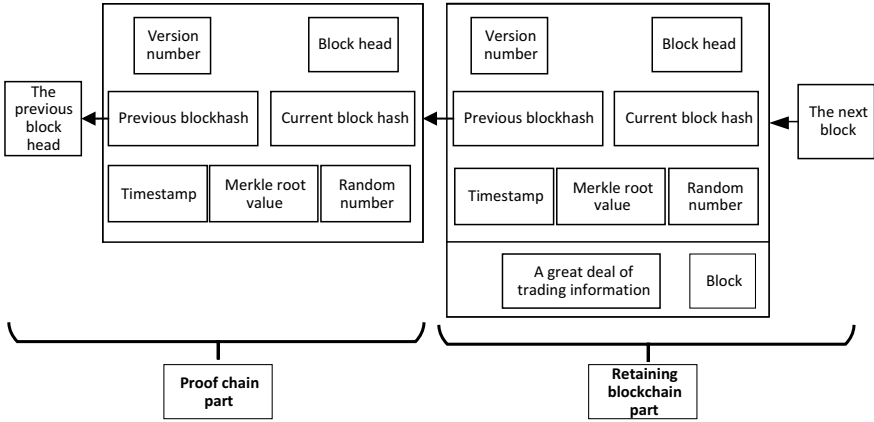


Fig. 5. The blockchain structure after cleaning

The chain structure of the block header in the figure is the part of the proof chain. Since the hash value that maintains the connection between blocks remains in the block header, the block that removes the block body can still be connected to the latest block retained. The chained structure can be traced back from the latest block to the recognized Genesis Block, so as to guarantee the authenticity of the transaction data in the current blockchain and resist the attacks such as secret mining in the network.

3 Experiment and Analysis

3.1 Introduction to the Experimental Environment

The Bitcoin trade statistics analysis experiment is written in the python language, whose version is 2.7.12. The database is Mysql database, whose version is 5.7.20. The Bitcoin client version is v0.15.0.0-g3751912e8e.

3.2 Transaction Data Statistics Experiment

In order to get real data of the Bitcoin blockchain, first install the Bitcoin core client and download Bitcoin blockchain complete data. Read operation block data by using Python language and port data through the bitcoinrpc expansion pack called Bitcoin core client PRC API [10]. Python Bitcoinrpc expansion pack can send the script to Bitcoin client to request and receive the data. The data storage module reads and stores data by connecting Mysql database with the Mysql database and the python expansion pack MySQLdb. Part of the data in the database tables is shown in Fig. 6.

block index	txnum	vin	vout
484951	310	523	1125
484952	1734	5509	4206
484953	2233	4967	5117
484954	2334	4634	6214
484955	1870	4880	5417
484956	2961	4754	6505
484957	2495	4688	5737
484958	2473	4667	6138
484959	2443	5037	5556
484960	1638	4705	4421
484961	1596	4965	4691
484962	2010	4614	4092
484963	1763	4794	4261
484964	2140	4395	4807
484965	1861	5271	3971
484966	948	4526	2056
484967	681	2106	1535
484968	1502	3312	3649
484969	491	1418	1167
484970	1568	4593	3714
484971	1780	3958	4499
484972	1618	3942	3955
484973	458	1199	944
484974	33	43	69
484975	1373	4143	3300
484976	492	1328	1137
484977	742	1460	1677
484978	1835	4735	4594
484979	1377	5240	2949
484980	1268	2808	3095

30 rows in set (0.00 sec)

Fig. 6. Table structure

After reading, the database contains 495,000 pieces of data, corresponding to the Bitcoin blockchain from the creation block to the 495,000 block.

Bitcoin Transaction Output Statistics Experiment. The contents include the trade output in the Bitcoin trading history and the STXO. This experiment takes one thousand blocks as the basic unit. Statistics on the number of total and all available trade outputs at different stages of the Bitcoin blockchain start from the Genesis Block. Statistics on historical Bitcoin trade output are shown in Fig. 7.

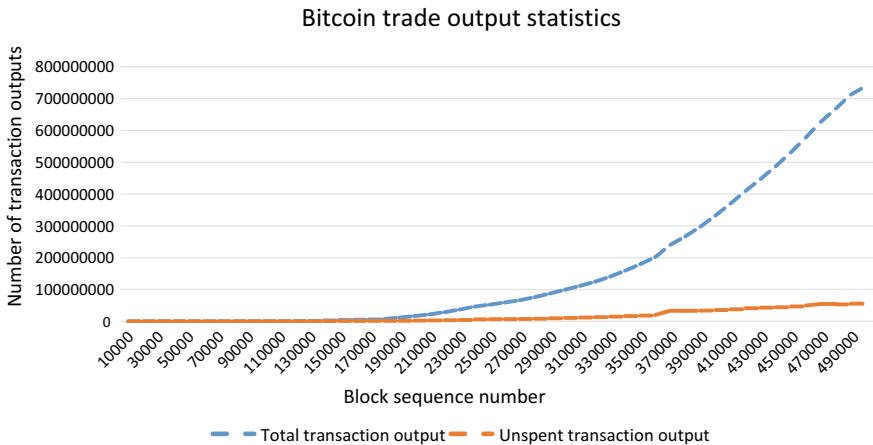


Fig. 7. The statistical graph of output

The curve in the chart shows that from the appearance of the Genesis Block to the 190,000th or so, the volume of Bitcoin transactions is relatively small. Starting nearby the 190,000th block, the trade output increases gradually and shows a trend of accelerating growth. Compared with the total volume of trade output, the UTXO grows more slowly. As subsequent blocks continued to emerge, the gap between the historical total TX0 and the available UTXO is gradually widening, so the assumption that the Bitcoin blockchain contains plenty of optimizable space is valid.

The STXO Proportional Statistics Experiment. The purpose of the experiment is to calculate the STXO ratio under different spans, and finally select the appropriate value of the span to carry on the follow-up experiment. According to the theoretical derivation of this paper, The STXO is an optimization space. The final result value is obtained by adding the ratio of saving space after recording each summary and calculating the average value. The peculiar distribution of the scale is shown in Fig. 8.

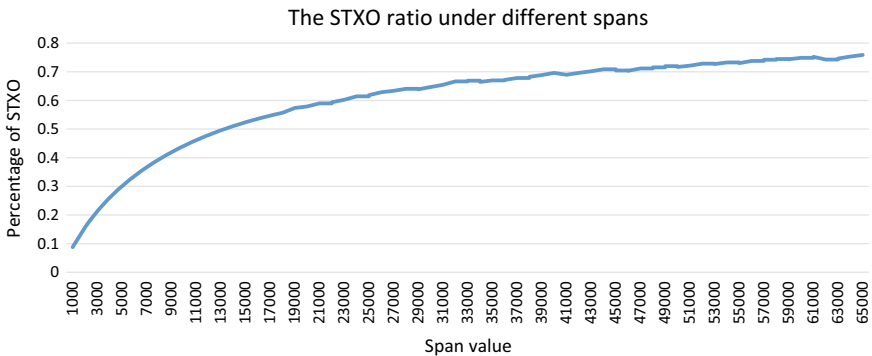


Fig. 8. Comparison of storage conditions

According to the curve of the Bitcoin real transaction data analysis chart, when the interval of summary blocks is about 25,000 blocks, the system has a more appropriate average optimization efficiency. The average optimization efficiency is about 60%. For bitcoin systems that have 8 years of operation and have produced up to 500,000 blocks, it is acceptable to sum up and clean up Bitcoin systems within half a year.

3.3 Storage Optimization Experiment

The purpose of the experiment is to verify the effect of the improved scheme. The experiment has the same data set and carries out the original Bitcoin scheme and the improved scheme separately. In the optimization scheme, the value of

block span is selected as 25,000. The collection of experimental data is mainly accomplished by recording the space occupation of block files on the client side of the two schemes in the disk periodically. The comparison of the changes of storage space data with the increase number of the blocks in the two groups is shown in Fig. 9.

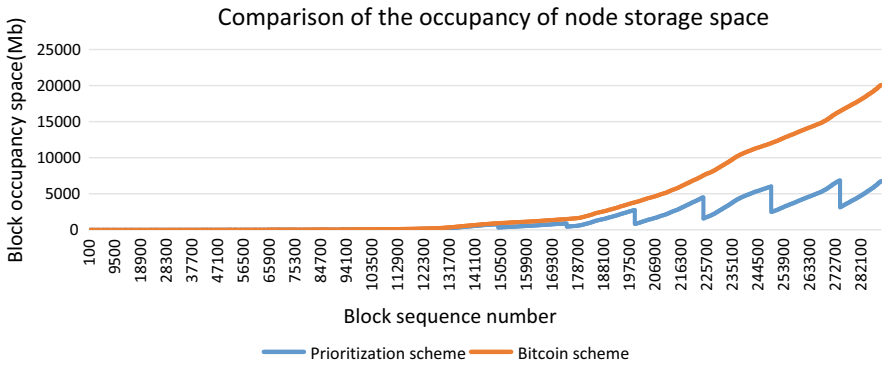


Fig. 9. Comparison of storage conditions

By observing the growth trend in dashed line, we can see that with the continuous production of blocks, the historical transaction information increases gradually and presents the trend of acceleration after more than 200,000 blocks. From the observation of real line segments, we can see that when the number of blocks is more than 150,000, the number of transaction increases rapidly and space usage of blocks in the Bitcoin scheme accelerate. But space usage of magnetic disk in prioritization scheme increases relatively slowly, and every 25,000 blocks in the line segment produce a decline. There are 11 summary blocks produced in this experiment and the average ratio of deleting data in each summary period is about 49.55. The final blockchain data occupies about 19.6 Gb in Bitcoin scheme and 6.6 Gb in optimization scheme. Therefore, as the blocks continue to grow, the optimization scheme can effectively reduce the space occupation of blockchain.

4 Conclusion

In the statistic analysis experiment of Bitcoin transaction data, it confirms the inference that a large number of optimal data is included in the blockchain.

In the capacity optimization experiment, the proposed scheme can reduce the occupied space of blockchain significantly. In the simulated environment, the transaction historical information can be summarized periodically and the UTXO file can be generated correctly and uploaded to the IPFS network. In normal mining, miner node that has cleaned up historical trading data are still able to verify the correctness of the transaction.

Therefore, by analyzing the simulation experiments in this paper, we can see that the spatial optimization of miner node has certain effect.

References

1. Bonneau, J., Miller, A., Clark, J., et al.: Research perspectives and challenges for bitcoin and cryptocurrencies. In: *Security and Privacy*, pp. 104–121 (2015)
2. Bhme, R., Christin, N., Edelman, B., et al.: Bitcoin: economics, technology, and governance. *J. Econ. Perspect.*, 213–238 (2015)
3. Blockmeta: The Blockchain Data of Bitcion. <http://blockmeta.com//btc-stat>
4. Bruce, J.D.: *Purely P2P Crypto-Currency with Finite Mini-Blockchain* (2013)
5. Jia, D., Xin, J., Wang, Z., et al.: Storage capacity scalable model of blockchain (In Chinese). *Comput. Sci. Explor.*, 1–10 (2017)
6. Donet, J.A.D., Prez-Sol, C., Herrera-Joancomart, J.: The bitcoin P2P network. In: *The Workshop on Bitcoin Research*, pp. 87–102 (2014)
7. Jiang, R., Wei, C.: Application and value of blockchain (In Chinese). *Gansu Financ.* **19–21** (2016)
8. Nakamoto, S.: *Bitcoin: a peer-to-peer electronic cash system*. Consulte (2008)
9. Curran, T., de Graaff, B.: Analysing the Performance of IPFS During Flash Crowds (2016)
10. Djilali, S.: P2P-RPC: programming scientific applications on peer-to-peer systems with remote procedure call. In: *IEEE/ACM International Symposium on Cluster Computing and the Grid*, pp. 406–413. Tokyo (2003)



Improved Ant Colony Optimization Algorithm for Optimized Nodes Deployment of HAP-Based Marine Monitoring Sensor Networks

Jianli Duan^{1,2(✉)}, Yuxiang Liu¹, Bin Lin^{1(✉)}, Yuan Jiang³, Fen Hou⁴,
and Wantong Li¹

¹ Dalian Maritime University, Dalian 116026, China

duanjianli@qut.edu.cn, binlin@dlmu.edu.cn

² Qingdao University of Technology, Qingdao 266033, China

³ University of Chinese Academy of Science, Beijing 100049, China

⁴ University of Macau, Macau 999078, China

Abstract. Territorial ocean safety and ocean development make it important to establish a large-scale, long-term, and low-energy integrated ocean monitoring sensor network (OMSN). In this paper, we introduce the high attitude platform-based ocean monitoring sensor network (HAP-OMSN) architecture and the basic ant colony optimization (ACO) algorithm first. And then, we propose an improved ant colony optimization algorithm for the node deployment of the HAP-OMSN architecture. Finally, we solve the multi-types node deployment (MTND) problems in HAP-OMSN using this algorithm. The final experiment results indicate that the improved ACO algorithm has good efficiency to find optimal solution.

Keywords: Ant colony optimization (ACO) · Ocean monitoring sensor network (OMSN) · The high attitude platform-based OMSN (HAP-OMSN)

1 Introduction

With the rapid development of intelligent ocean engineering, network deployment has become an important issue in the ocean monitoring sensor networks (OMSN). The integrated HAP-OMSN as shown in Fig. 1, which can reduce the cost, improve the coverage of the network, and increase data transmission speed, is a useful network architecture for remote OMSN [1]. Based on the HAP-OMSN [1] architecture, optimizing the path of information transmission can reduce power consumption and prolong the network lifetime. The ant colony optimization (ACO) algorithm which has been proposed in the early 1990s by Italy scholars Colormi et al. is a good choice to solve the path optimization problem; ACO is a population-based heuristic bionic evolution algorithm through studying foraging behavior of real ants [2]. Although ACO has the advantages of universality and global convergence, its convergence time is longer and its search is prone to “stagnation”. Therefore, improving ant colony algorithm has become a hotspot of researchers.

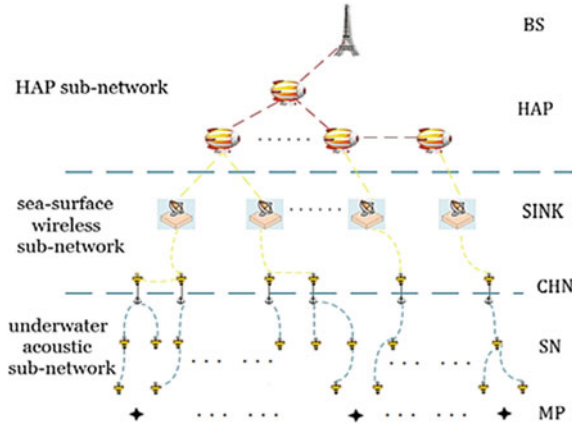


Fig. 1. The multitiered architecture of HAP-OMSN

The research in improving ACO attracts great interest of researchers. They have done a lot of work to improve ACO in the path selection, pheromone update criteria, local search and global search, speed of convergence, and so on. German scholars Stutzle T. and Hoos J. propose the MAX-MIN ant system (MMAS) and local search for the traveling salesman problem (TSP) [3], which improve the search at the initial stage of ACO by limiting the range of pheromone concentration. Meng X., Huang T. and Chen S present an improved ACO based on pheromone updating and evaporation factor adjusting [4]. For the problem such as slow convergence and being easy to fall into local optimum of classical ACO, Li L. and Yu H. propose an improved algorithm for robot path planning in complex environment [5]. Ref. [6–8] propose corresponding improved algorithms according to the characteristics of the environment in the path planning problem. However, ACO suffers from some problems. In this paper, we propose an improved heuristic algorithm based on ACO to solve the multi-types node deployment (MTND) problems in HAP-OMSN; we hope to obtain a near-optimal solution in large-scale network scenario.

The remainder of the paper is organized as follows. In Sect. 2, we describe the ant colony algorithm. In Sect. 3, we show the improvement of ant colony algorithm. The simulation results are also acquired in Sect. 3. Finally, we give the conclusion of the research in Sect. 4.

2 The Basic ACO Algorithm

ACO algorithm has strong robustness and is easy to be combined with other methods in optimization, but it has the limitation of stagnation and is easy to fall into local optimums [9]. ACO algorithm is the first to be used to solve traveling salesman problem, which is a very famous NP-hard problem. Traveling salesman problem can be described as follows: a salesman is going to N cities to sale good, and he starts from

one city to other cities through one shortest path, in addition, each city can be passed only once.

The processing of ACO is introduced as following based on the most basic traveling salesman problem. (1) Initialization: M ants are put in N cities at random. Paths exist among every city and its adjacent cities, and the initial pheromone trails of these paths are set to the same concentration; (2) path selection: ant selects the next path according to the transition probability; (3) the pheromone trail update: the pheromone trails are updated by the ant system while each ant goes through N time points (TPs); and (4) repeat the above three steps until the desired result is reached.

In general, the k th ant move from city i to city j with probability

$$P_{ij}^k(t) = \begin{cases} \frac{\tau_{ij}^\alpha(t)\eta_{ij}^\beta(t)}{\sum_{s \in \text{allowed}_k} \tau_{is}^\alpha(t)\eta_{is}^\beta(t)} & j \in \text{allowed}_k \\ 0 & \text{otherwise} \end{cases}$$

where $\tau_{ij}(t)$ is the amount of pheromone deposited for transition from city i to j at TP t , $0 \leq \alpha$ is a parameter to control the influence of $\tau_{ij}(t)$, $\eta_{ij}(t)$ is heuristic factor for transition from city i to j at TP t (typically $\eta_{ij}(t) = 1/d_{ij}$, where d is the distance), and β is a parameter to control the influence of $\eta_{ij}(t)$.

$$\tau_{ij}(t + n) = \rho\tau_{ij}(t) + \Delta \tau_{ij} \tag{1}$$

$$\Delta \tau_{ij} = \sum_{k=1}^m \Delta \tau_{ij}^k \tag{2}$$

$$\Delta \tau_{ij}^k = \begin{cases} \frac{Q}{L_k} & \text{if } k\text{th ant uses edge}(i, j) \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Equations (1)–(3) express the update rules of the pheromones concentration. $\Delta \tau_{ij}$ is the amount of information which are left by ants upon the path from city i to j at TP $t + n$; ρ represents the degree of attenuation of pheromones; $\Delta \tau_{ij}^k$ shows the amount of information which are left by the k th ant upon the path from city i to j at TP $t + n$; and Q is a constant and L_k is the cost of the k th ant’s tour (typically length).

3 The Improved Ant Colony Algorithm for MTND

3.1 MTND for the HAP-OMSN

The inputs, objects, and constraints of MTND are as follows.

Inputs:

- (1) The amount of data produced at MPs one day.
- (2) The location of the land BS, HAPs and MPs.

- (3) The location of the candidate points (CPs) for the SINKs, CHNs, and SNs.
- (4) The effective sensing radius of the SN.
- (5) The receiving radius of the HAP, SINK, and CHN.
- (6) The capacity of the CHN.
- (7) The maximum number of hops.
- (8) The HAPS' communication transmission power and radio reception power.
- (9) The radio reception power and transmission power of the CHN, underwater acoustic reception power.
- (10) The reception power and transmission power of the SN.
- (11) The sending parameters of CHN and SN.
- (12) The initial energy of HAPS, CHN, and SN.

Objects:

- (1) To minimize F, and the smaller the value of F, the longer the network lifetime.
- (2) To minimize the energy consumption of the SNs, CHNs, and HAPs per day.

Constraints [1]:

- (1) Coverage: Each MP should be covered by the SN.
- (2) Connection: HAP-OMSN should transmit the perceived data to BS.
- (3) Topology: HAP-OMSN should keep the tree structure. Loop is not allowed.
- (4) Delay: The maximum number of multi-hops from MP to BS is upper bounded.
- (5) Reliability: Data can be transmitted to CHN through two different Paths.

3.2 The Improved Ant Colony Algorithm

Many scholars have proposed improvement methods for ACO, such as ant colony system, MAX-MIN ant system, mutation ant colony algorithm, and so on [10]. These algorithms improve the search capability of AC, but cannot solve the MTND problem well, especially in lifetime. For example, the original intelligent factor will make all ants choose the optimal path, and also represent that the nodes on the optimal path will become the hotspots, which will lead to premature death of hot nodes, and as a result, the best choice of individuals will bring about the worst consequence to the collective. In this paper, reasonable improvements on the basis of AC are proposed to match the mathematical models established before.

The improved AC mainly improves the following four aspects:

- (1) Modify the intelligence factor

θ_{ij} is used to represent the intelligence factor instead of η_{ij} according to the real node energy consumption model, and the local update formulas of pheromone is set. The improved formulas are as follows:

$$p_{ij}^k = \begin{cases} \frac{[\tau(i, j)]^\alpha \theta_{ij}^\beta}{\sum_{s \in J_k(i)} [\tau(i, s)]^\alpha \theta_{is}^\beta} & j \in J_k \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

$$\tau(i, j) = (1 - \rho)\tau(i, j) + \rho \Delta \tau(i, j) + \rho \Delta \tau'(i, j) \quad (5)$$

$$\Delta \tau(i, j) = q_1 T_life \quad \text{if}(i, j) \quad j \in J_k \quad (6)$$

$$\Delta \tau'(i, j) = q_2 T_{best_life} \quad \text{if}(i, j) \quad j \in T_{bestpath} \quad (7)$$

$$P_j = \begin{cases} P_{ST} d_{jk}^{k_2} 10^{\frac{\alpha(f)}{10} d_{jk}} & j \in \{SN\} \\ \beta d_{ij}^{k_1} & j \in \{CHN\} \end{cases} \quad (8)$$

$$\theta_{ij} = \frac{E_i E_j}{P_i} \quad (9)$$

The initial value of $\tau(i, j)$ is set to 1. Equation (6) indicates all paths ants go by should be updated while net is formed. Equation (7) shows the update of the best paths. Equation (8) expresses the power of node i . Equation (9) denotes that ants incline to move to those nodes that can maximize their lifetime. T_life and T_best_life are the lifetime and the best lifetime of networks, respectively. E is the node energy.

(2) Modify the pheromone updating mode

The pheromones update mode is related to transfer probability. For each node, if the transition probability is greater than 85%, the pheromone of the path will not increase, and if the transfer probability is less than 10%, the pheromone of the path will be increased by 2 T_life . This pheromone update mode can slow down the convergence rate of the algorithm and make the whole network converge.

(3) Modify the random disturbance mode of path selection

The iterative termination condition we use is keeping the best network in the 600 generation. The adaptive disturbance is adopted and the probability formula is show in Eq. (10), where U is the number of iterations.

$$P = 0.3 * U/600 \quad (10)$$

(4) Modify the ant movement rules

In each iteration, ants will be placed at monitoring point (MP) and stop when they arrive at BS. We use a feasible path table to specify the travel direction of ants to avoid blinding search and breaking the multi-hop rule.

3.3 The Process of the Improved ACO Algorithm

The improved ACO algorithm can be described as follows:

Procedure: MTND-IACO

Input: coordinates of network nodes and CPs, other parameters.

Output: Optimal solution

INIT. The coordinates of HAP, MP, CPs for SINK, CPs for CHN, and CPs for SN, $\alpha, \beta, \rho, q_1, q_2, T_{life}, T_{bestlife}, D, E, P$, the pheromone matrix: PH, the probability transfer matrix: PR, the node selection matrix: ND and the path selection matrix: PT.

NETWORKING. Nodes enter the network in forward direction and opposite direction.

TABLE-BUILDING. Fill nodes of the network into the feasible path table

While (NOT satisfy stop condition)

```
{
    precluster_node_path_selection; /*select SN and
    CHN nodes from CPs and build paths from MP to CHN*/
```

```
    preHAP_node_path_selection; /*select SINK and
    HAP nodes from CPs and build paths from CHN to HAP */
```

```
    node_complectpath_selection; /*fulfill the
    node selection and complete pathway selection*/
```

```
    Update the iteration pheromone and global opti-
    mality pheromone, update PR
```

```
}
```

Output the optimal solution.

3.4 Computational Results

To validate the improved ACO for MTND, we solve the problem in large-scale scenarios using both gurobi and our algorithm. Table 1 shows the size of one scenario, and the result is shown in Figs. 2 and 3.

Table 1. Size of the scenario 3

	CPs for SN	CPs for CHN	CPs for SINK	MP
Quantity	8	36	79	16

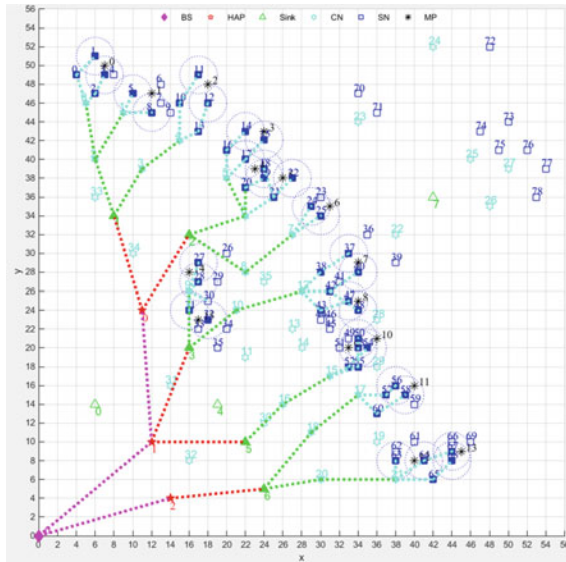


Fig. 2. Gurobi solving result diagram

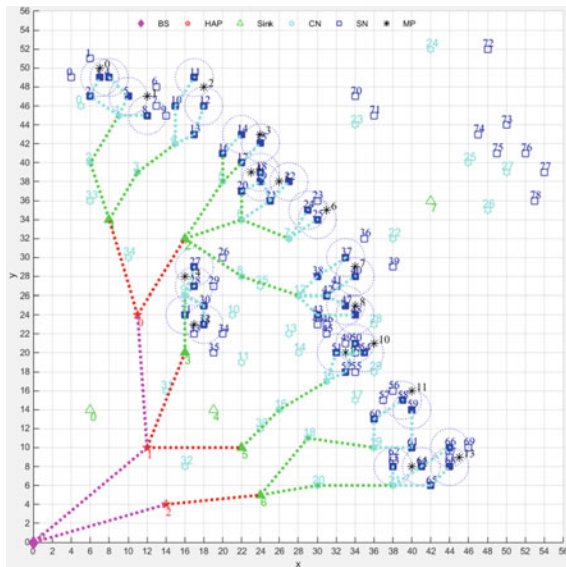


Fig. 3. The improved ACO solving result diagram

Compare Fig. 3 with Fig. 2, the backbone network part of algorithm optimization is the same as that of gurobi, but the network lifetime reduces from 116 (by gurobi) to 107, and the energy consumption has also increased from 188.6 (by gurobi) to 199.9.

Therefore, the network deployment result of the improved ACO is a suboptimal solution.

Based on the validation above, we make further efforts to examine the efficiency of the algorithm. Table 2 exhibits optimization results in several increasing scale scenarios. These results show that the optimization time of this algorithm is less than that of gurobi especially in large-scale scenarios; in addition, it can solve larger scenarios that gurobi cannot handle, such as scenario 5. As shown in Fig. 4, we can see clearly that the optimization time of the improved ACO is controlled within $O(n)$. Therefore, it has high efficiency in solving MTND of HAP-OMSN.

Table 2. Optimization time comparison of gurobi and the improved ACO

Scenario	Quantity				Running time(S)	
	CPs for SINKs, CHs, SNs	BSs, HAPs, MPs	Selected SINKs, CHs, SNs		Gurobi	Algorithm
			Gurobi	Algorithm		
1	4, 16, 30	1, 2, 6	3, 8, 12	3, 8, 13	5.89	0.9
2	6, 24, 52	1, 3, 10	4, 13, 30	4, 13, 31	61.17	3.2
3	8, 36, 79	1, 3, 16	5, 18, 46	5, 16, 45	1226.9	8.7
4	9, 38, 88	1, 3, 18	5, 19, 53	6, 21, 51	202436.7	3.3
5	12, 50, 120	1, 5, 24	Unsolvable	8, 23, 67	Unsolvable	10.2

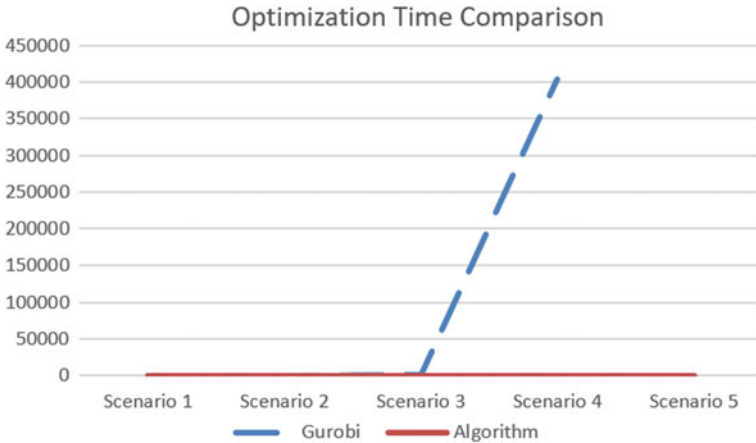


Fig. 4. Optimization time comparison of gurobi and the improved ACO

4 Conclusion

In this paper, the basic ACO and related works are introduced first, and then, an improved ACO algorithm is proposed to solve MTND of HAP-OMSN networks. The improvement method and process of this algorithm are indicated in detail. The final

experiment results indicate that the new algorithm has good efficiency to find optimal solution. In future research, ACO can be combined with other bionic algorithms to make the optimization algorithm more robust.

Acknowledgements. This study is sponsored by National Science Foundation of China (NSFC) No. 61371091 and No. 61301228, Liaoning Provincial Natural Science Foundation of China No.2014025001, and Program for Liaoning Excellent Talents in University (LNET) No. LJQ2013054.

References

1. Liu, Y.: Research on Optimized Nodes Deployment of HAP—Based Marine Monitoring Sensor Networks. Dalian Maritime University, Dalian (2018)
2. Gan, R., Guo, Q., Chang, H., et al.: Improved ant colony optimization for the traveling salesman problems. *J. Syst. Eng. Electron.* **21**(2), 329–333 (2010)
3. Stutzle, T., Hoos, H.: The MAX-MIN ant system and local search for the traveling salesman problem. In: *IEEE 4th International Conference on Evolutionary Computation*, pp. 308–313 (1997)
4. Meng, X., Huang, T., Chen, S.: Improved ant colony optimization algorithm based on pheromone updating and evaporation factor adjusting. *J. Chengdu Univers. Nat. Sci. Ed.* **34**(1), 48–51 (2015)
5. Li, L., Yu, H.: Improved ant colony algorithm in complex environments on the robot path planning. *J. Chin. Comput. Syst.* **38**(9), 2067–2071 (2017)
6. Xiang, U., Liang, Z., Wei, Z., et al.: Dynamic path planning in RoboCup rescue simulation competition. In: *The 27th Chinese Control and Decision Conference*, pp. 4341–4344 (2015)
7. Qu, H., Huang, L., Ke, X.: Research of improved ant colony based robot path planning under dynamic environment. *J. Univ. Electron. Sci. Technol. China.* **44**(2), 260–265 (2015)
8. Liu, J., Yan, Q., Ma, Y., et al.: Global path planning based on improved ant colony optimization algorithm for geometry. *J. Northeast. Univ. (Nat. Sci.).* **36**(7), 923–928 (2015)
9. Duan, H., Wang, D., Zhu, J., et al.: Development on ant colony algorithm theory and its application. *Control Decis.* **19**(12), 1321–1320 (2004)
10. Fu, Y.: *The Improvement and Application of Ant Colony Algorithm*. Shanghai Maritime University, Shanghai (2006)



A Resource Allocation Algorithm for D2D Multicast in Internet of Vehicles

Wei Wu^{1(✉)}, Muchen Yu¹, Xuanli Wu¹, and Guoan Bi²

¹ School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150080, China

kevinking@hit.edu.cn

² School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, Singapore

EGBI@ntu.edu.sg

Abstract. D2D multicast communications can reduce communication cost and decrease interference in Internet of Vehicles (IoV). In this paper, the resource allocation algorithm for Device-to-Device (D2D) multicast communication reusing LTE network uplink resource is studied to improve the reliability of networks. The optimization objective of the algorithm is to minimize the outage probability of D2D multicast group links on the basis of guaranteeing the constraint of cellular link quality. To solve this optimization problem, a low-complexity heuristic algorithm is designed for spectrum resource allocation and power control. By comparing performance through simulation, it can be found that the outage probability of the heuristic algorithm designed in this paper can be very close to that of the Hungarian algorithm, but the computational complexity is greatly reduced compared with the latter.

Keywords: D2D multicast · IoV · Resource allocation · Heuristic algorithm

1 Introduction

As a part of 5G mobile communication, IoV has become a hot research topic. Communications in IoV require low delay and high reliability. D2D communication in IoV not only can reduce the load of the base station but also can largely increase the capacity of communication system, reduce the transmission delay, improve the cell edge user performance, and enlarge the coverage of network [1]. For information distribution in IoV, multicast communication can reduce the cost and interference. It can be seen that the application of D2D multicast communication in IoV has certain advantages compared with other vehicle communication methods.

About the application of D2D multicast technology and D2D technology in IoV, some scholars have done some relevant works. Literature [1] grouped D2D vehicles according to the geographic location, maximizing the sum rate through channel reuse and power control. Literature [2] proposed a heuristic resource allocation mechanism related to vehicle location information for D2D communication in IoV, which not only met the requirements of communication security service but also reduced signaling cost and communication interference. In literature [3], a joint power and channel allocation

optimization scheme for D2D multicast cluster, based on the social attributes and geographical factors, was proposed by extending the bipartite graph matching algorithm. In literature [4], a resource allocation algorithm was present to improve the system capacity, according to the number of users and SINR (Signal to Interference pulse Noise), a channel allocation optimization scheme based on distance constraint is proposed to reduce the complexity.

The above researches provide a good theoretical basis for the application of IoV. However, few of them focus on the reliability of the network. In this paper, we will study the resource allocation optimization for D2D multicast in IoV to improve the system reliability.

2 System Model and Problem Formulation

Figure 1 is a scenario of D2D multicast communication in IoV, D2DM (D2D Multicast groups) mainly exists in the communication between vehicles (V2V) [5]. In this paper, D2DM communications reuse uplink resources of CUs (Cellular Users). There are a set $C = \{1, \dots, C\}$ of CUs and a set $K = \{1, \dots, K\}$ of D2DMs in one cell. The interference is divided into two types: BS (Base Station) is interfered by D2DM transmission, and D2DM receivers are interfered by uplink signals of CUs.

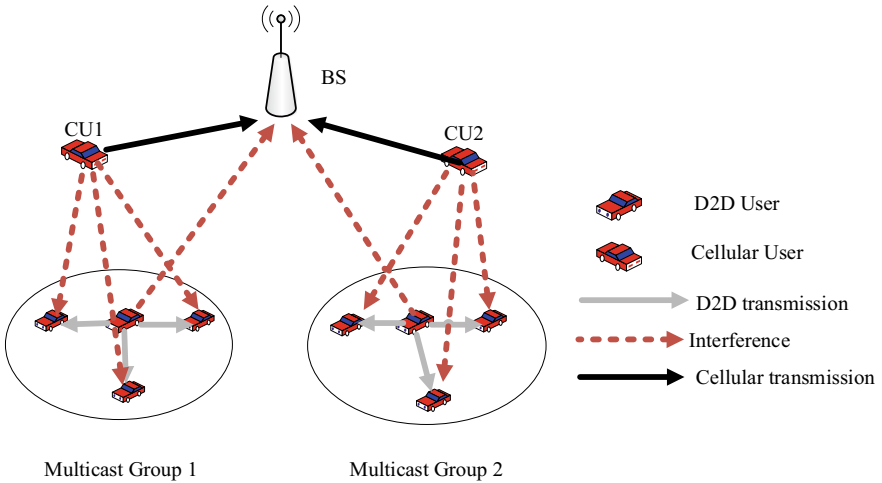


Fig. 1. D2D multicast communication in IoV.

The SINR of CUs' uplink is denoted as

$$\gamma_c = \frac{p_c G_c^b}{n_0 B_0 + \sum_{k=1}^K y_{k,c} p_k G_{k,m}^b} \tag{1}$$

where the p_c , p_k are the transmitted power of c th CU and k th D2DM senders, respectively; n_0 is the additive white Gaussian noise which is the distribution $N(0, n_0)$;

B_0 is the system bandwidth; and $y_{k,c}$ is a binary variable which equal to 1, if the k th D2DM reuses the c th CU's Resource, else $y_{k,c} = 0$; $G_c^b, G_{k,m}^b$ are the link gains from c th CUs and k th D2DM sender to BS, respectively.

Similarly, the SINR of D2D receiver is denoted as

$$\gamma_d = \frac{p_k G_{k,d}^{D2D}}{n_0 B_0 + \sum_{c=1}^C y_{k,c} p_c G_{k,c,d}^{C2D}} \tag{2}$$

where $G_{k,d}^{D2D}$ is the gain of link from k th D2DM sender to the D2D receivers in the same group; $G_{k,c,d}^{C2D}$ is the gain of link from c th CUs and one of k th D2DM receivers.

In order to guarantee the Quality of Service (QoS) of D2D multicast communication network, the SINR requirements of cellular and D2D links are defined as γ_{cmin} and γ_{dmin} , respectively. The corresponding rate requirements are R_{cmin} and R_{dmin} . Obviously, when the SINR of a cellular link is less than SINR requirements, the link is interrupted. According to Lemma 1 in [6], P_{cout} and P_{dout} , the outage probability expressions of the two kinds of links can be obtained as

$$P_{cout} = 1 - \frac{p_c (d_{c,b})^{-\alpha}}{p_c (d_{c,b})^{-\alpha} + \sum_{k=1}^K y_{k,c} p_k (d_{D_t,b})^{-\alpha} \gamma_{cmin}} \exp\left(-\frac{n_0 B_0 \gamma_{cmin}}{p_c (d_{c,b})^{-\alpha}}\right) \tag{3}$$

$$P_{dout} = 1 - \frac{p_k (d_{D_t,D_r})^{-\alpha}}{p_k (d_{D_t,D_r})^{-\alpha} + \sum_{c=1}^C y_{k,c} p_c (d_{c,D_r})^{-\alpha} \gamma_{dmin}} \exp\left(-\frac{n_0 B_0 \gamma_{dmin}}{p_k (d_{D_t,D_r})^{-\alpha}}\right) \tag{4}$$

where $d_{c,b}, d_{D_t,b}, d_{D_t,D_r}, d_{c,D_r}$ are the distances between CU and BS, D2DM sender and BS, D2DM sender and D2DM receiver, and CU and D2DM receiver, respectively.

Improving the reliability of the D2D multicast of IoV can be transformed to minimize the outage probability of the D2D multicast links. As a result, the problem can be formulated as

$$\begin{aligned} & \min \sum_{c=1}^C \sum_{k=1}^K y_{k,c} P_{dout} \\ & \text{s.t.} \left\{ \begin{array}{l} p_{kmin} \leq p_k \leq p_{kmax}, \forall k \in K \\ p_{cmin} \leq p_c \leq p_{cmax}, \forall c \in C \\ \sum_{k=1}^K y_{k,c} \leq 1, \forall c \in C \\ \sum_{c=1}^C y_{k,c} = 1, \forall k \in K \\ P_{cout} \leq P_{cth}, \forall c \in C \\ y_{k,c} \in \{0, 1\}, \forall k \in K, c \in C \end{array} \right. \tag{5} \end{aligned}$$

3 Resource Allocation Algorithm for D2D Multicast in IoV

The problem above is a Mixed Integer Nonlinear Programming (MINLP) problem, which is a channel and power allocation combinatorial optimization problem. The complexity is very high. Therefore, the problem is divided into two parts: the first part is power control for CUs and D2D users, the second part is channel allocation with the assumption that D2DMs transmit at the controlled power.

3.1 Power Control Algorithm

The power control of D2D multicast communication in IoV is a nonlinear programming problem. In order to solve this problem, the feasible region needs to be analyzed. The constraint can be divided into two parts: power constraint and rate constraint, where power constraint is expressed as $\Omega_1 = \{(p_c, p_k) : p_{cmin} \leq p_c \leq p_{cmax}, p_{kmin} \leq p_k \leq p_{kmax}\}$, and rate constraint is expressed as $\Omega_2 = \{(p_c, p_k) : \gamma_c \geq \gamma_{cmin}, \gamma_d \geq \gamma_{dmin}\}$, then the feasible region is $\varphi = \Omega_1 \cap \Omega_2$ which is shown in Fig. 2.

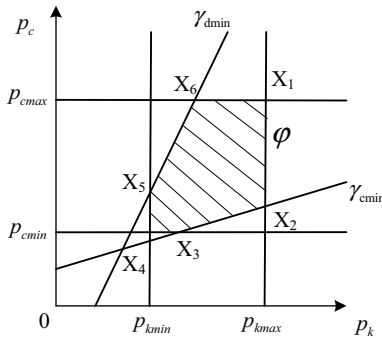


Fig. 2. The feasible region of power control.

We introduce Lemma 2 in literature [7]. According to this lemma, by traversing all corner points $\{X_1, X_2, X_3, X_4, X_5, X_6\}$ in Fig. 2, the point at which the objective function of Eq. (5) is minimized can be obtained, and the power value of the corresponding CU and D2D users (p_c^*, p_k^*) is acquired.

3.2 Channel Allocation Algorithm

To solve the problem of channel allocation, Hungarian algorithm is a good way. However, the complexity of Hungarian algorithm is very high. In consequence, a heuristic algorithm is proposed in this paper. The core idea of this algorithm is to choose CU whose link quality is higher and the D2DM sender reuses the channel of this CU. The description of heuristic algorithm based on link quality is given below.

- Step 1 Initialize the outage probability P_{dout} and set $y_{k,c} = 0$, denote the link gain from CU to D2DM is $G_{k,c}^{C2D} = \min_{d \in D_k} G_{k,c,d}^{C2D}, \forall k \in K, c \in C$.
- Step 2 When $\sum_{c=1}^C y_{k,c} = 1, \forall k \in K$ is not satisfied,
- (1) Traverse all the $k \in K$, calculate the set $C' = \left\{ \forall c \in C \mid \sum_{k=1}^K y_{k,c} < 1 \right\}$.
 - (2) If $C' = \emptyset$ of $\sum_{c=1}^C y_{k,c} < 1$, calculate the index $c^* = \arg \min p_c G_{k,c}^{C2D}$ for all $c \in C'$.
 - (3) Set $y_{k,c^*} = 1$.
 - (4) Calculate the outage probability of every corner point according to the power control algorithm above, and find the index value $\{P_c^*, P_k^*\}$ which satisfies the object function of Eq. (5).
 - (5) Calculate the P_{cout} and P_{dout} in Eqs. (3) and (4).
 - (6) Check whether P_{cout} and P_{dout} satisfy the constraints. If yes, set $y_{k,c^*} = 1$; else $y_{k,c^*} = 0$ and return to substep (1).
- Step 3 Check whether $y_{k,c}$ acquired from Step 2 satisfies $\sum_{c=1}^C y_{k,c} = 1, \forall k \in K$. If yes, output $y_{k,c}$, else, repeat Step 2.

The computation complexity is much smaller than that of Hungarian Algorithm.

4 Simulation Results and Analysis

In the simulation, we compare the performance of outage probability between the Hungarian algorithm which can find the optimum solution and the heuristic algorithm proposed in this paper, varying the parameters such as the D2DM number, radius of D2DM and SINR constraint of multicast communications (Table 1).

Table 1. Parameters setting

Parameters	Value
CU number (C)	100
D2DM number (K)	10
D2DM receiver number (D_k)	10
Maximum power of CU (p_{cmax})	18 dBm
Maximum power of D2DM sender (p_{kmax})	18 dBm

Figure 3 shows the relationship between D2DM outage probability and D2DM radius. With the increase of multicast group radius, the outage probability of each communication link is also increased, that is, the reliability is decreased. It can be seen that when the multicast group radius is 30 m, the average outage probability of multicast group link is approximately 1×10^{-4} .

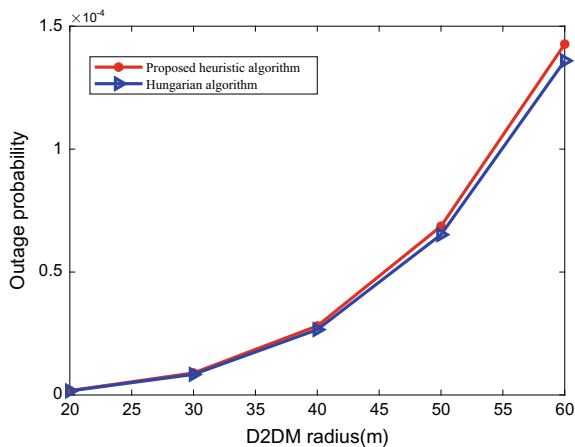


Fig. 3. D2DM outage probability versus D2DM radius.

Figure 4 shows the relationship between D2DM outage probability and the D2DM number. Since the simulation result is the average outage probability of each link, the outage probability is not significant monotonic change as the D2DM number increases. It can be seen that when the number of D2DM is [10, 20], the average outage probability is maintained at about 8×10^{-5} .

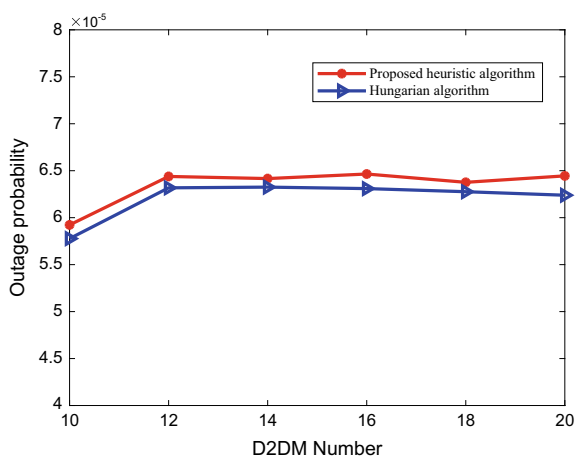


Fig. 4. D2DM outage probability versus D2DM number.

Figure 5 shows the relationship between D2DM outage probability and D2DM link SINR constraint. If the D2DM link SINR is lower than the constraint, this link will be deemed to be interrupted. With the increase of SINR constraint of multicast group, the

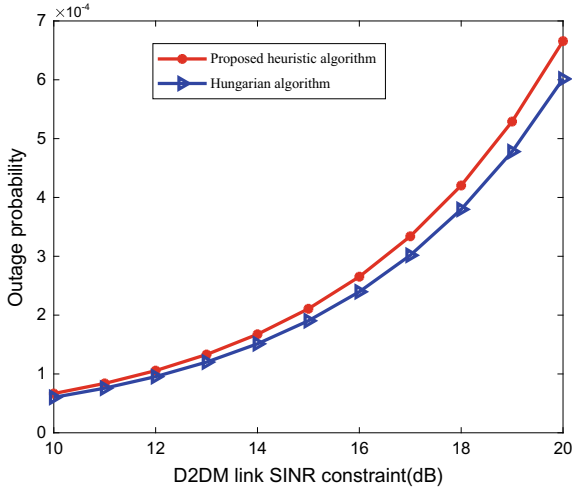


Fig. 5. D2DM outage probability versus D2DM link SINR constraint.

outage probability of each communication link is also increased. It can be seen that when the SINR constraint of multicast group is 10 dB, the average outage probability of multicast group link is about 7×10^{-5} .

From the simulation results, it can be found that the performances of the proposed heuristic algorithm are very close to that of Hungarian algorithm. However, the proposed heuristic algorithm has much lower computation complexity.

5 Conclusion

In this paper, the system model of D2D multicast communication is established. Then the objective function and constraints of the resource allocation problem are proposed for the application scenario according to the analysis of the quality information of related links. Power control and channel allocation algorithms are proposed to improve the reliability. The simulation results show that the performances of the proposed heuristic algorithm are very close to that of Hungarian algorithm, but computation complexity is much lower.

Acknowledgements. This research is supported by Natural Science Foundation of China (Grant No. 61501136).

References

1. Yu, S., Langar, R., Wang, X.A.: D2D-multicast based computation offloading framework for interactive applications. In: 2016 IEEE Global Communications Conference (GLOBECOM), 2016, 4–8 December (2016)
2. Sun, W., Yuan, D., Str, E.G., et al.: Cluster-based radio resource management for D2D-supported safety-critical V2X communications. *IEEE Trans. Wireless Commun.* **15**(4), 2756–2769 (2016)
3. Dongyu, W., Xiaoxiang, W., Yuan, Z. (eds.): An interference coordination scheme for device-to-device multicast in cellular networks. In: Vehicular Technology Conference (VTC Fall), 2012 IEEE, 2012, 3–6 September (2012)
4. Pushpalatha, M., Shruthi, N.V., Ramesh, T.K., Konda, S.K. (eds.): Efficient multicast algorithm for dynamic intra cluster device-to-device communication for small world model paper. In: 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2015, 10–13 August (2015)
5. Peng, H., Li, D., Ye, Q., Abboud, K., Zhao, H., Zhuang, W., et al. (eds.): Resource allocation for D2D-enabled inter-vehicle communications in multiplatoons. In: 2017 IEEE International Conference on Communications (ICC), 2017, 21–25 May (2017)
6. Bhardwaj, A., Agnihotri, S. (eds.): A resource allocation scheme for device-to-device multicast in cellular networks. In: 2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), 2015 August 30, 2015-September 2 (2015)
7. Wu, X., Chen, Y.: Joint resource allocation for cellular and D2D multicast based on cognitive radio. *Ksii Trans. Internet Inf. Syst.* **8**(1), 91–107 (2014)



Services Ranking Based Random Access Scheme for Machine-Type Communication

Lu Dai¹, Yunjian Jia¹(✉), Zhengchuan Chen¹, Liang Liang¹, and Guojun Li²

¹ College of Communication Engineering, Chongqing University,
Chongqing 400044, P. R. China

{dailu,yunjian,cxc,liangliang}@cqu.edu.cn

² Chongqing University of Posts and Telecommunications,
Chongqing 400065, P. R. China
skywavecommnavi@126.com

Abstract. There is a considerable pressure for random access networks to access numerous devices with limited preambles, while the different delay requirements of diverse applications exacerbate this situation. In this paper, we propose a service ranking scheme to ensure that delays of different applications are within a reasonable range. We classify applications by latency requirements and dynamically partition preambles for serving these applications. In particular, distribute queue is used to coordinate delay-critical applications. Delay-tolerant applications are partially prohibited when congestion occurs. The average success delay is minimized under the preambles constraint. Simulation and numerical results show that the proposed scheme can effectively reduce the average success delay by 30% while guaranteeing the success ratio of delay-critical applications.

Keywords: Machine-type communications · Random access · Massive access requests · Service ranked · Distributed queuing

1 Introduction

In the fourth-generation cellular networks, random access is widely used, which is fully applicable to the traditional Human-Type Communications (HTC). However, the application of Machine-Type Communications (MTC) brings challenges to the existing technologies. The main characters of MTC are numerous devices and small packets which need to be transmitted to eNB more frequently. The high overhead and congestion caused by drastically increase of MTC devices in a cellular are the major contradictions. As 3GPP technical report proposes, the maximum number of MTC devices covered by a cellular raises to 30,000 [1] while the number of preambles in a cellular is only 64. Due to this reason, a more flexible and efficient MAC protocol to manage access resource for numerous nodes is critical in the future system.

Many algorithms have been proposed to improve the average success delay with a limited number of random access (RA) preambles. Release 8 proposes Access Class Barring (ACB) scheme to classify users into different groups [2]. ACB scheme forbids the lower class transmission when access is congested. However, the service for the lower class may be totally unavailable when the traffic congested and the utilization rate of the preamble is low. Besides, the class of a MTC device's applications varies under different scenes. Some schemes use split preamble to prioritize UE to enhance Quality of Service in the RA procedure [3]. The split preamble is also used to distinguish the transmission of MTC and HTC. Disjoint allocation and joint allocation are adopted in these schemes [4, 5]. These schemes improve the average success delay and the success ratio comparing with ACB. UE priorities are proposed to solve the restriction of the preamble, and the different priorities of applications are not taken into consideration.

There are many algorithms proposed to simplify the random access. Tagged preambles are put forward to detect collision at the first step of RA procedure [6], while the traditional ways detect collisions at the third step. The authors of [7] use partial preamble sequences to avoid collisions. These algorithms simplify the random access procedure while they increase the complexity in the transmitters and receivers. Comparing with these algorithms, the distributed queue (DQ) algorithm has a good performance while introducing negligible complexity increase [8-10]. It allocates more resources to the collided devices and releases the preambles to lower priority applications' transmission when there is no congestion.

In general, MTC devices contain delay-critical applications such as breakage of high voltage wire and delay-tolerant applications such as general information update. HTC applications need to assure the priority to lower the interference of MTC applications. In this paper, we focus on classifying MTC applications to ensure that devices with delay-critical applications can be served in a timely manner. We consider that all the devices are possibly locked with delay-critical applications and delay-tolerant applications. Thus, devices should be able to select the preamble to transmit these different priority applications. DQ is used to transmit the higher priority applications. This tree algorithm allocates the collided devices more resources and retransmits these devices until they succeed. Besides, in order to ensure the transmission of higher priority applications, the lower priority applications are partly forbidden when congestion occurs. The priority of HTC is considered by guaranteeing preamble resources not be occupied too much by MTC devices. Therefore, the preamble is dynamically divided into different groups to transmit these different applications.

The rest of the paper is organized as follows. The services ranking based scheme description is detailed in Sect. 2. In Sect. 3, the model of HTC and MTC is presented and studied. The total delay of the system is also derived in this section. Section 4 explains the ranked-application scheme to process the delay-critical applications and delay-tolerant applications. Section 5 presents the performance simulation. Finally, we make a conclusion of our work in Sect. 6.

2 Service Ranking Based Scheme

In LTE networks, each UE that wants to transmit data to the eNB needs to send the random access preamble to reserve uplink resources in Physical Uplink Shared Channel (PUSCH). The transmission is organized into slots. Random access succeeds when the current preamble is selected by only one terminal in one slot. Random access fails when a preamble is selected by multiple terminals in one slot. And these failed terminals need to retransmit in the next slot until they succeed or reach the maximum number of retransmissions.

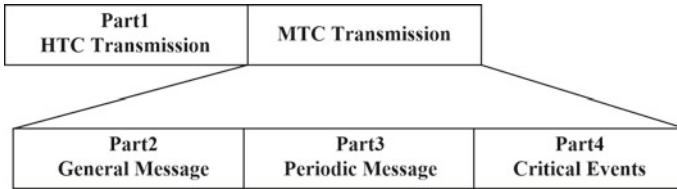


Fig. 1. Preamble distribution. The preamble is divided into four parts. Part1 for HTC is used for HTC transmission. Part2, Part3, and Part4 are used for MTC different applications' transmission respectively.

The preambles are split into four parts. Part1 is used for HTC transmission. Part2, Part3, and Part4 are used for MTC transmission. The partition line distinguishes the transmission of MTC and HTC, which ensures the preambles allocated to HTC UE within a reasonable range. As mentioned above, α is the corresponding partition line between HTC transmission and MTC transmission in Fig. 1. Part2, Part3, and Part4 are used for general information transmission, periodic information transmission, and critical events transmission respectively. General information transmission is converted to periodic information transmission when the number of retransmissions reaches the threshold. Terminals of Part1 and Part2 randomly select preambles. Terminals of Part3 and Part4 use DQ algorithm to select preambles.

DQ works through the distributed devices. The device itself decides the time to send the preamble according to the feedback information of eNB. The devices can get the length of Collision Resolution Queue (CRQ) in the feedback information [7]. Every device that has periodic information and critical events should start four integer counters inside. These counters count two queues among CRQs. One is CRQ0, the other is selected by CRQ0 among the idle CRQs. We denote these counters as RQ0, RQ, pRQ0, and pRQ. RQ0 is the number waiting for resolution in the CRQ0 queue. RQ is the number waiting for resolution in the selected CRQ queue. pRQ0, and pRQ are the position of device in CRQ0 and selected CRQ. The value of these counters is initially set as 0 and updated according to the feedback information.

In CRQ0, the first preamble is used for the transmission of periodic information and the second and the third preambles are allocated to transmit critical

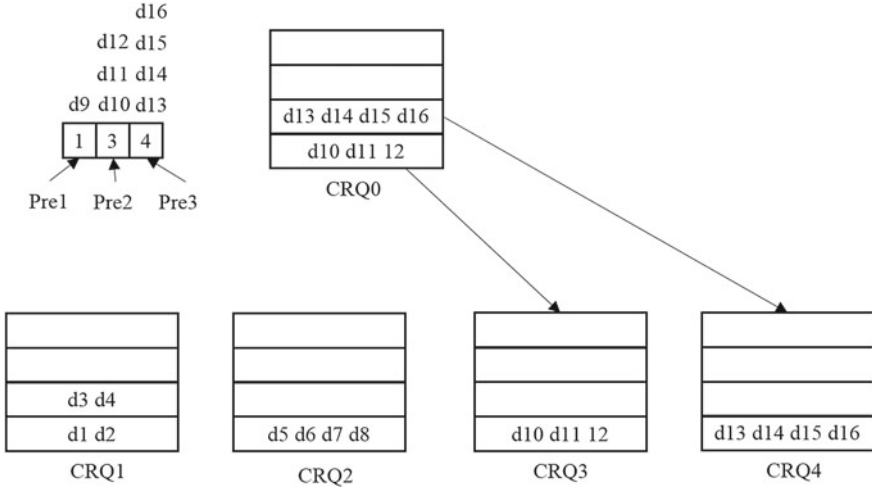


Fig. 2. The DQ procedures. The collision devices are allocated to null CRQs.

events. The device with these delay-critical application transmits using these fixed preambles. If they succeed, the start transmitting the data. If they fail, the value of CRQ0 increases one. The collided devices get the CRQ0's value and set their pRQ0 values equal to it. CRQ0 finds the idle CRQ and assigns the idle CRQs to the devices in the queue. The value of CRQ0 and pRQ0 decreases one when it assigns one idle CRQ to the devices in the queue. When there is no CRQ, the devices wait in the queue until there are idle ones. The devices get the selected queue and set their CRQ and pRQ as 1. The value of CRQ and pRQ decrease one in each slot. Meanwhile, the length of CRQ increases one when a collision happens in a slot. The device sends a preamble to access to eNB when its RQ counter decreases to zero. The new collision devices set their RQ0 counter equal to the length of CRQ0, such that it is placed itself at the end of the queue. The collided devices in the CRQ reset their RQ counter equal to the length of CRQ which means it is rearranged at the end of the queue.

The critical events and periodic information use DQ in random access procedure to prompt the success rate and delay to access to eNB. In this system, UE transmits the definite preamble when the device has the first or second rank service to transmit. As Fig. 2 shows, CRQ contains Pre1, Pre2, and Pre3. The devices chose Pre1 and Pre2 randomly to transmit the critical event and chose Pre3 to transmit periodic information. Then, the collision devices are put into the CRQ which is empty or wait in the CRQ0 until there are empty ones. The idle CRQ detected by CRQ0 among CRQ1 to CRQ4 in each slot. If they collide in the CRQ in a slot, CRQ increases one. PCRQ sets the value equal to CRQ until it successfully accesses to eNB.

As described in Fig. 2, d9 chooses Pre1 to transmit the critical event and succeeds. In the Pre2, d10 to d12 have collisions and they are put into the

CRQ3. The periodic information transmission of d13 to d16 uses the Pre3 and they are put into the CRQ0 too. Then, CRQ0 detects the empty CRQ. CRQ1 and CRQ2 are occupied by the other devices. CRQ3 and CRQ4 are empty, so CRQ0 notifies the collision devices with the number of empty CRQ. They are used to transmit the collision devices d10 to d12 and d13 to d16. That is the way DQ operates when critical events and periodical information transmission occur.

eNB broadcasts feedback information to devices so that the idle CRQ can be used for general message transmission. Once the idle CRQ is assigned to the collided devices in the queue, the transmission of general message is forbidden. When the transmission is congested and the success ratio is less than the threshold, the transmission is partially prohibited to ensure the transmission is not deadlock.

3 System Model

In this system, MTC devices and HTC UEs access to eNB through different preambles. In the transmission of MTC. Devices with different application select different preambles respectively. It is valuable to mention that the general message is partially prohibited when congestion happens.

Assuming that there are M available preambles can be randomly selected to reserve resources. The preambles are divided into two groups according to the results of HTC UE and MTC device success rate. One is used for HTC UE and is denoted as M^{H2H} . The other is used for the MTC devices and represented as M^{M2M} . M^{H2H} accounts for α of the total preambles M .

$$\begin{cases} M^{H2H} = M \times \alpha \\ M^{M2M} = M \times (1 - \alpha) \end{cases} \tag{1}$$

3.1 Human-Type Communication

Let us assume that the number of newly arrived preamble of each slot in RA procedure follows the Poisson distribution which mean number is λ , and the number of new arrivals per slot is independent from others slots. We use Z_N^{H2H} indicate the new arrival UEs in the n th slot, and $E[Z_N^{H2H}] = \lambda$. In order to ensure the successful access rate of each UE within the maximum number of retransmissions, we should ensure that the success rate of each slot is within a stable range. We set the success probability of the preamble in a slot to be constant and be expressed as P_S^{H2H} . Note that failed contending UEs will transmit in the next RA procedure and can be defined as (Fig. 3)

$$Z_F^{H2H} = Z^{H2H} \times (1 - P_S^{H2H}) \tag{2}$$

where Z^{H2H} refers to the total number of contending UEs. Z_C^{H2H} indicates the canceled UEs which reach the maximum number of retransmission W . It can be defined as

$$Z_C^{H2H} = Z_N^{H2H} \times (1 - P_S^{H2H})^W \tag{3}$$

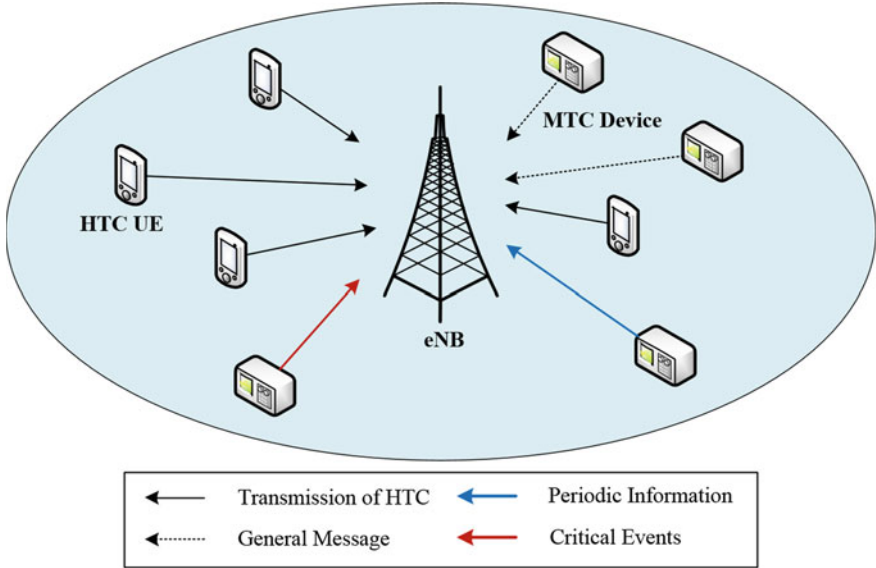


Fig. 3. The system is organized by the transmission of MTC and HTC. They use different preambles to transmit their message respectively. Critical events and periodic message are transmitted using fixed preambles. We use DQ algorithm for transmission when congestion occurs.

Thus, the total number of contending UEs includes new arrivals in this slot and failures in the previous transmission, and subtract cancelations reaching the maximum retransmission. It can be expressed as

$$Z^{H2H} = Z_N^{H2H} + Z_F^{H2H} - Z_C^{H2H} \tag{4}$$

UE randomly selects the preamble in each slot. The successful preamble is selected only by one UE. And the unsuccessful preamble is either selected by multiple UEs or selected by no UEs. Thus, as shown in [11], the failure UEs can be expressed as

$$\begin{aligned} E[Z_F^{H2H}] &= E[Z^{H2H}(1 - (\frac{M_{H2H} - 1}{M_{H2H}})^{Z^{H2H}})] \\ &\approx \sum_{i=1}^{\infty} \frac{(Z^{H2H})^k}{k!} e^{-Z^{H2H}} k [1 - (\frac{M_{H2H} - 1}{M_{H2H}})^{k-1}] \\ &= Z^{H2H} [1 - e^{-Z^{H2H}} \exp(Z^{H2H} \frac{M_{H2H} - 1}{M_{H2H}})] \\ &= Z^{H2H} [1 - e^{-Z^{H2H}}] \end{aligned} \tag{5}$$

Using (2) and (5), the success rate can be expressed as

$$P_S^{H2H} = e^{-Z^{H2H}} \tag{6}$$

Since the transmission of HTC can be maintained in a stable state, the retransmission of each time slot can also be kept stable. UEs that are successfully transmitted in each time slot can be expressed as the sum of the success of all transmission UEs. Therefore, the successful UEs in each slot can be described as

$$Z_S^{H2H} = \lambda \times [1 - (1 - P_S^{H2H})^W] \tag{7}$$

The delay of successful UEs can be expressed as

$$D_S^{H2H} = \tau_1 \times \sum_{n=0}^{W-1} n \frac{(1 - P_S^{H2H})^n P_S^{H2H}}{\sum_{k=0}^{W-1} (1 - P_S^{H2H})^k P_S^{H2H}} \tag{8}$$

where τ_1 is the duration of one slot.

3.2 Machine-Type Communication

MTC devices randomly select the preamble in M^{M2M} and send a signaling message to eNB. The number of MTC devices will not change in most cases. The UE that sends the preamble selected only by itself in a slot successfully transmits Msg1. The UE that sends the preamble selected by multiple UEs in a slot has failed to transmit Msg1. The failed MTC device will retransmit Msg1 until it reaches the maximum retransmission times and the rank of the MTC device is upgrade. The number of request time is calculated by UE itself and increases by one in each slot. Assume that the total number of MTC devices that select preambles is a constant denoted by Z_N^{M2M} . In order to guarantee the success of HTC UEs and MTC devices, some general message transmissions is prohibited when current preambles are insufficient to transmit all the requests. β is used to describe the transmission rate of MTC devices. The number of MTC devices transmitting in the n th slot can be given as

$$Z^{M2M} [n] = (Z_N^{M2M} + Z_F^{M2M} [n - 1]) \times \beta \tag{9}$$

in which $Z_F^{M2M} [n - 1]$ refers to the failure transmission in the previous slot. It can be expressed as

$$Z_F^{M2M} [n] = (1 - P_S^{M2M} [n]) \times Z^{M2M} [n] - Z_C^{M2M} [n] \tag{10}$$

The canceled general message transmission is updated to periodic message. the canceled general message transmission can be written as

$$Z_C^{M2M} [n] = Z_N^{M2M} [n] \times \prod_{i=n-1}^{n-W-1} (1 - P_S^{M2M} [i]) \tag{11}$$

The preambles for the MTC devices' general message transmission in the n th slot can be written as $M_N^{M2M} [n]$, which varies with the condition of α and M_{DQ}^{M2M} . It can be given as

$$M_N^{M2M} [n] = M^{M2M} [n] - E [M_{DQ}^{M2M}] \tag{12}$$

in which $M^{\text{M2M}}[n]$ is the total preambles assigned to the MTC devices, and can be calculated through (1). $E[M_{DQ}^{\text{M2M}}]$ indicates the average preambles used by the DQ. We assume that the arrival devices follow the Poisson distribution and the average number of newly arrived devices is 2. We simulate the DQ procedures and the average number of CRQ per slot is 1.6. Thus, the average number of preambles occupied by the DQ procedures is 4.8. All the MTC devices randomly select the preamble and independent to others. An MTC device transmits Msg1 successfully only when it selects a preamble only chosen by itself. Hence, the probability of the MTC devices' general message transmission in the n th slot can be written as

$$P_S^{\text{M2M}}[n] = \binom{M_N^{\text{M2M}}[n]}{1} \frac{1}{M_N^{\text{M2M}}[n]} \left(1 - \frac{1}{M_N^{\text{M2M}}[n]}\right)^{Z^{\text{M2M}}[n]-1} \quad (13)$$

The total number of UEs that transmit Msg1 successfully in the n th slot can be calculated as

$$Z_S^{\text{M2M}}[n] = Z^{\text{M2M}}[n] \times P_S^{\text{M2M}}[n] \quad (14)$$

4 Problem Formulation

The objective in this part is to minimize the average success delay of MTC and HTC transmission. The problem is constructed as an optimization problem that can minimize the average success delay of MTC transmission and guarantee the transmission of HTC UEs and MTC devices. The average success delay can be defined as

$$\mathbb{E}[\tau] = \frac{NZ_S^{\text{H2H}}D_S^{\text{H2H}} + \tau_2 \sum_{n=1}^N (Z_N^{\text{M2M}}[n] + Z_F^{\text{M2M}}[n] - W \times Z_C^{\text{M2M}}[n])}{NZ_S^{\text{H2H}} + \sum_{n=1}^N Z_S^{\text{M2M}}[n]} \quad (15)$$

where $\mathbb{E}[\tau]$ is the average delay of system. τ_2 corresponds to the delay general message transmission of MTC. The optimization problem can be described as

$$\begin{aligned} & \min \mathbb{E}[\tau] \\ & \text{s.t. } 0 < \alpha < 1 \\ & 0 < \beta < 1 \\ & M - M \times \alpha > 1 \\ & 1 - (1 - P_S^{\text{H2H}})^W > 0.9999 \end{aligned} \quad (16)$$

The transmission of HTC must follow the conditions in (16), which can promise the successful transmission of most UEs and keep the transmission in stable state. HTC transmission cannot reach the specified success rate when the number of current preambles is reduced and the transmission success rate of HTC will not increase too much when the number of current preambles increases. Meanwhile, assigning the same number of preambles to MTC can increase more successful terminals. Therefore, the preamble utilization rate is the highest when

the preamble allocated to HTC is just enough to achieve the specified success rate. Therefore, the number of preambles allocated to HTC can be calculated and α is determined.

$$\begin{cases} Z^{\text{H2H}} = \lambda \frac{1 - (1 - P_s^{\text{H2H}})^w}{P_s^{\text{H2H}}} \\ M_N^{\text{H2H}} = -Z^{\text{H2H}} (\ln P_s^{\text{H2H}})^{-1} \end{cases} \quad (17)$$

Once the preambles allocated to HTC is determined, the number of preambles allocated to MTC can be calculated. We find the reciprocal of $Z_S^{\text{M2M}}[n]$ and calculate the number of devices participating in the competition when the number of successful devices is the maximum value, which can be expressed as

$$Z^{\text{M2M}}[n] = \min \left\{ (Z_N^{\text{M2M}} + Z_F^{\text{M2M}}[n-1]), - \left[\ln \left(1 - \frac{1}{M_N^{\text{M2M}}[n]} \right) \right]^{-1} \right\} \quad (18)$$

5 Simulation Results and Discussions

In this section, we evaluate the performance of the proposed service ranking scheme in terms of average success delay and success access ratio. The performance comparison of random selection scheme and ACB has been simulated. In the simulation, the number of newly arrived HTC UEs per slot in the RA procedure follows a Poisson distribution, and the number of MTC devices follows a uniform distribution. There are 54 preambles and the incidence of events per device follows the Poisson distribution.

5.1 Average Success Delay

To compare the performance of three kinds of services, we evaluate the average success delay first. Fig. 4a shows average success delay of device. The random selection scheme has the same delay with ACB at first because ACB transmits with no limitations for the lower priorities. Service ranking scheme has higher delay because it scarifies the devices' delay to support LTE UE and higher rank services. The average success delay of LTE UE in Fig. 4b shows that service ranking scheme has best performance as it allocates preambles to LTE UE dynamically. The delay grows with the increasing number of devices. The random selection scheme has the longest delay. The critical event's transmission is shown in Fig. 4c, which indicates that service ranking scheme has more stable performance than other schemes. In these performance comparisons, it is observed that the proposed scheme can keep the delay of critical events within a small range while preventing LTE UEs from being interrupted.

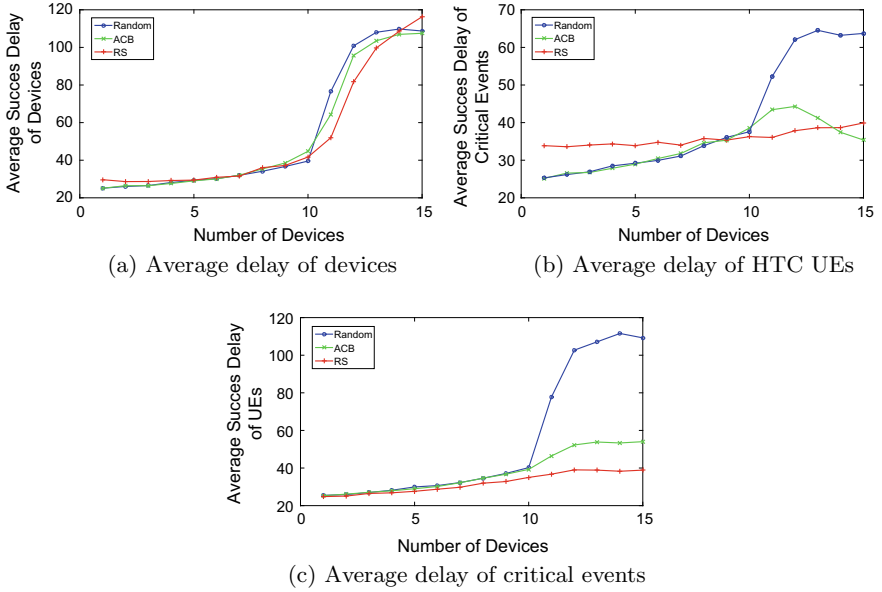


Fig. 4. Performance comparison in terms of average delay

5.2 Average Success Ratio

The success ratio is another important index of system performance. We compare the success ratio of three schemes. The random selection scheme in Fig. 5a has the same success rate because all UEs randomly select the preamble and have the same probability of success. ACB scheme sets LTE UE as first class to guarantee its transmission, while sacrificing the success ratio of devices. Since the periodical message transmission guarantees the success rate of the device, the service ranking scheme has the best performance. It is seen from Fig. 5b that the success ratio of the LTE UE remains at a high level, and when the access request becomes more frequent, the success ratio of the device decreases. ACB can hardly protect the transmission of device’s critical event while it can release congestion through barring devices’ transmission as described in Fig. 5c. Service ranking scheme has the best performance, and DQ guarantees success ratio. We can observe from the three figures that the proposed schemes can guarantee the success ratio of critical event and LTE. It can avoid the congestion of preambles.

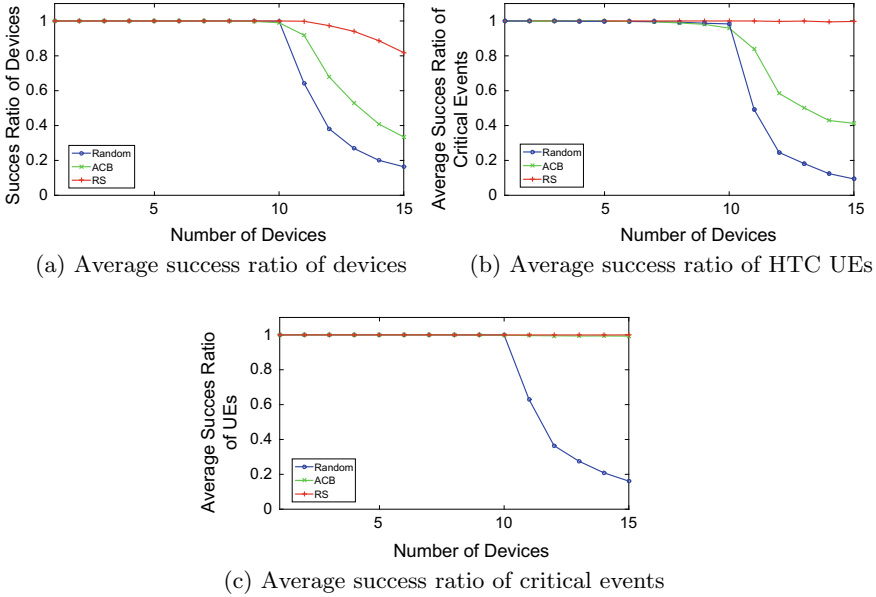


Fig. 5. Performance comparison in terms of success ratio

6 Conclusion

In order to guarantee the success rate of critical events and LTE transmissions, we propose a service ranking scheme which restricts access to the general message of devices. Genetic algorithm is used to reach the suboptimal solution. The simulation results show that our scheme can serve varies applications within their delay requirements. The proposed scheme reduces 70% delay and keeps success ratio in a promising level comparing with random selection scheme. In future work, we will consider adopting queuing theory to solve the higher level message’s transmission of MTC device.

Acknowledgments. This work was supported in part by the Scientific Research Foundation of the Ministry of Education of China-China Mobile under Grant MCM20150102; in part Ministry of Education of China-China Mobile under Grant MCM20150102; in part Ministry of Education of China-China Mobile under Grant MCM20150102; in part and universities in Chongqing, China under Grant CXTDX201601006.

References

1. Huawei: Technical specification group radio access network; study on RAN improvements for machine-type communications. 3GPP, Technical Report, 37.868 (2011)
2. Larmo, A., et al.: The LTE link-layer design. In: LTE Part II: 3GPP Release 8 (2009)
3. Pourmoghadas, A., Poonacha, P.G., Khoenihay, N.: Performance analysis of split preamble RAN overload control method for massive access of M2M devices in LTE networks. In: IEEE International Conference on Internet of Things (2017)
4. Astudillo, C.A., Andrade, T.P.C.D., Fonseca, N.L.S.D.: Allocation of control resources with preamble priority awareness for human and machine type communications in LTE-advanced networks. In: IEEE International Conference on Communications (2017)
5. Pourmoghadas, A., Poonacha, P.G.: Performance analysis of a machine-to-machine friendly MAC algorithm in LTE-advanced. In: International Conference on Advances in Computing, Communications and Informatics (2014)
6. Han, S., Jang, et al.: An early preamble collision detection scheme based on tagged preambles for cellular M2M random access. In: IEEE Transactions on Vehicular Technology (2017)
7. Kim, T., et al.: An enhanced access reservation protocol with a partial preamble transmission mechanism in NB-IoT systems. In: IEEE Communications Letters (2017)
8. Laya, A.: Goodbye, aloha!. In: IEEE Access (2016)
9. Changbae, Y.: Distributed queuing with preamble grouping for massive IoT devices in LTE random access. In: International Conference on Information and Communication Technology Convergence (2016)
10. Laya, A., Alonso, L., Alonso-Zarate, J.: Efficient contention resolution in highly dense LTE networks for machine type communications. In: IEEE Global Communications Conference (2015)
11. Alonso-Zarate, T.R.: Performance Models for LTE-Advanced Random Access. Arizona State University (2014)



A Novel Energy Harvesting Scheme in Interference Networks with UAVs

Bingcai Chen^{1,2(✉)}, Manrou Yang¹, and Yu Chen¹

¹ School of Computer Science and Technology, Dalian University of Technology,
Dalian 116024, China

cbc9@qq.com

² School of Computer Science and Technology, Xinjiang Normal University,
Urumqi 830054, China

Abstract. Along with the development of the communication network technology and computer technology, the diversity of information in modern society and the geographical area scale has placed limits on the single unmanned aerial vehicle (UAV) in the tasks, so that cooperative mission of multiple UAVs will be the main form to be utilized in the future information society. Interference and energy limit are two key issues in the network, so we propose a novel opportunistic interference alignment (OIA) scheme with wireless energy harvesting (EH), which can eliminate the interference while optimizing the energy efficiency. Extensive simulation results are presented to show the effectiveness of the proposed OIA scheme with wireless EH.

Keywords: Interference alignment · Energy harvesting · Power splitting · Opportunistic communications

1 Introduction

In order to meet the increasing demand of data transmission as well as the different types of data from diverse areas, UAVs gradually come into people's horizons which emerge as a new kind of data transmission method and data acquisition technology [1–5].

To effectively eliminate the interferences, we introduce interference alignment (IA) [6, 7] method in our system. Gomadam et al. propose a distributed numerical method to solve the problem which utilizes the network reciprocity to do the iterations, and the feasible conditions of IA are analyzed in previous papers. Besides, opportunistic IA (OIA) can be utilized to improve the energy efficiency of the system, and simultaneous wireless information and power transfer (SWIPT) which is first proposed by Varshney has become the hot point in the wireless networks.

The remaining part of the paper is arranged as follows. In Sect. 2, we will describe the system model, and the feasibility analysis and iterative IA algorithm

are proposed in Sect. 3. In Sect. 4, we will propose a novel scheme, and apply it to the system we proposed. In Sect. 5, we will conduct simulations and analyze the results.

2 System Model

We consider a wireless network with UAVs in which some UAVs play the role as relays. There are K transmitters, K relays which are full-duplex and K receivers in the system. We assume M antennas are equipped on each transmitter, N_r receiving antennas and N_t transmitting antennas on the relays and N antennas on each receiver. The receiving signals are also discussed in two parts. First, the recovered signal of the i th relay can be expressed as follows:

$$\begin{aligned}
 \mathbf{y}_b^{[i]} = & \mathbf{u}^{[i]\dagger} \mathbf{h}^{[ii]} \mathbf{v}^{[i]} \mathbf{x}^{[i]} + \sum_{l=1, l \neq i}^K \mathbf{u}^{[i]\dagger} \mathbf{h}^{[il]} \mathbf{v}^{[l]} \mathbf{x}^{[l]} \\
 & + \sum_{k=1, k \neq i}^K \mathbf{u}^{[i]\dagger} \mathbf{G}^{[ik]} \mathbf{V}^{[k]} \mathbf{x}_1^{[k]} + \mathbf{u}^{[i]\dagger} \mathbf{z}^{[i]}
 \end{aligned} \tag{1}$$

The receiving signal of i th receiver can be described as follows:

$$\mathbf{y}^{[i]} = \mathbf{U}^{[i]H} \mathbf{H}^{[ii]} \mathbf{V}^{[i]} \mathbf{x}_1^{[i]} + \sum_{k=1, k \neq i}^K \mathbf{U}^{[i]H} \mathbf{H}^{[ik]} \mathbf{V}^{[k]} \mathbf{x}_1^{[k]} + \mathbf{U}^{[i]H} \mathbf{z}^{[i]} \tag{2}$$

3 Energy Harvesting in IA-Based Interference Networks with UAVs

When information is transmitted, the interference between users becomes more serious. Therefore, IA can be used in wireless communication network to effectively remove interferences. According to Bezout’s theory, a general polynomial system can be solved when and only if the number of equations does not exceed the number of variables. According to the formulas above, the number of equations and the number of variables can be expressed as follows.

$$N_\varepsilon = 3 \times K \times (K - 1) \times d^2 \tag{3}$$

$$N_v = K \times d \times (M + N + N_r + N_t - 4 \times d) \tag{4}$$

When the feasible condition is satisfied, that is $N_v \geq N_\varepsilon$, then IA can be performed in the system and the interferences can be perfectly eliminated. If IA is feasible, the pre-coding matrix and decoding matrix can be obtained iteratively.

Mobility and flexibility of the UAVs can effectively improve the performance of information transmission rate, but because of the energy of UAVs battery is very finite, energy efficiency has become an important aspect in the communication with UAVs. So in this section, we propose an OIA scheme which takes

both information transmission rate and energy harvesting into consideration. As the antenna selection scheme and power splitting scheme both have their advantages and disadvantages. We would like to combine two methods to get an optimal solution. In the antenna selection scheme, antennas are selected to do either information decoding or energy harvesting. We assume that the received power by the selected antennas is sufficient for information decoding, so there exists another part of the received power which can be harvested too. We can apply power splitting scheme on them as well.

4 Simulation Results and Discussions

In the simulation, we consider a three-user wireless network and $M = 3, N = 3, N_t = 3, N_r = 3, K = 3$. The transmit power of each UAV is set to P_1 , and the transmit power of each relay is set to P_2 .

First, the performance of the combined algorithm in sum rate is investigated. Figure 1 shows the sum rate obtained by common IA, antenna selection, power splitting and the combined scheme, respectively. Considering $P_1 = P_2 = 20W, d = 1$. From the results, we can observe that the proposed scheme has a better performance of sum rate than other schemes at the same SNR.

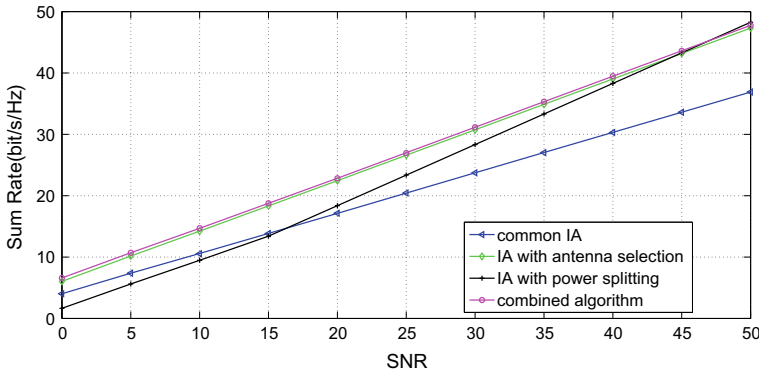


Fig. 1. Performance of sum rate. $P_1 = P_2 = 20W, d = 1$

Then we change the value of d to investigate the influence of degree of freedom. Considering $d = 2$, Fig. 2 shows the sum rate of different schemes. Compared with Fig. 1, the SNR at which the performance of power splitting scheme exceeds the combined scheme decreases. So that low degree of freedom can improve the performance of sum rate of the proposed scheme. This property suits the communication system with UAVs.

Another issue we considered is the energy harvesting. The harvested energy of different schemes when $P_1 = P_2 = 20W, d = 1$ is shown in Fig. 3. From the

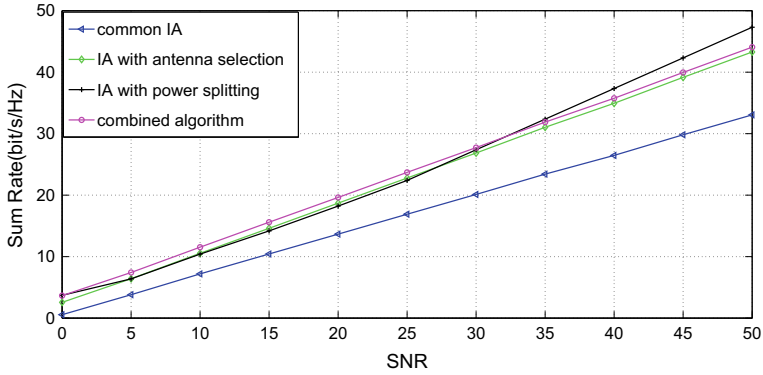


Fig. 2. Performance of sum rate. $P_1 = P_2 = 20W, d = 2$

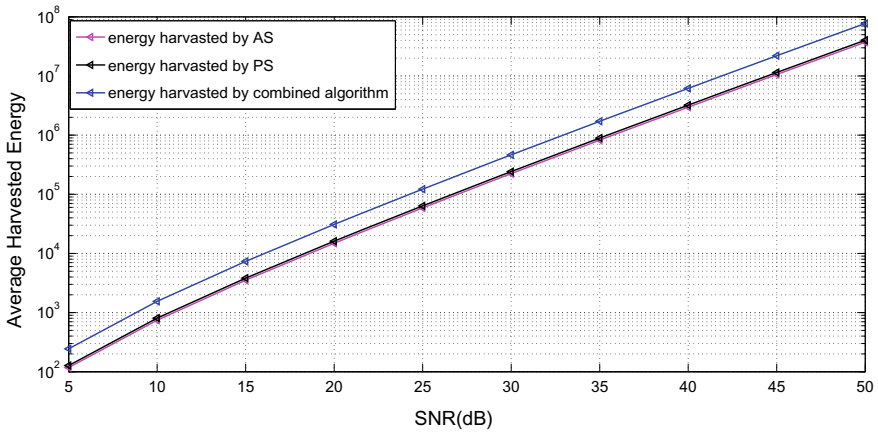


Fig. 3. Performance of harvested energy.

results, we can find that the proposed scheme can harvest more energy than both antenna selection and power splitting schemes at the same SNR. So that the performance of the proposed scheme in energy harvesting is better than other schemes.

5 Conclusion

In this paper, we proposed a combined scheme of antenna selection and power splitting schemes, which performed power splitting scheme on the selected antennas in the IA-based network to further improve the performance of system with UAVs. The sum rate of proposed scheme is higher than other schemes when the condition suits the system with UAVs. And the energy harvested by the

proposed scheme is higher than previous schemes. Compared with conventional energy harvesting schemes, the proposed has a better overall performance.

Acknowledgements. This work was supported by National Natural Science Foundation of China (NSFC) under Grant 61771089.

References

1. Zhao, N., Cheng, F., Richard Yu, F., Tang, J., Chen, Y.F., Gui, G., Sari, H.: Caching UAV assisted secure transmission in hyper-dense networks based on interference alignment. *IEEE Commun. Soc.* **66**, 2281–2294 (2018)
2. Mozaffari, M., Saad, W., Bennis, M., Debbah, M.: Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage. *IEEE Commun. Lett.* **20**, 1647–1650 (2016)
3. Lyu, J., Zeng, Y., Zhang, R., Lim, T.J.: Placement optimization of UAV-mounted mobile base stations. *IEEE Commun. Lett.* **21**, 604–607 (2017)
4. Rohde, S., Putzke, M., Wietfeld, C.: Ad hoc self-healing of OFDMA networks using UAV-based relays. *Ad Hoc Netw.* **11**, 1893–1906 (2013)
5. Sharma, V., Bennis, M., Kumar, R.: UAV-assisted heterogeneous networks for capacity enhancement. *IEEE Commun. Lett.* **20**, 1207–1210 (2016)
6. Cadambe, V.R., Jafar, S.A.: Interference alignment and degrees of freedom of the K-user interference channel. *IEEE Trans. Inf. Theory* **54**, 3425–3441 (2008)
7. Shin, W., Lim, J.B., Choi, H.-H., Lee, J., Poor, H.V.: Cyclic interference alignment for full-duplex multi-antenna cellular networks. *IEEE Trans. Commun.* **65**, 2657–2671 (2017)



Low-Power Wide Area Networks: Changes for Smart Grid

Liang Wan¹, Yirui Huang¹, Weihua Li², Yu Zhang³,
and Zhijian Zhang¹(✉)

¹ School of Electronic Science and Engineering,
Nanjing University, Nanjing, China

njuzzzj@nju.edu.cn

² Unit 31106 of PLA, Pudong, China

³ Unit 31618 of PLA, Pudong, China

Abstract. Nowadays, a hybrid of communication technologies including a variety of wired and wireless technologies have been used in smart grid. Each wireless communication technology has its advantages and disadvantages. Short-range links, high power consumption, high cost, or a mesh topology are the main characters for wireless technologies. Due to wireless connection problems in mesh topology and a large amount of infrastructure required to build a reliable network, the costs of deployment for smart grid are still amazing. Recently, a new wireless technology Low-Power Wide Area Network (LPWAN) has challenged typical wireless connection. It is characterized by low-cost and long-range transmission technologies, and enables power efficient communication over very long distances using a simple topology. In this paper, we introduce the new wireless technology to implement the applications in smart grid, and discuss its advantages and the potential opportunities and challenges that can bring to smart grid.

Keywords: Smart grid · LPWAN · Long range · LoRa · NB-IoT

1 Introduction

The next-generation electricity grid, known as the “smart grid” or “intelligent grid”, is a modern electric power grid infrastructure for providing more robustness, efficiency, security, reliability, flexibility, and quality of service, through two-way communications, digital technologies, advanced sensing and metering technologies, software abilities, and modern energy management techniques [1]. Smart grid can be predicated as a two-way flow of power in electrical network and information in communication network between the two concerned entities, i.e., grid and consumer [2]. A reliable real-time or near real-time data transmission and exchange between all components is essential for smart grid’s successful operation. Communication provides the channel for the data and plays an important role.

A hybrid of communication technologies including wired technologies such as fiber optics, power line communication (PLC) systems, and a variety of wireless technologies, e.g., ZigBee, GSM, GPRS, WiMAX, and WLAN, have been used in

supporting some monitoring or controlling applications [3]. Advanced wireless systems can provide the benefits of low costs, rapid deployment, widespread access, and mobile communications which wired technologies cannot provide [4]. However, Typical wireless technologies are characterized by short-range and the mesh topology, and these induced the costs of the communication infrastructure is still amazing, due to wireless connection problems in mesh topology and a large amount infrastructures required to build a reliable network.

The Low-Power Wide Area Network (LPWAN) which enables power efficient communication over very long distances represents an exciting new trend in the development of the wireless communication systems, and its main characteristics are long lifetime of nodes, low transceiver chip cost and large coverage area [5]. Most LPWAN technologies use a star network instead of a mesh topology network. LPWAN can be used when other wireless networks are not a good fit. For example, ZigBee and WiFi are unsuitable for long-range performance, and cellular networks are costly as far as hardware and services, and have higher energy consumption and limited capacity. A typical application scenario for LPWAN technologies is citywide meter reading collection where devices send readings at very low frequency over a long distance to a data concentrator (one-hop networks) [6]. It is foreseen that LPWANs will be widely used for a broad range of possible applications, e.g., smart energy metering and smart grids monitoring, and it will bring a great change for smart grid communication network architecture and reduce the costs greatly.

In this paper, we will provide an introductory overview of the LPWANs' main technological interpretations and focus on the advantages of LPWANs' connectivity with respect to the traditional solutions operating in the unlicensed spectrum for smart grid applications. To confirm this argumentation, we have conducted some experiments based on LoRa network with commercially available hardware.

The rest of this paper is outlined as follows: Sect. 2 describes the smart grid communications network architecture, and Sect. 3 describes the LPWAN products available today, focusing in greater detail on LoRa and NB-IoT. In Sect. 4, we discuss the experience gained with some experimental test of a LoRa network. Conclusions and final remarks can be found in Sect. 5.

2 Smart Grid Communication Network Architecture

A smart grid Communication is represented by a hierarchical multilayer architecture, this three-layered architecture comprises home area network (HAN), neighborhood area network (NAN), and wide area network (WAN). The three main tiers that are located between these three networks are the access point, backhaul distribution, and the core backbone [7]. Specific network requirements for major smart grid applications in terms of their data rate and coverage range requirements are summarized in Table 1.

HAN is at the customer end of the network architecture and supports communications for customer domain applications. HAN is deployed and operated within a small area and its coverage distance is up to 100 m [8]. HAN applications can be implemented over communication technologies including ZigBee, WLAN, PLC, and M-Bus.

Table 1. Network Requirements for major smart grid applications and communication technologies that can meet the requirements

Network	Application	Data rate	Coverage range	Communications
HAN	Home automation	<100 Kbps	Up to 100 m	PLC/ZigBee/ WLAN/Z-Wave
NAN	On-demand meter reading	>100 Kbps	Up to 10 km	PLC/Wireless Mesh/WiMAX/Cellular
	Multi-interval meter reading	>100 Kbps		PLC/Wireless Mesh/WiMAX/Cellular
	Load management	>50 Kbps		/Wireless Mesh/WiMAX/Cellular
	Distribution automation	>18 Kbps		Wireless Mesh/WiMAX/Cellular
WAN	Synchrophasor	>2 Mbps	Up to 100 km	Fiber Optic/WiMAX/Cellular
	Backhaul/core/metro networks	>10 Mbps		Fiber Optic/WiMAX/Cellular

NAN, which also can be called field area network (FAN), acts as a bridge between customer premises and substations with collectors, access points, and data concentrators [9]. NAN/FAN is deployed and operated within the areas of urban–suburban and rural environments, and its coverage distance could be up to 10 km [8]. NAN/FAN applications can be implemented over communication technologies including WLAN, WiMAX, Wireless Mesh, PLC, and cellular technologies.

WAN is the backbone of the smart grid communication network and its coverage can be up to 100 km [8]. The WAN applications can be implemented over various communication technologies such as Fiber Optic, PLC, WiMAX, and cellular technologies.

3 A Review of LPWANs Communications

3.1 Sigfox

Sigfox was proposed for the first time by the French company Sigfox in 2009. The main purpose of Sigfox technology is to meet the requirements of the applications which need long device battery life cycle, low device cost, long range, and high access capacity. Sigfox has served 24 countries, and now it is being introduced to establish global IoT network to collect and exchange millions of objects’ data. Sigfox uses the ultra narrow band modulation and operates in 868 or 915 MHz ISM bands [10]. It uses star topology which is akin to a cellular architecture, but rolls out a lightweight protocol to deal with downlink and uplink message. In Sigfox, the coverage of one base station might reach an area of dozens of square kilometers. The network data is usually small, for a downlink message, it is up to 8-bytes payload, and 12-bytes payload for an uplink message. The devices in Sigfox network do not connect a specific base station, and it is

no need to establish and maintain network connections, the message is received by any base station in the range [11].

3.2 LoRa

LoRa, which stands for “Long Range”, is a private physical layer LPWAN solution belonged to Semtech Corporation. The technology employs a spread spectrum method based on chirp spread spectrum modulation (CSS) technique used for radar applications since 1940s [12, 13]. This technique helps LoRa to resistant against multipath fading, Doppler effect, and in-band jamming interferers, and it also improves receiver’s sensitivity due to the respective processing gain of the spread spectrum technology.

In LoRa modulation, the spreading of the spectrum is achieved by generating a chirp signal that continuously varies in frequency. The raw data rate (R_b) can be expressed as follows [12]:

$$R_b = SF * \frac{CR}{\lceil \frac{2^{SF}}{BW} \rceil} (\text{bit/sec}) \quad (1)$$

where SF is the spreading factor, BW is the modulation bandwidth, and CR is the code rate. For LoRaWAN in the EU bands, SF can be set to 7/8/9/10/11/12, BW can take the value of either 125 or 250 kHz, and CR can be 4/5, 4/6, 4/7 or 4/8. Thus, three parameters—SF, BW, and CR—determine the bit rate of a point-to-point LoRa link. Higher SF values result in lower bit rate, but at the same time, higher sensitivity.

The LoRa Alliance promotes an open standard for LoRa-based networks called LoRaWAN, which was developed by Semtech Corporation, IBM Research, and Actility. The LoRaWAN specification defines the network architecture and the communication protocol. A typical LoRaWAN network is laid out in a star-of-stars topology [14], which consists of end devices, gateways or concentrators and a server. As shown in Fig. 1, in “star-of-stars” topology, end devices are connected to one or many gateways or concentrators via a single-hop LoRa link, while gateways or concentrators are connected to a common Network Server via standard IP protocols.

In the LoRaWAN network, end devices can be defined as three different classes according to the trade-off between network downlink communication latency versus battery life, as shown in Fig. 1 [15, 16]. (a) Class A (bidirectional end devices). It has maximum battery lifetime and must be supported by all other devices. The end devices transmit to the gateway when needed. After transmission, the node opens a receive window to obtain queued messages from the gateway. (b) Class B (bidirectional end devices with scheduled receive slots). All end devices start and join the network as end devices of Class A and then can decide to switch to Class B. Class B end devices must receive a time-synchronized beacon from the gateway to receive window at the scheduled time. (c) Class C (bidirectional end devices with maximal receive slots). Class C have almost continuously open receive windows, only closed when transmitting.

3.3 NB-IoT

Narrowband Internet of things (NB-IoT) was introduced as a new radio interface to the specification 3 GPP Rel. 13 in 2016. It is designed to support high network capacity,

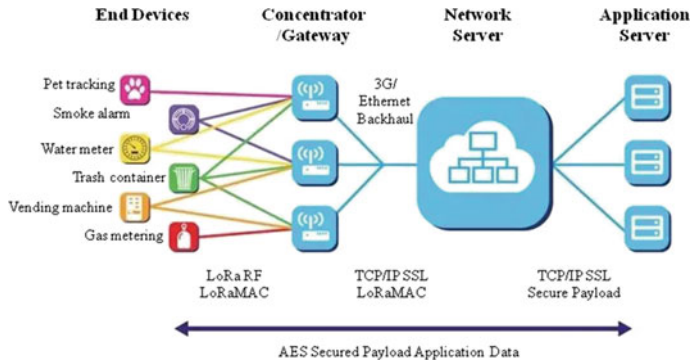


Fig. 1. LoRa network architecture

low data rate, low device power consumption, extreme coverage, and ultralow device cost. NB-IoT could be considered as new narrower radio interface, but its basic functionalities are inherited from the legacy LTE system. NB-IoT, GSM, and legacy LTE could serve the same region without interfering with each other, which means they could coexist together [17]. NB-IoT employs QPSK modulation and utilizes the same frequency numbers as used in legacy LTE [18].

NB-IoT utilizes one or more physical resource block (PRB) of LTE (180 kHz). The downlink consists of Narrowband Physical Broadcast Channel (NPBCH), Narrowband Physical Downlink Control Channel (NPDCCH), and Narrowband Physical Downlink Shared Channel (NPDSCH), while the uplink includes Narrowband Physical Uplink Shared Channel (NPUSCH) and Narrowband Physical Random Access Channel (NPRACH). These channels are based on existing LTE channels with necessary modifications to fit into the narrow bandwidth used by NB-IoT [18]. Enhanced coverage and low power consumption are the most prominent features of NB-IoT.

According to 3Gpp Rel.13, NB-IoT could operate in three types of deployment scenes: stand-alone, guard-band, and in-band [19, 20]. In stand-alone operation, NB-IoT can be deployed with any available frequency band, and the current GSM frequencies might be a good choice. In guard-band operation, NB-IoT could operate in the unused edge frequency band of LTE carrier's, that is guard-band. In general, the NB-IoT carriers are only located on one side guard-band of the LTE carriers. In in-band operation, NB-IoT utilizes the same resource block as LTE within LTE carriers. LTE and NB-IoT share the transmit power. NB-IoT carriers are located in the middle of the LTE carriers. There would be more interfaces for NB-IoT, and some restrictions are also for LTE operation [18].

4 Measurement Setup Using a LoRa Network

There are three typical key challenges for LPWAN [21]: (a) low cost, nodes should be cheap; (b) low battery consumption, nodes should have a lifetime of up to ten years when powered from a battery; and (c) coverage rang, the distance between the base

station and a node may exceed 10 km. Under the LoRa network, the two former ones are easy to reach and they have been solved, but the actual coverage range is still a debated aspect. For applications in smart grid, this is a crucial for a correct designing the number of the base station and estimation of the costs.

To gain insight in this respect, we conducted a coverage experimental test of LoRa networks in the city of Nanjing, China. During the measurements, the gateway was fixed on the top of a five-storey building, in an area where high buildings are present, and it was equipped with an external antenna with gain 5 dBi. The end device was mounted on a car, powered from a laptop via a USB cable. The output power of the end device was set to 20 dBm (100 mW); an end device was equipped with an external antenna with gain 2 dBi. Spreading factor SF was set to 12, and bandwidth was set to 125 kHz. The car moved and when it came to an agreed point, the end device transmitted to the station a packet including a sequence number, node's status and GPS coordinates for 20 times.

Figure 2 shows the results of the test. A base station could cover a cell of about 5 km of radius when using the lowest bit rate, with low margin for possible interference

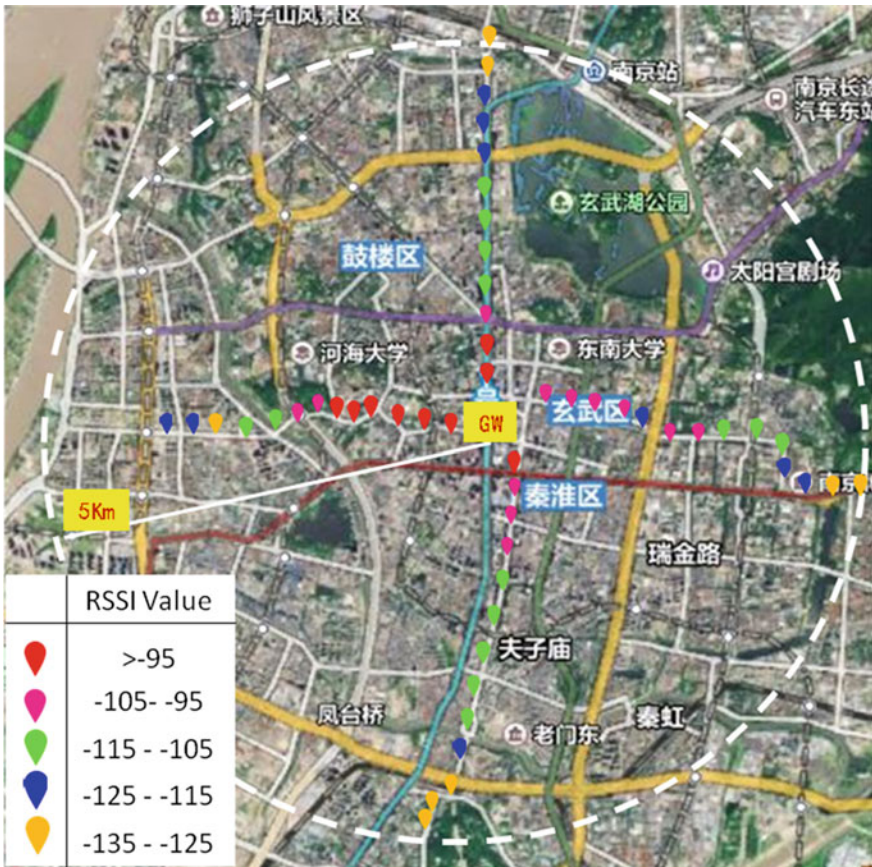


Fig. 2. LoRa system single cell coverage in Nanjing, China.

or to link budget changes. Not surprisingly, when the Received Signal Strength Indicator (RSSI) was at the edge of the radiosensitivity limit, the connection will be lost. It is also worth noticing that the distance is not the only factor affecting the RSSI. In some cases, the closer test points showed lower RSSI than the farther. There are many reasons which may have caused this behavior, such as changing radio environment, different surroundings, and others.

From this test, we can find that the LoRa actual coverage range is much larger than the typical wireless communication technologies, and it is suitable for the low data rate, widely deployment applications.

5 Discussion and Conclusion

In this paper, we have described the most popular wireless technology Low-Power Wide Area Networks (LPWANs). The most prominent LPWAN technologies, i.e., Sigfox, LoRa, and NB-IoT, have been introduced and compared to the current short-range communication standards. Larger coverage, high robustness, and low power consumption are the main characteristics for the LPWANs. The experiment which has been performed based on LoRa networks in an urban environment has shown LPWANs' larger coverage that based on the long-range link.

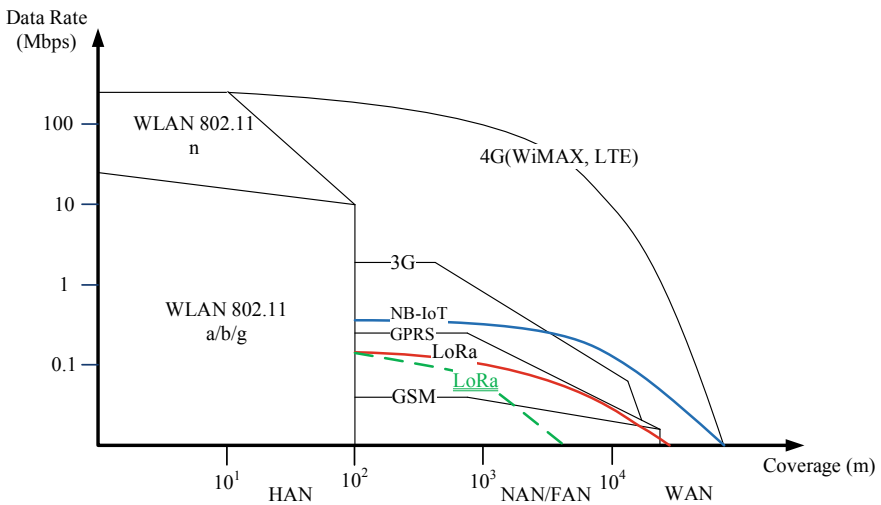


Fig. 3. Data rate and coverage range of wireless technologies. The solid lines depict the maximum theoretical data rate and coverage range of wireless technologies, and the dotted line shows actual coverage range based on the experiment in an urban environment which has been described in Sect. 5.

Figure 3 depicts the maximum theoretical data rate and coverage range of some wireless technologies which can meet the smart grid applications requirements. It shows that the new wireless technologies—LPWANs—could take the place of other wireless technologies, e.g., GPRS, GSM, to implement some applications that require

low data rate in FAN/NAN of the smart grid. The star topology will replace the mesh, and one gateway (base station) will cover a larger number of the end devices. This will bring opportunities and challenges for network topology of the smart grid and reduce the costs of the communication infrastructure greatly.

References

1. Farhangi, H.: The path of the smart grid. *IEEE Power Energ. Mag.* **8**(1), 18–28 (2010)
2. Baimel, D., Tapuchi, S., Baimel, N.: Smart grid communication technologies-overview, research challenges and opportunities. In: *International Symposium on Power Electronics, Electrical Drives, Automation and Motion* (2016)
3. Yan, Y., Qian, Y., Sharif, H., Tipper, D.: A survey on smart grid communication infrastructures: motivations, requirements and challenges. *IEEE Commun. Surv. Tutor.* **15** (1), first quarter (2013)
4. EPRI Technical Report: Assessment of wireless technologies in substation functions Part-II: substation monitoring and management technologies (2006)
5. Raza, U., Kulkarni, P., Sooriyabandara, M.: Low power wide area networks: a survey (2016). [arXiv:1606.07360](https://arxiv.org/abs/1606.07360)
6. Bor, M., Vidler, J., Roedig, U.: LoRa for the Internet of Things. ResearchGate (2016)
7. *Communications Requirements of Smart Grid Technologies*. Department of Energy, Washington, DC (2010)
8. Kuzlu, M., Pipattanasomporn, M., Rahman, S.: Communication network requirements for major smart grid applications in HAN, NAN and WAN. *Comput. Netw.* **67**, 74–88 (2014)
9. Wang, W., Xu, Y., Khanna, M.: A survey on the communication architectures in smart grid. *Comput. Netw.* **55**, 3604–3629 (2011)
10. Raza, U., Kulkarni, P., Sooriyabandara, M.: Low power wide area networks: an overview. *IEEE Commun. Surv. Tutor.* (2017)
11. Augustin, A., Yi, J., Clausen, T., Townsley, W.M.: A study of LoRa: long range & low power networks for the internet of things. *Sensors*, **16**, 1466 (2016)
12. Springer, A., Gugler, W., Huemer, M., Reind, L., Ruppel, C., Weigel, R.: Spread spectrum communications using chirp signals. In *Proceedings of the IEEE/AFCEA Information Systems for Enhanced Public Safety and Security (EUROCOMM 2000)*, Munich, Germany, pp. 166–170, 19 May 2000
13. Berni, A.J., Gregg, W.D.: On the utility of chirp modulation for digital signaling. *IEEE Trans. Commun.* **21**, 748–751 (1973)
14. LoRa Alliance: LoRaWAN What is it. Technical Marketing Workgroup1.0 (2015)
15. Alliance, LoRa: White Paper: A Technical Overview of Lora and Lorawan; The LoRa Alliance: San Ramon, CA, USA (2015)
16. Bardyn, J.P., et al.: IoT: the era of LPWAN is starting now. In: *Proceedings of the 42nd European Solid-State Circuits Conference Lausanne*, pp. 25–30 (2016)
17. Ratasuk, R., Vejlgard, B., Mangalvedhe, N., Ghosh, A.: NB-IoT system for M2M communication. In: *IEEE WCNC, Doha* (2016)
18. 3GPP TR 36.802, Narrowband Internet of Things (NB-IoT), Technical Report TR 36.802 V1.0.0, Technical Specification Group Radio Access Networks (2016)
19. Rohde, D., Schwarz, J.: Narrowband Internet of Things (2016). <https://www.rohde-schwarz.com/us/applications/narrowband-internet-of-things-application-note-56280-314242.html>
20. Nokia: Nokia LTE M2M—optimizing LTE for the internet of things. In: *White paper* (2014)
21. Ratasuk, R., Mangalvedhe, N., Zhang, Y., Robert, M., Koskinen, J.-P.: Overview of Narrowband IoT in LTE Rel-13. In: *IEEE Conference on Standards for Communications and Networking* (2016)



A Low Energy Consumption Ocean Environment Information Collection System Designing

Qiuming Zhao, Hongjuan Yang, and Bo Li^(✉)

Harbin Institute of Technology (Weihai), Weihai, China
libol1983@hit.edu.cn

Abstract. Research on sensory networks for marine environmental information acquisition has received more and more attention. However, the underwater wireless sensor network has the characteristics of limited bandwidth and energy consumption. Therefore, by using the sparseness of sensor node data in frequency domain and space domain, this paper proposes a dual-domain compressed sensing (DCS) low-energy ocean environment information collection scheme. The scheme uses the sparsity of the frequency domain for information recovery, thereby further saving the control overhead of the sink node's downstream sending address frame. Through the result of simulation, it is found that the scheme proposed in this paper is better than the previous IDMA multiple access detection scheme in terms of energy consumption.

Keywords: Compressed sensing · Underwater wireless sensor network · Multiuser detection · Multiple access

1 Introduction

Because of the complexity and particularity of the underwater communication channel environment, most underwater communication uses sound waves as the carrier. Candes and Donoho et al. proposed a compressed sensing theory [1]. The main principle is based on the sparsity of the original information, using a noncoherent measurement matrix to realize the dimensional reduction of the original information, and then at the receiver solves an optimization problem through the dimensionality reduction observation vector, thus reconstructing the original data information with high probability [2]. Since most of the marine environment information is sparse, it can be considered to apply the theory of compressed sensing to marine sensor networks. This can provide a solution to the problems of underwater sensor network resources.

The rest of this paper is organized as follows. Section 2 introduces the system model. In Sect. 3, a marine environment information acquisition scheme based on dual-domain compressed sensing (DCS) is proposed. System simulation results are shown in Sect. 4. Finally, conclusions are drawn in Sect. 5.

2 System Model

In this paper, the sensor nodes are deployed in the underwater two-dimensional plane, and the sink nodes are deployed in shallow seawater depths, the sensor nodes are evenly distributed on the two-dimensional detection plane to collect the required marine environment information and then transmitted to the sink node through the uplink multiple access channel. Therefore, this paper mainly studies an efficient data acquisition and transmission scheme in a large ocean monitoring underwater sensor network. Besides a single-channel multiple access underwater channel model is used, assuming that the power required by each sensor node at the receiving end is P_0 . The distance between the sink node and the sensor node is d , the carrier frequency is f , the transmit power of the node is $P_0 \cdot A(d, f)$, where $A(d, f) = d^c \cdot a(f)^d$, general $c = 1.5$, $a(f) = 10^{\alpha(f)/10}$.

3 Proposed Scheme

In this chapter, we use the sparseness of sensor node data in frequency domain and airspace, and propose a marine environment information collection scheme based on dual-domain compressed sensing (DCS). The system block diagram of DCS scheme is shown as in Fig. 1.

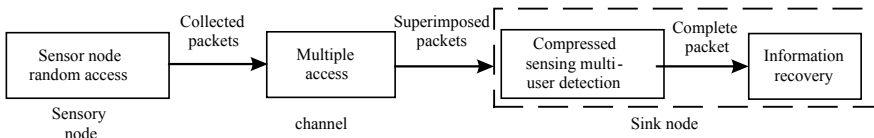


Fig. 1. DCS system block diagram

In traditional sensor networks, all sensor nodes participate in data acquisition and transmission. Fatemeh Fazel et al. proposed a random access compression sensing scheme [3] based on the sparseness of information on the marine environment and proved that the original ocean monitoring information can be reconstructed from the information of some sensor nodes at the sink node. The matrix of the resulting active nodes is defined as Φ , the element ϕ_{ij} needs to satisfy.

$$\begin{aligned}
 \sum_{i=1}^{N_{act}} \phi_{ij} &\leq 1, j = 1, 2, \dots, N \\
 \sum_{j=1}^N \phi_{ij} &\leq 1, i = 1, 2, \dots, N_{act}
 \end{aligned} \tag{1}$$

It shows that there is only one element per line in Φ , which is the position of the active node in the sensor network. There are a total of N_{act} active nodes.

When the data packets collected by the active nodes are delivered to the sink nodes, a stable and efficient multiple access method is needed to process the symbol frames so as to achieve multiuser transmission. Then the signal received at the sink node is

$$y = \sum_{n=1}^N H_n A_n d_n + n \quad (2)$$

where H_n is the matrix form of the impact response of node block fading channels, and A_n is a matrix form of node multiple access processing. d_n is the symbol frame data that the node can potentially transmit, in this paper, using IDMA as a multi-access underwater wireless sensor.

In the literature [4], a CS-CBC multiuser detection algorithm based on compressed sensing in IDMA is proposed, and it is proved that the performance is superior to the traditional compressed sensing algorithm. Therefore, the CS-CBC algorithm is adopted as a multiuser detection algorithm in the DCS scheme.

Because the original information f of the marine environment has sparseness in the frequency domain, it can be recovered by compressive sensing reconstruction technology. In the previous section has obtained a random extraction measurement matrix Φ and symbol frame data d_Γ containing active nodes through the multiuser detection technology. The relationship is

$$d_\Gamma = \Phi f = \Phi \Psi \theta \quad (3)$$

where θ is a sparse representation of the frequency domain of f , and Ψ is a Fourier transform matrix containing N elements. This information reconstruction can be solved by solving the following optimization problems:

$$\min_{\theta} \|\theta\|_1 \quad s.t. d_\Gamma = \Phi \Psi \theta \quad (4)$$

According to compressed sensing reconstruction theory, when the sampled data satisfy the condition, the original data can be recovered with almost no error.

4 Simulation and Analysis

This section systematically evaluates the energy consumption of the network bandwidth based on the dual-domain compressed sensing ocean environment information collection scheme and compares it with the scheme proposed in previous studies. Simulation settings: Acquisition probability $p_a = 0.45$, number of subframes $I = 4$, and single-node ID length is $L_a = 6 \text{ bit}$. Under different signal-to-noise ratios, the simulation curves of DCS scheme and IDMA-based CS scheme reconstruction error and single-node power consumption are shown in Fig. 2.

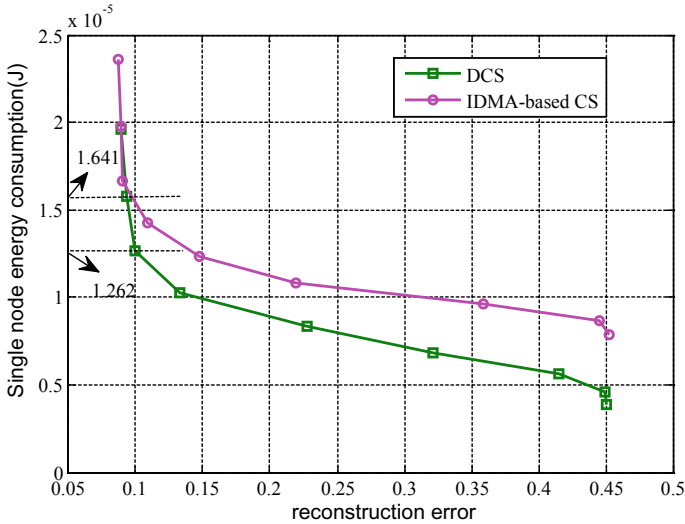


Fig. 2. Curve of the reconstruction error between DCS scheme and single-node energy consumption

Figure 2 shows that under the condition that the accuracy of ocean information recovery does not exceed 0.1, the energy consumption of DCS node is 1.262×10^{-5} J, and the energy consumption of single node of IDMA-based CS scheme is 1.641×10^{-5} J, so the DCS scheme saves 23.1% of single-node energy consumption.

5 Conclusions

This paper studies multiuser detection based on compressed sensing theory in sporadic communication scenarios, continuously exploiting the sparseness of sensor nodes’ data in frequency domain and airspace, and proposes a dual-domain compressed sensing data collection scheme to construct a high-efficiency multipoint stable communication network. The simulation results show that the DCS scheme can not only ensure the reconstruction error performance of the marine environment monitoring network but also outperforms the IDMA-based CS scheme in energy consumption performance. The design of this paper is of great significance to the research and development of marine environmental information monitoring and provides an important reference for the next generation of large-scale, three-dimensional, high-resolution marine environment monitoring sensor network construction.

Acknowledgments. This work is supported in part by National Natural Science Foundation of China (No. 61401118 and No. 61671184), Natural Science Foundation of Shandong Province (No. ZR2018PF001 and No. ZR2014FP016), the Fundamental Research Funds for the Central Universities (No. HIT.NSRIF.201720 and HIT.NSRIF.2016100), and the Scientific Research Foundation of Harbin Institute of Technology at Weihai (No. HIT(WH)201409 and No. HIT (WH)201410).

References

1. Donoho, D.L.: Compressed sensing. *IEEE Trans. Inf. Theory* **52**(4), 1289–1306 (2006)
2. Liu, G., Kang, W.: IDMA-based compressed sensing for ocean monitoring information acquisition with sensor networks. *Math. Probl. Eng.* **1**, 1–13 (2014)
3. Fazel, F., Fazel, M., Stojanovic, M.: Random access compressed sensing for energy-efficient underwater sensor networks. *IEEE J. Sel. Areas Commun.* **29**(8), 1660–1670 (2011)
4. Li, B., Du, R., Kang, W., Liu, G.: Multi-user detection for sporadic IDMA transmission based on compressed sensing. *Entropy* (2017)
5. Wang, B., Dai, L., Mir, T., et al.: Joint user activity and data detection based on structured compressive sensing for NOMA. *IEEE Commun. Lett.* **20**(7), 1473–1476 (2016)
6. Liu, G., Kang, W.: IDMA-based compressed sensing for ocean monitoring information acquisition with sensor networks. *Math. Probl. Eng.* **2014**(1), 1–13 (2014)

System Security



Cyberspace Security Evaluation Technology on the Condition of Attack and Defense Confrontation

Zhang Bo, Wen Tao^(✉), Lei Jing, Yang Yunxiang, and Guo Jing

China Academy of Electronics and Information Technology, Beijing 100041,
China

{edchang, wt_bj_2016, yyxsdu}@126.com,
leijanet@163.com, 572179714@qq.com

Abstract. The blurred synthesis evaluation arithmetic based on analytic network process and systemic security evaluation target system is designed, and the main technology and system realization of cyberspace security evaluation system are presented, which is realized based on fight theory's attack and defense affection evaluation and systemic security evaluation technology. The technology and system can provide a systemic security evaluation to information system target's security and put forward security resolve scheme suggestions, which have important effect and meaning on accelerating China's cyberspace war ability.

Keywords: Cyberspace · Attack and defense confrontation · Security evaluation

1 Introduction

Cyberspace is a globalized information environment that radiates all aspects of human production, life, social activities, and even ideology, as well as the stability of the government, the security of national defense, social stability, economic development, and scientific and technological progress. It has become an important part of national sovereignty and is a new area of national interest [1]. A series of incidents such as “Prism” and “Seismological Network” also show that cyberspace confrontation has been changed from speculation to reality, and information security conflicts at the national level have become increasingly fierce. However, our country lacks a sound defense system, infrastructure, and advanced self-controllable technologies and products in cyberspace security construction, which puts our country in a serious threat to cyberspace security [2].

Research on cybersecurity assessment under the conditions of attack and defensive confrontation can provide methods and means for assessing and verifying network attack and defense confrontation test, offensive and defensive countermeasure

equipment, and security defense system, thereby improving our cyberspace operational capability and testing and evaluation capability.

Traditional information security assessment [3] is the process of evaluating information systems and security attributes such as the confidentiality, integrity, and availability of information handled, transmitted, and stored by them in accordance with relevant information security technology and management standards. The accurate assessment of the security status of the entire information system can help managers grasp the security vulnerabilities and operational risks existing in the whole network, identify the security vulnerabilities that the system needs to be patched, the security measures to be taken, propose security solutions, and improve Security protection, to ensure the safety of information.

In this paper, the evaluation index system and evaluation algorithm are developed to systematically evaluate the security of various complex heterogeneous networks, cyberspace attack and defense confrontation test, counterattack and defense equipment, and information system security, and at the same time, evaluation of the effectiveness of the implementation. At the same time, it also enables system-level, integrated, and automated testing and evaluation of various real or theoretical security systems. It also provides a variety of security solutions, security model, security offensive and defensive theory, security technology measures, and safety product evaluation and verification.

2 Evaluation Algorithms and Evaluation Index System Design

2.1 Evaluation Algorithm

This paper adopts the fuzzy comprehensive evaluation algorithm based on the Analytic Network Process (ANP). The evaluation algorithm is described in detail below.

2.1.1 Network Analysis

The Analytic Network Process (ANP), proposed by Prof. Saaty TL in 1996, is a scientific decision-making method that adapts to a complex structure. It uses the network structure to represent the relationship between elements in the system instead of A simple hierarchical structure is a new decision-making method developed based on the ANP method. Its theory describes the relationship between objective things more accurately and is a more effective and practical decision-making method.

2.1.2 Fuzzy Comprehensive Evaluation Method

Fuzzy comprehensive evaluation method was developed by the United States of America cybernetic expert Zadeh, L.A. of fuzzy sets (fuzzy sets), which uses the principle of fuzzy relationship synthesis, with a variety of attributes, the attributes of the boundary between the unclear, difficult to quantify a reasonable combination of things with the overall judgment.

2.2 Evaluation Index System

Indicators are tools for evaluation and are indicators of the attributes of an object. The index system is a set of relevant indicators based on the requirements of assessment objectives and assessment contents, and collects the relevant information materials of the assessment objects to reflect the basic appearance, characteristics, and level of the assessment objects.

System-level security assessment index tree contains five levels, the most abstract system security step by step to implement the most specific information equipment, the system parameters to ensure that the system-level security has tangible content, security assessment has operable Sex. Among them, the Tier 3 indicators are related to each other and have a different influence on the high-level indicators. Therefore, their determination of the upper-level indicators is based on multiple experiments and expert experiences.

The corresponding assessment method takes into account the mutual enhancement, supplement, and weakening of the security components of the information system components, and deduces the score of the system-level security assessment index from the source data of each component of the target system acquired, and calculates the overall system security assessment score. Finally, with the evaluation criteria, the conclusion of the safety evaluation of the target system is given, the weak point of the system is pointed out, and some suggestions for improvement are made. System-level security assessment indicator tree is shown in Fig. 1.

3 Cyberspace Security Assessments of the Main Technologies

Cyberspace Security Assessments of the Main Technologies contains (1) Based on Game Theory Attack and Defense Confrontation Effect Evaluation Techniques; (2) System-level Security Assessment Techniques.

- (1) Based on Game Theory Attack and Defense Confrontation Effect Evaluation Techniques is a set of strategy choice theory, a branch of applied mathematics [4]. Based on the game theory of offensive and defensive countermeasure evaluation technology is to study in a particular cyberspace, offensive and defensive confrontation may be adopted by both the behavior planning and its probability. Using prior knowledge to assess the possible probability distribution of the two side's offensive and defensive strategy, get the best strategy of both attack and defense. At the same time, comprehensive consideration of operating costs, negative impact, participation losses, and other factors to achieve the offensive and defensive confrontation under the conditions of the implementation of the evaluation.
- (2) By analyzing and comparing various existing safety assessment standards and interoperability standards, according to the applicability of each standard and combining with the actual demand of network attack and defense confrontation test, the appropriate standard is selected as the basis for safety assessment [5].

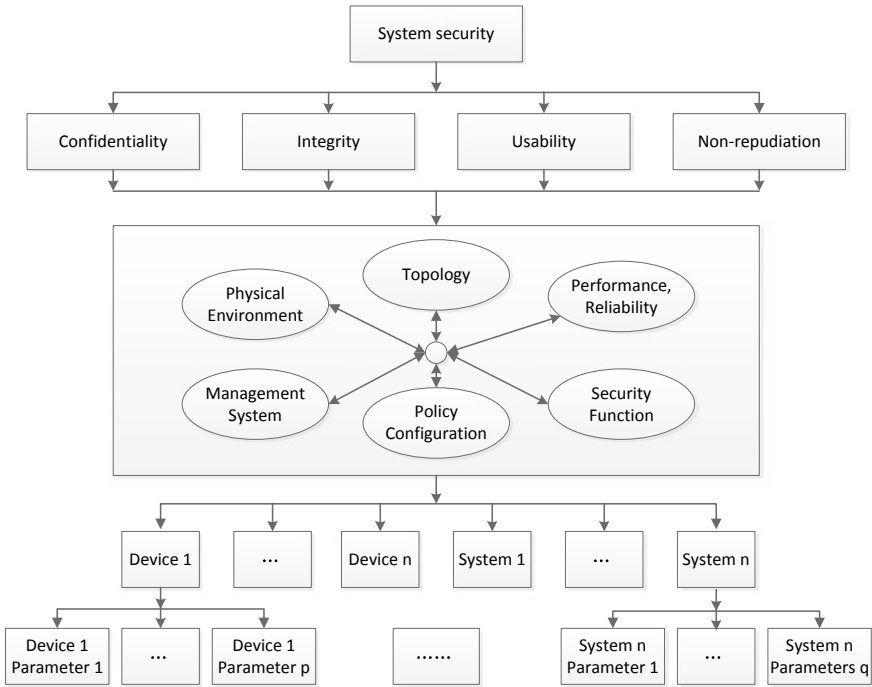


Fig. 1. System-level security assessment indicator tree

4 Safety Assessment System Technologies

The purpose of safety assessment is to find out the weak points of the system, find out and correct the weaknesses and loopholes in the system, and let the decision makers carry out the risk disposal, that is, to adopt a comprehensive strategy to solve the risk and put forward the safety solution to ensure the safety of the system.

Security assessment is a complex large-scale systems engineering which we must use to solve the complex system of large-scale solution, and we use a large-scale network of high-speed fast reproduction techniques, through loading, simulation of high-fidelity business and red and blue attack and defense confrontation to achieve the goal System Security Test Evaluation.

4.1 Based on Offensive and Defensive Game Safety Assessment Tools

Based on offensive and defensive game safety assessment is through the establishment of attack and defense model to simulate the offensive and defensive sides. Offensive and defensive process is the process of the game between the two sides, as the attacking party and defensive side of the game are affected by the environment and each other. Risk Assessment According to the assessment criteria, a systematic safety assessment is conducted by assessing the attack/defensive effects on both sides.

Based on offensive and defensive game safety assessment tool is mainly composed of three parts: the game red side, game blue square, risk assessment module. Among them, the game Red Square and Game Blue Square are both parties playing games with each other. They define their own attack models and defense models, and make use of offensive and defensive tactics to carry out offensive and defensive game. The attack model analyzes the assets of the system and exploits the vulnerabilities and security vulnerabilities to simulate the attack. The defense model protects the integrity, confidentiality, and non-repudiation of the system assets from being damaged. Risk assessment module is according to the offensive and defensive effects of the two sides of the game and using evaluation criteria to give the system risk assessment level. The overall structure is shown in Fig. 2.

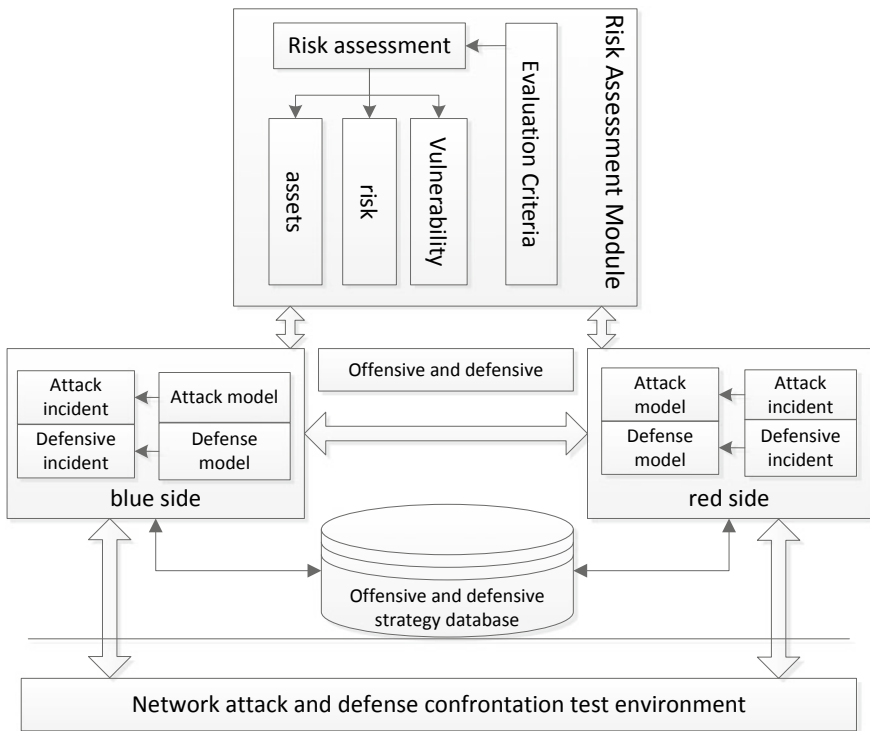


Fig. 2. Attack and defense game-based security assessment tool

The final game evaluation tool for real-time data collection of current network status information, the collected data information such as logic-based data model data log language description. Combined with the attack knowledge base, the collected data are inferred, so as to obtain the possible nodes in the current network and possible attacks. Combining the impact of offensive and defensive knowledge base on each threat, the offensive and defensive game model is used to calculate the probability of threat and the risk of the current system attack is calculated according to the results of

the first two parts. The corresponding system threats are drawn from the risk information obtained from the game risk calculation module Risk change chart.

4.2 System-Level Security Assessment Tools

System-level security assessment based on the in-depth study of the information security assessment standards at home and abroad, the evaluation index system is completed and a mathematical model is established to develop a product-system-level security assessment tool with independent intellectual property. Through the establishment of a scientific and complete evaluation index system, combined with a variety of technical testing methods used to assess the network offensive and defensive safety against the experimental environment.

The purpose of system-level security assessment is to discover the vulnerabilities and defects of the target system, to form assessment reports and suggestions for improvement so as to take corresponding protective measures to enhance the security.

The process of general risk assessment mainly includes asset identification, threat identification, vulnerability identification, security measures analysis, probability analysis, impact analysis, and final security decision [6]. In this paper, the characteristics of the target information system and the design of the system-level security assessment of the main process are shown in Fig. 3.

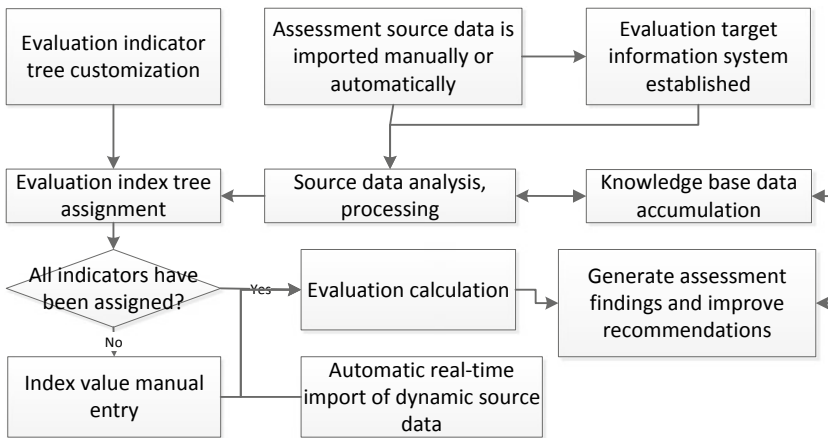


Fig. 3. System-level security assessment flowchart

Evaluation tools to assess the network offensive and defensive confrontation test environment, combined with the assessment results obtained, the need for the target system security protection system defects and loopholes provide a positive solution, the only way to further improve the safety of the target system to provide a powerful help for the information system security to provide a reference and guidance.

System-level security assessment tool itself is based on a certain standard norm for the assessment based on the overall level of security of the system under test to conduct

an in-depth study and effective exploration, so after the conclusion of the assessment can be targeted for a specific standard regulate, provide the solution that meets the requirements.

5 Conclusions

In this paper, the network attack and defense confrontation test environment, the network security equipment, security strategy, business systems, offensive and defensive modeling and simulation studies, and carried out under attack confrontation cyberspace security assessment technology research to achieve the evaluation criteria, method test and evaluation result verification, a set of system-level cyberspace security assessment index system is proposed to realize the analysis and evaluation of the typical target system's security protection effectiveness. The system can be targeted to the security problems of the target information system-level cyberspace security solutions which are proposed to lay a solid foundation for enhancing the combat capability of cyberspace in China.

Acknowledgements. This paper is supported by Beijing NOVA Program (Z1811000062 18041).

References

1. Liu, X., Chen, M., Li, Y.-B., Yu, J.: Cyber War, pp. 2–3. National Defense Industry Press, Beijing (2011)
2. Zeng, Q., Xu, F., Zhang, Y.: Information Security Framework, pp. 15–20. Publishing House of Electronics Industry, Beijing (2010)
3. Wang, C., Chen, J., Hu, C., Sun, Y.: Kernel matrix learning with a general regularized risk functional criterion. *J. Syst. Eng. Electron.* **21**(1), 72–80 (2010)
4. Niu, T., Tian, Z.: The network active defense research based on the game theory and the analysis of network vulnerabilities. *Intell. Comput. Appl.* **2**(3), 57–58 (2012)
5. Zhang, X., Yao, S., Tang, C.: Assessing the risk situation of network security for active defense. *Wuhan Univ. J. Nat. Sci.* **11**(6), 1718–1719 (2006)
6. Xiang, H., Li, F., Banghua, Z.: Information Security Test with Evaluation and Risk Assessment, pp. 313–314. Beijing: Publishing House of Electronic Industry (2009)



Interoperability Performance Analysis and Assessment of B2a/L5/E5a

Weiye Shuai^(✉), X. R. Dong, Di Yan, Jun Wang, and Wei Fu

Space Engineering University, Beijing, China
18803348431@163.com

Abstract. Compatibility and interoperability are the main developed directions for the future of GNSS. The B2a signal transmitted by BDS-3 will interoperate with GPS L5 signal and Galileo E5a signal. The performance of interoperable signals will directly affect the performance of user receivers. This paper systematically analyzes and assesses the interoperability of B2a/L5/E5a signals. By briefly introducing and analysis the signal systems and key technologies interoperability, the theoretical basis of the performance assessment of the interoperable signals is laid down. The interoperability performance is mainly assessed from two aspects: signal character and service performance. In terms of signal character, the same carrier frequency, chip rate, and signal bandwidth are selected for the interoperable signals, and the similar signal modulation method is adopted to realize low power consumption, small size, and low cost of the receiver terminal. Seeing from the service performance, through the unification of coordinate and time reference systems, interoperable signals achieve better performance in the accuracy, continuity, and availability than the single-system signal, and achieve the design goal of interoperable signals.

Keywords: B2a/L5/E5a signals · Interoperability · Performance · Assessment

1 Introduction

Compatibility and interoperability have become a hotspot research in the field of GNSS in recent year with the modernization of GPS and GLONASS and the springing up of BDS and Galileo [1]. Compatibility and interoperability can use multifrequency observation provided by multi-constellation for PNT applications without limit and reduce the dependence on a single constellation, then improve the service performance of navigation system. Galileo's E5a and E1 will interoperate with GPS on the L5 and L1 bands, respectively. GLONASS has also planned to add CDMA signals on the L5 and L1 bands to interoperate with GPS in its modernization plan. QZSS will also be fully compatible and interoperable with GPS on L1, L2, and L5 bands. BDS-3 adds B1C and B2a signals which have better performance and becomes the first system to achieve compatibility and interoperability with the other three global systems. And BDS-3 will also enable interoperability with GPS and Galileo on the B1C/L1C/E1 and B2a/L5/E5 signals [2].

Currently, the research on compatibility mainly focuses on the mutual interference of signals between different systems. Unlike compatibility, interoperability is more extensive and relates not only to signals but also to space and time benchmarks from various GNSS, and has a more direct influence on users. Therefore, this paper takes B2a/L5/E5 signals as an example to analyze and assess the performance of interoperable signals from aspects of signal system, key technologies, signal designed parameters, and service performance. The proposed method can provide some reference for the design of multi-constellation receiver and practical application of multi-constellation compatibility and interoperability.

In Sect. 2, the analysis of signal systems of B2a/L5/E5 is briefly introduced. Section 3 focuses on the analysis of key technologies involved in signal interoperability. The assessment result of interoperable signals is given in Sect. 4. The results of the measured data verify the excellent performance of the interoperable signals. And the proposed method can provide some reference for the design of multi-constellation receiver and practical application of multi-constellation compatibility and interoperability.

2 Signal Systems of B2a/L5/E5a

1. B2a signal

B2a signal is the second civil signal of BDS-3, which is used to replace the B2I signal of BDS-2. It mainly serves dual-band or tri-band receivers and can be used for high-performance services such as life safety services and high-precision measurement. The carrier frequency of B2a signal is 1176.45 MHz and the bandwidth is 20.46 MHz. The data and pilot quadrature structure (QPSK) is adopted. The data component is modulated with the navigation message data and the ranging code, while the pilot component is modulated with ranging code only. They both adopt BPSK(10) modulation.

2. L5 signal

The GPS L5 signal is the third civil GPS signal transmitted on the Block IIF satellites in the process of GPS modernization. Its signal carrier frequency is the same as that of B2a, with a bandwidth of 24 MHz. The signal includes both in-phase and quadrature components, namely, L5I and L5Q. The L5I component is modulated with I5 pseudocode and CNAV message data. The L5Q signal component is the pilot component which is modulated the Q5 pseudocode only. Their modulation modes are both BPSK (10). The power ratio of the two signals is 1:1. The signal has a stronger anti-interference ability owing to a wider bandwidth and higher transmitted power [3].

3. E5a signal

Galileo E5 signal shares the same carrier frequency with B2a and L5, whose bandwidth is 51.15 MHz and modulation mode is AltBOC. For the receiver, the E5 signal can be regarded as being consisted of E5a and E5b, which are two independent sideband

signals. Each sideband signal can be considered as the data and pilot components with BPSK(10) modulation [4]. The spectrum of B2a signal is coinciding with the GPS L5, mainly for free public service. What should be specially explained here is that the B2a signal transmitted by BDS-3 test satellites also adopts the AltBOC modulation.

In summary, B2a signal, L5 signal, and E5a signal are not only having the same carrier frequency but also having the same signal modulation method, chip rate, period, and other parameters, which facilitate achieving the compatible signal tracking and positioning of B2a/L5/E5a in one receiver by configuring different software parameters, thereby improves the compatibility and interoperability of multisystem.

3 Interoperability Key Technologies of B2a/L5/E5a

The research on interoperability is often focused on both the system and signal levels. At the system level, interoperability can be considered as the ability of providing consistent solutions by multiple navigation systems signals with a single-system signal under the same boundary constraints. In other words, a multiuser receiver and a single-system user receiver can provide the same precision navigation solution. At the signal level, interoperability can be considered as that users' navigation receivers can receive similar signals transmitted by different satellite navigation systems without major changes. To achieve interoperability, we must consider the following factors:

1. Signal Characteristic Parameters

The main characteristic parameters of satellite navigation signals are composed of carrier frequency, modulation mode, signal bandwidth, signal power, spreading code, text format, and so on. From the perspective of the user terminal, in order to ensure a multisystem receiving device with good performance, low power consumption, small size, and low cost, the above parameters are always expected to be as similar as possible so that to share the hardware and software in one receiver. In particular, the main characters that are closely related to the spectral characteristics of the signal, such as carrier frequency, signal bandwidth, and modulation method, should preferably be exactly the same.

2. Space Reference Coordinates

Due to the different coordinate frames adopted by the four major global satellite navigation systems, the relative deviation of the coordinate frames will affect the interoperability. In practical applications, a common method is to move the satellite coordinate involved in each system to a unified reference frame and then convert to the reference coordinate frame required by the user. The common method of coordinate frame conversion is the Bursa model.

Assuming that the coordinate in coordinate frame 1 is $[X_1 \ Y_1 \ Z_1]$, while the coordinate in coordinate frame 2 is $[X_2 \ Y_2 \ Z_2]$. $x_0, y_0, z_0, \varepsilon_x, \varepsilon_y, \varepsilon_z, m$ are conversion parameters. The Bursa model is shown below:

$$\begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \begin{bmatrix} m & \varepsilon_z & -\varepsilon_y \\ -\varepsilon_z & m & \varepsilon_x \\ \varepsilon_y & -\varepsilon_x & m \end{bmatrix} \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} \quad (1)$$

In this paper, the final positioning results are in the CGCS2000 coordinate frame. Select ITRF 2008 as a unified reference coordinate system, and convert satellite coordinates to the ITRF 2008. Then convert them to the CGCS2000. The conversion parameters between the four major navigation systems coordinate frames and ITRF 2008 are given in [5].

3. Time Reference System

The BDT-GNSS time synchronization parameters are transmitted in the navigation message of B2a signal, being used to calculate the time deviation between the BDS and other GNSS systems. The calculation method is as follows:

$$\begin{aligned} \Delta t_{\text{Systems}} &= t_{\text{BD}} - t_{\text{GNSS}} \\ &= A_{0\text{BGTO}} + A_{1\text{BGTO}}[t_{\text{BD}} - t_{0\text{BGTO}} + 604800(WN - WN_{\text{BGTO}})] + \\ &\quad A_{2\text{BGTO}}[t_{\text{BD}} - t_{0\text{BGTO}} + 604800(WN - WN_{\text{BGTO}})]^2 \end{aligned} \quad (2)$$

where $A_{0\text{BGTO}}$ is the bias coefficient, $A_{1\text{BGTO}}$ is the drift coefficient, $A_{2\text{BGTO}}$ is the drift rate coefficient, $t_{0\text{BGTO}}$ is the reference time number, and WN_{BGTO} is the reference week number.

4 Interoperability Key Technologies of B2a/L5/E5a

At the signal level, interoperability is manifested in that it requires the frequency, signal system, time scale, and space coordinate frame to meet the requirements. At the user level, receivers are required to be able to achieve equivalent or better performance, while being required low power consumption, small size, and low cost, to achieve the optimal configuration of interoperable receiver. Therefore, interoperability not only needs to be assessed from the signal level, but it also needs to be assessed from the user's service performance. To this end, a comprehensive assessment method of interoperability is proposed in terms of both signal character and overall service performance [6].

4.1 Signal Characteristic Parameters Analysis

Interoperability of space signals is mainly achieved by sharing carrier frequency and spectrum overlap, while modulation mode, multiple access mode, chip rate, and chip shaping are similar. It can be seen from the data in Table 1 that the carrier frequency, modulation mode, signal bandwidth, chip rate, and the like parameters of the three types of signal are basically consistent, and meanwhile, it ensures that the user's receiver can receive these navigation signals without major changes. Therefore, from the perspective of the signal characteristic parameters, the three types of signals have good interoperability.

Table 1. Signal characteristic parameters

Signals and components	Signal characteristic parameters				Carrier frequency (MHz)	Modulation	Bandwidth (MHz)	Coordinate system	Time system
	Pseudocode		Navigation message						
	Chip rate (Mcps)	Type	bit/symbol						
B2a	B2a_data	10.23	CNAV2		1176.45	BPSK(10)	20.46	CCGS2000	BDT
	B2a_pilot		Pilot			BPSK(10)			
L5	L5I	10.23	CNAV		1176.45	BPSK(10)	20.46	WGS-84	GPST
	L5Q		Pilot			BPSK(10)			
E5a	E5a-I	10.23	F/NAV		1176.45	BPSK(10)	20.46	GTRF	GST
	E5a-Q		Pilot			BPSK(10)			

4.2 Service Performance Assessment

The ultimate goal of interoperability is to give users better service. A systematical method to assess the performance of interoperable signals is proposed below. The service performance parameters are shown in Fig. 1.

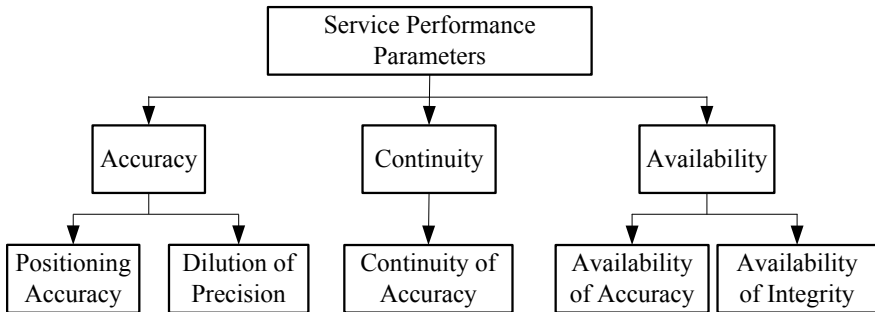


Fig. 1. Frame diagram of service performance parameters

In this paper, the service performance assessment of B2a/L5/E5a signals is performed using 12-h dynamic measurement data of BDS/GPS/Galileo compatible receiver from 8 a.m. to 8 p.m. on October 9. The test environment is an open and unobstructed area. B2a signal is transmitted by BDS-3 test satellites. The reference for positioning is provided by the Novatel dlv3 GPS post-PPP results. The service performance of interoperable signals is assessed by being compared with the GPS L5 single-frequency signal.

1. Assessment results of accuracy

Figure 2a is a comparison of the positioning results of L5 single-frequency signal and interoperable signals. Figure 2b is the PDOP of satellites participating in position for the two kinds of signals. As can be seen in the figures, the positioning results and PDOP of the interoperable signals are better than that of the GPS L5 single-frequency signal. Judging from the entire observation period, the horizontal positioning accuracy of interoperable signals 5.65 m and the vertical positioning accuracy is 7.88 m. Mean PDOP is 1.78.

2. Assessment results of accuracy

Accuracy continuity refers to the probability that the system will continuously meet the accuracy requirements during the test period. It is usually calculated by the mean interval time (MTBF) of the accuracy fault. The probability of system accuracy continuity can be expressed as:

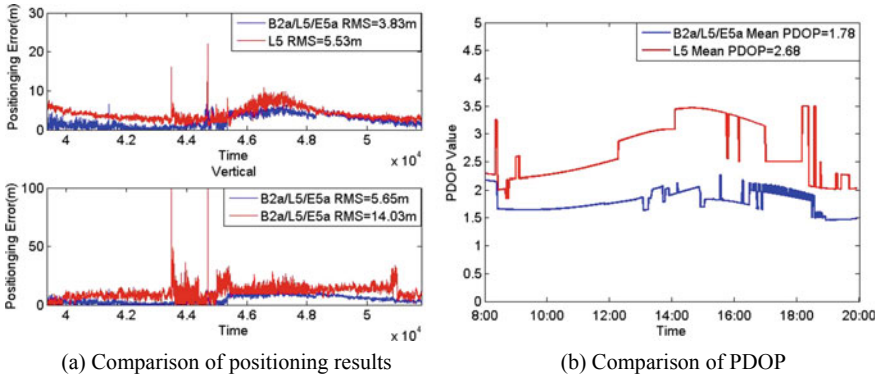


Fig. 2. Positioning accuracy results schematic

$$P_{con} = e^{-(1/MTBF)} \tag{3}$$

Reference [7] is selected as the assessment criteria. The threshold of horizontal positioning error is 9 m and that of the vertical positioning error is 15 m. When the positioning error exceeds the threshold, the accuracy fault occurs. This paper argues that if any fault occurs in any direction, the accuracy fault occurs on that epoch. Judging from (3), it can be calculated that the MTBF of interoperable signals and L5 signals is 210.56 s and 74.89 s, and the corresponding probability of continuity is 99.53% and 98.67%, respectively.

3. Assessment results of availability

For users, the availability of the navigation system is the probability that the system will meet the requirements of availability of accuracy and integrity, which can be calculated as follows:

- Availability of accuracy.

$$Ava_{\delta} = \frac{\sum_{t=t_{start}, inc=T}^{t_{end}} bool\{\delta_t \leq T_{acc}\}}{1 + \frac{t_{end} - t_{start}}{T}} \tag{4}$$

where $[t_{start}, t_{end}]$ is the observation session, and T is the time interval. T_{acc} is the threshold accuracy availability, and the values are the same as those in continuity. $bool\{\}$ is the Boolean function.

- Availability of integrity

In the user segment, the integrity of the system mainly relies on the RAIM algorithm. At this level, the intact availability of the system is the availability of the RAIM algorithm, which is often characterized by RAIM black holes. In this paper, the Fault Detection Exclusion Holes (FDE black holes) is used. When the number of visible satellites is less than, RAIM algorithm cannot realize fault detection. In this case, FDE black holes are occurred and calculated as follows:

$$Ava_{\delta} = \frac{\sum_{t=t_{start, inc=T}}^{t_{end}} bool\{\delta_t \leq T_{acc}\}}{1 + \frac{t_{end} - t_{start}}{T}} \tag{5}$$

Then availability can be expressed as:

$$Ava = Ava_{\delta} \cdot Ava_{FDE} \tag{6}$$

It can be seen from Fig. 3 that the number of visible satellite was 6 only in some epochs around 16:00 during the whole observation period, that is, the FDE black hole occurred at those epochs. Combined with the positioning results above, the availability of interoperable signals and L5 signal can be calculated as follows:

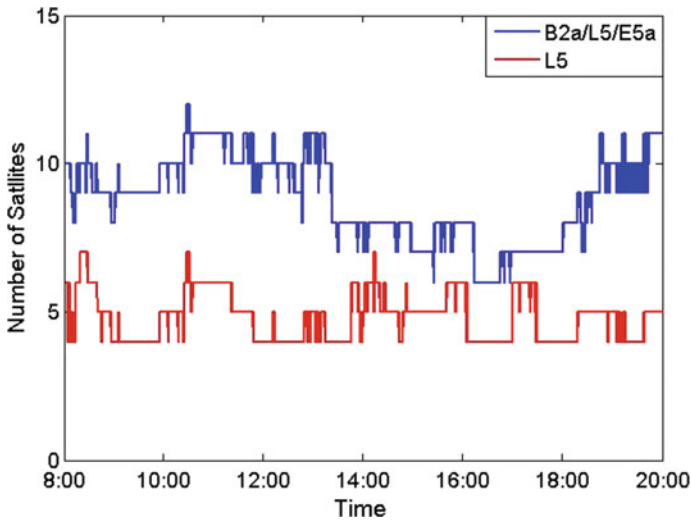


Fig. 3. Number of visible satellites

$$\begin{aligned}
 A_{va_{B2a/L5/E5a}} &= A_{va_{B2a/L5/E5a}} \delta A_{va_{B2a/L5/E5a}FDE} = 99.99\% \\
 A_{va_{B2a/L5/E5a}} &= A_{va_{L5}} \delta A_{va_{L5}FDE} = 64.76\%
 \end{aligned}
 \tag{7}$$

Above all, the service performance of interoperable signals is obviously better than that of L5 signal. So the B2a/L5/E5a signals achieve the design goal of interoperable signals.

5 Conclusions

Satellite navigation signals are the carriers of various positioning, navigation, and timing applications and have also become the most important symbol of the upgrading of satellite navigation systems. Compatibility and interoperability are the future trends of the satellite navigation system. This paper analyzed and assessed the interoperability of B2a/L5/E5a signals from several aspects such as signal systems, key technologies, signal character, and service performance. From the signal level, the signals share the same carrier frequency, chip rate, signal bandwidth, and other characteristic parameters. Together with using a similar modulation mode, the receiver does not need to make major changes to receive the three kinds of signals. Therefore, the goal of low power consumption, small size, and low cost of receiver terminals has been achieved. From the perspective of service performance, interoperable signals realized better navigation and positioning performance than single-system signal in terms of accuracy, continuity, and availability. At the same time, the method presented in this paper is universal and can be used for the performance analysis and assessment of the B1C/L1C/E1 interoperable signals in the next step.

References

1. Yang, Y.X.: Some notes on interoperability. *Acta Geod. Cartogr. Sin.* **45**(3), 253–259 (2016)
2. Tan, S.S.: Innovative development and forecast of BeiDou system. *Acta Geod. Cartogr. Sin.* **46**(10), 1284–1289 (2017)
3. Xie, G.: *Principles of GNSS: GPS, GLONASS, and Galileo*. Publishing House of Electronics Industry, Beijing (2013)
4. Hein, G., Godet, J., Issler J., Martin J., Erhard P., Lucas-Rodriguez, R., Pratt, T.: Status of Galileo Frequency and Signal Design. *ION GNSS*, Portland, OR, pp. 24–27, Sept 2002
5. Zeng, A.M.: A methodology for assessing interoperability for global satellite navigation system. In: *International Symposium on GPS/GNSS 2010*, Taipei, Taiwan (2010)
6. Zeng, A.M., Yang, Y.X., Ming, F., Jing, Y.F.: Comparison of ellipsoids for GNSS geodetic coordinate system. *J. Geomat.S Sci. Technol.* **33**, 551–556, June 2016
7. U.S.A. Department of Defense: *Global Positioning System Standard Positioning Service*, 4th edn, Sept 2008



Physical Layer Security of a Buffer-Aided Relay Selection for Underlay Cognitive Radio Network

Jian Jia, Ting Jiang^(✉), Wei Guo, and Xiaoying Qiu

Key Laboratory of Universal Wireless Communication, Ministry of Education,
Beijing University of Posts and Telecommunications, Beijing, China
{tjiang,qxy}@bupt.edu.cn, {bupt-jiaj,gwnyj}@163.com

Abstract. This paper investigates the physical layer security in underlay cognitive radio networks with multiple buffer-aided decode-and-forward relays, where an eavesdropper can intercept the data transmission from the source and relay nodes. Different from existing works, we propose a new buffer-aided relay selection scheme in multi-relay scenario instead of a link selection scheme in one relay scenario. Moreover, finite data buffers are assumed to be available at every relay to avoid selecting the best source-to-relay and relay-to-destination links concurrently, and the secrecy outage probability is derived. The proposed scheme is evaluated using simulations and theoretical results, which show that proposed scheme has performance advantage over the conventional unbuffered relay selection scheme.

Keywords: Physical layer security · Buffer-aided relay network · Cognitive radio network

1 Introduction

Buffer-aided relay is a promising technology to improve diversity gains, transmission efficiency and security performance. Incorporating the instantaneous strength of the wireless links as well as the status of the finite relay buffers, buffer-aided relay selection scheme can select different relays for transmitting and receiving, respectively. In contrast, for a data packet transmission, conventional unbuffered relay selection scheme, namely, max-min scheme [1], must determine the best transmitting and receiving links concurrently and adopt the same relay. We aim to investigate the security performance of buffer-aided relay selection scheme in cognitive radio network (CRN) in terms of the secrecy outage probability (SOP).

A max-link buffer-aided relay selection scheme, which made the relay selection decision based on the strongest available link, was proposed for decode-and-forward (DF) cooperative wireless network [1]. Inspired by the max-link scheme,

the authors proposed a max-ratio multi-relay selection scheme in DF buffer-aided cooperative network with an eavesdropper [2].

One promising application area for cooperative communication is CRN to solve the spectrum scarcity problem. In the underlay model of CRN, the SUs are permitted to utilize the licensed band as long as the PUs link quality can be guaranteed [3], which means that there is a power constraint in SUs so that the interference of primary receiver must be kept under a certain threshold. To improve the performance of SUs, the cognitive buffer-aided relay network has been investigated. The authors proposed a link selection scheme in underlay CRN with one DF buffer-aided relay where the throughput of the secondary networks was maximized [4]. A DF buffer-aided relay selection for underlay CRN was proposed where the best relay was selected with the highest signal-to-interference ratio (SIR) among all available links [5]. The authors investigated the secure performance of a link selection scheme with one DF buffer-aided relay in CRN [6].

To improve the physical layer security of underlay CRN, we propose a new relay selection scheme with multiple buffer-aided DF relays in the face of eavesdropping attack, which is different from existing works from the following three aspects:

- Since power constraints and mutual interference must be taken into consideration in CRN comprehensively, the relay selection schemes [1, 2, 7] for general cooperative networks is not available.
- Instead of link selection schemes of one relay [4, 6], we propose a relay selection scheme with multiple buffer-aided relays to take the full advantage of channel states in every time slots and achieve better security performance.
- Different from relay selection schemes without considering eavesdropping attack [5] and only considering the eavesdropping on AF relays [8], the proposed scheme focus on security performance in the face of eavesdropping on DF relays as well as secondary source.

We have derived the expression of SOP and analyzed the security performance under our method relative to that of conventional max-min relay selection scheme without buffer-aided relays. Furthermore, the performance of our proposed scheme is verified via simulation results.

The remainder of paper is organized as follows. Section 2 describes the system model and proposes a new relay selection scheme. Section 3 analyzes the secrecy outage probability of proposed scheme and conventional scheme. Section 4 gives simulation results to verify the proposed scheme. Section 5 concludes this paper.

2 System Model and Proposed Buffer-Aided Relay Selection Scheme

2.1 System Model

In underlay CRN, as show in Fig. 1, the PUs consist of one primary source p_s and one primary destination p_d , and the SUs have three parts, one secondary source s , one secondary destination d , and a cluster of half-duplex DF relays r_k ,

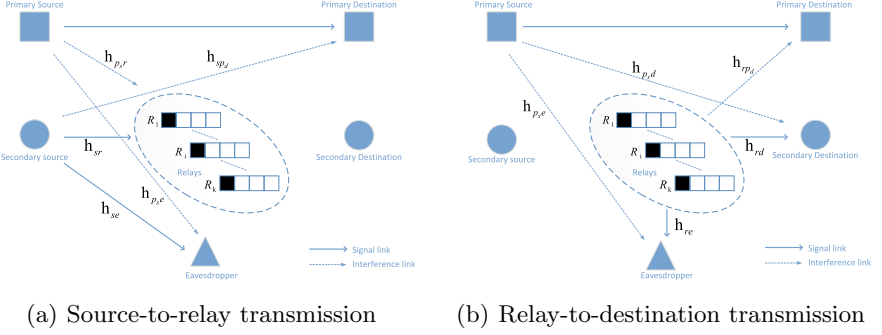


Fig. 1. Relay selection in the CRN.

$k \in (1, 2, \dots, K)$. The channels, $\{h_{ij}\}$, $i \in \{p_s, s, r_k\}$, $j \in \{r_k, p_d, d, e\}$, between the transmitter and receiver are independent quasi-static Rayleigh fading so that the channel coefficients remain unchanged during one packet duration but independently vary from one packet time to another. Channel coefficients, channel gains, and average channel gains are denoted as h_{ij} , $\gamma_{ij} = |h_{ij}|^2$, $\lambda_{ij} = E|h_{ij}|^2$ and channel gains γ_{ij} is exponential distributed with variance $1/\lambda_{ij}$. The secondary-transmission-related channels can be divided into four groups similar as [5]: secondary transmission channels h_{sr_k} and $h_{r_k d}$; interfering channels of relays and destination by primary source $h_{p_s r}$ and $h_{p_s d}$; interfering channel of eavesdropper by primary source $h_{p_s e}$; wiretap channels $h_{s e}$ and $h_{r e}$. We assume that channels within every group are independently and identically distributed fading, but channels for different groups may have different average gains, or we have $\lambda_{sr_k} = \lambda_{r_k d}$, $\lambda_{p_s r_k} = \lambda_{p_s d}$, $\lambda_{s e} = \lambda_{r_k e}$ for all k . We also assume that all channel state information (CSI) are available. The availability of the eavesdropper CSI is a commonly used assumption when eavesdropper is an active user of the system and not the intended destination for confidential messages [9].

As show in Fig. 1a, if the relay is selected to receive data from the secondary source, the received signal at relay and eavesdropper are respectively given by

$$\begin{aligned} y_{r_k} &= \sqrt{P_s} h_{sr_k} s + \sqrt{P_{p_s}} h_{p_s r_k} s' + n_{r_k}, \\ y_{e1} &= \sqrt{P_s} h_{s e} s + \sqrt{P_{p_s}} h_{p_s e} s' + n_e. \end{aligned} \quad (1)$$

Similarly, the received signal at destination and eavesdropper in Fig. 1b are given by

$$\begin{aligned} y_d &= \sqrt{P_{r_k}} h_{r_k d} s + \sqrt{P_{p_s}} h_{p_s d} s' + n_d, \\ y_{e2} &= \sqrt{P_{r_k}} h_{r_k e} s + \sqrt{P_{p_s}} h_{p_s e} s' + n_e, \end{aligned} \quad (2)$$

respectively where P_s , P_{p_s} , P_{r_k} are denoted as transmitting power of secondary source, primary source and relay, respectively; s , s' are transmission signal from secondary source and primary source, respectively. In underlay CRN, interfering power of primary destination must below a predefined threshold P_{th} so that we can get $P_s \gamma_{sp_d} < P_{th}$, $P_r \gamma_{rp_d} < P_{th}$.

Thus, in the source-to-relay transmission, the received SIR at the selected relay and eavesdropper is

$$SIR_{r_k} = \frac{P_s |h_{sr_k}|^2}{P_{p_s} |h_{p_s r_k}|^2} \quad SIR_{e1} = \frac{P_s |h_{se}|^2}{P_{p_s} |h_{p_s e}|^2}. \quad (3)$$

respectively. In the relay-to-destination transmission, the received SIR at secondary destination and eavesdropper is

$$SIR_d = \frac{P_{r_k} |h_{r_k d}|^2}{P_{p_s} |h_{p_s d}|^2} \quad SIR_{e2} = \frac{P_{r_k} |h_{r_k e}|^2}{P_{p_s} |h_{p_s e}|^2}. \quad (4)$$

In this scenario, interference power from the primary source is dominant relative to the noise so that the noise effects can be ignored and the capacity is mainly dependent on the SIR [10].

2.2 The Proposed Buffer-Aided Scheme

The target of the proposed buffer-aided scheme is to select the relay separately whether to receive or transmit packets, so that the selected relay transmission can get the best security performance and improve the diversity gains. As the definition [5], the SOP in this scenario is obtained as the probability that the selected link is in outage, i.e.

$$p_{sop} \triangleq \begin{cases} \mathbb{P} (C_{\Delta 1} < C_{th}) & R \text{ reception,} \\ \mathbb{P} (C_{\Delta 2} < C_{th}) & D \text{ reception.} \end{cases} \quad (5)$$

where

$$\begin{aligned} C_{\Delta 1} &= (1/2)\log_2 (1 + SIR_{r_k}) - (1/2)\log_2 (1 + SIR_{e1}) \\ C_{\Delta 2} &= (1/2)\log_2 (1 + SIR_d) - (1/2)\log_2 (1 + SIR_{e2}) \end{aligned} \quad (6)$$

where C_{th} is the target rate, and the factor 1/2 captures the fact that it takes two time slots to transmit any packet from the source to the destination.

According to above expression of the SOP, the proposed novel scheme select the best relay as follows:

$$R = \arg \max_{R_k} \left\{ \max_{R_k: \Psi(Q_k) \neq L} \left\{ \frac{SIR_{r_k}}{SIR_{e1}} \right\}, \max_{R_k: \Psi(Q_k) \neq 0} \left\{ \frac{SIR_d}{SIR_{e2}} \right\} \right\}. \quad (7)$$

Substituting (3), (4) into (7) and simplifying gives

$$R = \arg \max_{R_k} \left\{ \frac{\max_{R_k: \Psi(Q_k) \neq L} \left\{ \frac{|h_{sr_k}|^2}{|h_{p_s r_k}|^2} \right\}}{\frac{|h_{se}|^2}{|h_{p_s e}|^2}}, \frac{\max_{R_k: \Psi(Q_k) \neq 0} \left\{ \frac{|h_{r_k d}|^2}{|h_{r_k e}|^2} \right\}}{\frac{|h_{p_s d}|^2}{|h_{p_s e}|^2}} \right\}. \quad (8)$$

3 Secrecy Outage Probability Analysis

In this section, we derive the SOP of the proposed scheme and the conventional max-min scheme respectively.

3.1 The Proposed Buffer-Aided Scheme

First, we construct the state transition matrix of the Markov chain (MC).

In order to analyze the proposed relay selection scheme, we model the possible state transitions between the states as MC. Thus, due to K relays available which has a buffer of size L , there are $(L + 1)^K$ states totally and form a $(L + 1)^K \times (L + 1)^K$ state transition matrix of the MC as \mathbf{A} where $\mathbf{A}_{i,j} = \mathbb{P}(s_j \rightarrow s_i) = \mathbb{P}(X_{t+1} = s_i | X_t = s_j)$ is the transition probability to move from state s_j at time t to state s_i at time $t + 1$.

According to the number of data packets at every relay, namely Q_k , we can get one pair of $(\Phi_r(Q_k), \Phi_t(Q_k))$ which denotes the available links of relay for transmission and reception. So every state s_l corresponds to one pair of $(K_{1,l}, K_{2,l})$ where $K_{1,l} = \sum_{i=1}^K \Phi_t(Q_i)$ and $K_{2,l} = \sum_{i=1}^K \Phi_r(Q_i)$. Therefore, the target of our proposed scheme is to select best link among all $K_{1,l}$ source-to-relay links and $K_{2,l}$ relay-to-destination links. $p_{sop}^{s_l}$ is denoted as the SOP for every state s_l . Considering all possible $(L + 1)^K$ states, we derive the SOP as

$$P_{out} = \sum_{l=1}^{(L+1)^K} \pi_l \cdot p_{sop}^{s_l}, \quad (9)$$

where $\pi_l = p(s_l)$ is the stationary probability for state s_l which we obtain below. Due to all channel are independent and SIRs for all channel are independent, according to (8), the probability of selected relay link whether to receive data packet from secondary source or transmit data packet to secondary destination is equable as $\frac{1}{(K_{1,l} + K_{2,l})}$. Furthermore, the state remains unchanged if secrecy outage event occurs, the probabilities that state s_l move to a state in U_l is given by

$$\bar{p}_{out}^{s_l} = \frac{1 - p_{out}^{s_l}}{K_{1,l} + K_{2,l}}. \quad (10)$$

With these results, we have

$$\mathbf{A}_{i,j} = \begin{cases} p_{out}^{s_l}, & \text{if } i = j \\ \bar{p}_{out}^{s_l}, & \text{if } s_i \in U_l \\ 0, & \text{elsewhere} \end{cases}. \quad (11)$$

Because the transition matrix \mathbf{A} in (11) is column stochastic, irreducible and aperiodic [1], the stationary state probability vector is obtained as

$$\boldsymbol{\pi} = (\mathbf{A} - \mathbf{I} + \mathbf{B})^{-1} \mathbf{b}, \quad (12)$$

where $\boldsymbol{\pi} = [\pi_1, \pi_2, \dots, \pi_{(L+1)^K}]$, $\mathbf{b} = (1, 1, \dots, 1)^T$, \mathbf{I} is the identity matrix, and $\mathbf{B}_{i,j} = 1, \forall i, j$. Finally, substituting (11), (12) into (9), we can get

$$P_{out} = \sum_{l=1}^{(L+1)^K} \pi_l \cdot p_{out}^{s_l} = \text{diag}(\mathbf{A}) \cdot \boldsymbol{\pi} = \text{diag}(\mathbf{A}) \cdot (\mathbf{A} - \mathbf{I} + \mathbf{B})^{-1} \mathbf{b}. \quad (13)$$

Secondly, we derive the SOP of the proposed relay selection scheme.

In order to simplify the expression, combining $\gamma_{ij} = |h_{ij}|^2$ and $\lambda_{ij} = E|h_{ij}|^2$ as we mentioned before, let $m_k = \gamma_{sr_k}/\gamma_{p_s r_k}$, $m = \max\{m_k\}$, $n_k = \gamma_{r_k d}/\gamma_{r_k e}$, $n = \max\{n_k\}$, so we have $x = m/\frac{\gamma_{s e}}{\gamma_{p_s e}}$, $y = n/\frac{\gamma_{p_s d}}{\gamma_{p_s e}}$ and $z = \max\{x, y\}$. Concentrating on the secrecy capacity, we assume the channel SIR is high enough so that the decoding is always successfully at the relay and destination nodes as [2]. Thus, the probability of secrecy outage event at the high SIR is given by

$$P_{out}^{sI} = \mathbb{P}(z < 2^{2C_{th}}) = F_Z(z)|_{z=2^{2C_{th}}}, \quad (14)$$

where $F_Z(z)$ is the cumulative-probability-function(CDF) of z which is derived below.

On the one hand, γ_{sr_k} and $\gamma_{p_s r_k}$ are independent and both exponential distributions, the CDF of m is given by

$$F_M(m) = \left[\frac{m}{m + \alpha_1} \right]^{K_{1,l}}, \quad \alpha_1 = \frac{\lambda_{sr_k}}{\lambda_{p_s r_k}}, \quad (15)$$

The rest can be done in the same way. And we have

$$F_X(x) = \int_0^\infty \frac{\alpha_2}{(m + \alpha_2)^2} \cdot \left(\frac{xm}{\alpha_1 + xm} \right)^{K_{1,l}} dm, \quad F_Y(y) = \int_0^\infty \frac{\alpha_4}{(m + \alpha_4)^2} \cdot \left(\frac{ym}{\alpha_3 + ym} \right)^{K_{2,l}} dm, \quad (16)$$

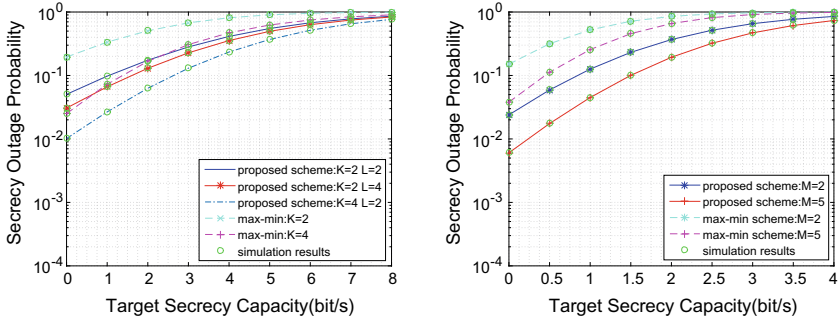
where $\alpha_1 = \frac{\lambda_{sr_k}}{\lambda_{p_s r_k}}$, $\alpha_2 = \frac{\lambda_{s e}}{\lambda_{p_s e}}$, $\alpha_3 = \frac{\lambda_{r_k d}}{\lambda_{r_k e}}$, $\alpha_4 = \frac{\lambda_{p_s d}}{\lambda_{p_s e}}$. Because x and y are mutually independent, we have

$$P_{out}^{sI} = F_Z(z)|_{z=2^{2C_{th}}} = F_X(2^{2C_{th}}) F_Y(2^{2C_{th}}). \quad (17)$$

4 Numerical Results

In this section, simulation results are given to verify the superiority of proposed scheme compared to conventional scheme in terms of the SOP. In the simulations below, we assume that transmission power of the primary transmitter is normalized to unit [2]. The buffer size $L = 4$ and the number of relays $K = 4$ are set as default. Simulation results are obtained by averaging 100000 independent runs.

In Fig. 2a, we plot the secrecy outage probability versus the target secrecy capacity with different schemes, different buffer size L , and different the number of relays K respectively. For comparison, we set $\lambda_{sr_k} = \lambda_{r_k d} = 30$ dB, $\lambda_{s e} = \lambda_{r_k e} = 6$ dB and $\lambda_{p_s r_k} = \lambda_{p_s d} = \lambda_{p_s e} = 10$ dB. First, we compare the performance of the proposed buffer-aided relay selection scheme with performance of conventional unbuffered max-min scheme where the number of relays is same. It is showed clearly that proposed scheme has better security performance than conventional scheme and the SOP is reduced significantly when the target secrecy capacity is lower than 2 whether $K = 2$ or $K = 4$. Second, with different combinations of buffer size and the number of relays, we find that



(a) The SOP of different combination of buffer size and the number of relays (b) The SOP of different average wiretap channel gains

Fig. 2. The SOP for different relay selection schemes

whether the increase of the number of relays or increase of the buffer size can decrease the SOP.

In Fig. 2b, we focus on different values of average channel gains, especially the wiretap channel average gains. We set $\lambda_{sr_k} = \lambda_{r_k d} = 30$ dB, $\lambda_{p_s r_k} = \lambda_{p_s d} = \lambda_{p_s e} = 10$ dB, but set $\lambda_{s e} = \lambda_{r_k e} = 6$ dB and $\lambda_{s e} = \lambda_{r_k e} = 15$ dB respectively. $M = \lambda_{sr_k} / \lambda_{s e}$ is denoted as average channel gains ratio. We can get the conclusion that the security performance will decrease with better wiretap channel state and proposed scheme have a lower SOP than max-min scheme in the same channel state.

5 Conclusion

This paper proposed a new buffer-aided relay selection scheme for underlay CRN in presence of an eavesdropper. We considered interference introduced by the primary transmitter and eavesdropping attack comprehensively. We analyzed the security performance in the underlay CRN with multiple buffer-aided relays for the first time. We used state transition matrix and stationary state vector of the MC to derive a new expression of the SOP. Both theoretical and simulations showed that the proposed scheme achieved a significant improvement of security performance than conventional scheme, making it an attractive way to realize secrecy transmission at the physical layer of the CRN.

Acknowledgements. This work was supported by National Natural Science Foundation of China (NSFC) (No. 61671075) and Major Program of National Natural Science Foundation of China (No. 61631003).

References

1. Krikidis, I., Charalambous, T., Thompson, J.S.: Buffer-aided relay selection for cooperative diversity systems without delay constraints. *IEEE Trans. Wirel. Commun.* **11**(5), 1957–1967 (2012)
2. Chen, G., Tian, Z., Gong, Y., Chen, Z., Chambers, J.A.: Max-ratio relay selection in secure buffer-aided cooperative wireless networks. *IEEE Trans. Inf. Forensics Secur.* **9**(4), 719–729 (2014)
3. Yang, P., Luo, L., Qin, J.: Outage performance of cognitive relay networks with interference from primary user. *IEEE Commun. Lett.* **16**(10), 1695–1698 (2012)
4. Darabi, M., Jamali, V., Maham, B., Schober, R.: Adaptive link selection for cognitive buffer-aided relay networks, *IEEE Commun. Lett.* **19**(4), 693–696 (2015)
5. Chen, G., Tian, Z., Gong, Y., Chambers, J.: Decode-and-forward buffer-aided relay selection in cognitive relay networks. *IEEE Trans. Veh. Technol.* **63**(9), 4723–4728 (2014)
6. Sun, A., Liang, T., Zhang, Y.: Secure performance analysis of buffer-aided cognitive relay networks under delay unconstraint case. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2164–2168, March 2016
7. Huang, J., Swindlehurst, A.L.: Buffer-aided relaying for two-hop secure communication. *IEEE Trans. Wirel. Commun.* **14**(1), 152–164 (2015)
8. Ding, X., Song, T., Zou, Y., Chen, X.: Relay selection for secrecy improvement in cognitive amplify-and-forward relay networks against multiple eavesdroppers. *IET Commun.* **10**(15), 2043–2053 (2016)
9. Duong, T.Q., Hoang, T.M., Kundu, C., Elkashlan, M., Nallanathan, A.: Optimal power allocation for multiuser secure communication in cooperative relaying networks. *IEEE Wirel. Commun. Lett.* **5**(5), 516–519 (2016)
10. Duong, T.Q., Yeoh, P.L., Bao, V.N.Q., Elkashlan, M., Yang, N.: Cognitive relay networks with multiple primary transceivers under spectrum-sharing. *IEEE Signal Process. Lett.* **19**(11), 741–744 (2012)



Generative Adversarial Network-Based Credit Card Fraud Detection

Xiaobo Xie^(✉), Jian Xiong, Liguo Lu, Guan Gui, Jie Yang,
Shangan Fan, and Haibo Li

College of Telecommunication and Information Engineering, Nanjing University
of Posts and Telecommunications, Nanjing 210000, China
18225905780@163.com, jxiong@njupt.edu.cn

Abstract. With the development of the financial industry, the number of credit cards has greatly increased. However, credit card fraud is still a major concern for financial security. Credit card fraud detection is a typical imbalanced classification problem, in which fraudulent cardholders are far less than non-fraudulent cardholders. The training on imbalanced samples will cause that the classifier ignores the minor fraudulent samples. To solve this problem, a generative adversarial network (GAN) based classification method is proposed for credit card fraud detection. GAN consists of a generative model and a discriminative model, and the two models are trained in a competitive way to get the Nash equilibrium. Specifically, the generative model tries to fit the real distribution of the non-fraudulent samples. The discriminative model determines the probability of a sample belongs to the distribution of non-fraudulent samples. To improve the discrimination performance, the fraudulent samples are also used in the training of discriminative model, which is different from the traditional GAN training scheme. The experimental results show that the recall reaches 82.7%.

Keywords: Fraud · Imbalanced · GAN · Discriminative model

1 Introduction

With the rapid advancement of machine learning, the demand for data is increasing. However, there are still many imbalanced datasets that are difficult to handle conventionally, which has attracted more and more attention in the academia and industry. For example, credit card fraud detection is a typical imbalanced classification problem. Using the existing methods to classify balanced data can generally get a good classification performance. However, the data in the real world is often imbalanced. If these methods are used to classify imbalanced data, the performance of the classifier will be degraded. Therefore, it is important to study the classification method for handling imbalanced datasets.

At present, the relevant methods of imbalanced data classification are mainly studied on the data preprocessing. Oversampling and undersampling are generally used. Oversampling is the most commonly used method for sampling imbalanced data. The basic idea is to eliminate data imbalance by increasing the size of minor class [1].

The undersampling method improves the classification performance of minor class by reducing the size of the major class [2]. The simplest method is to reduce the size of the major class by randomly removing some of the major class samples. The disadvantage is that some important information of major class will be lost, such that it cannot make full use of existing information [3–5].

In this work, a generative adversarial network (GAN) based classification method is proposed for imbalanced data classification. GAN consists of a generative model and a discriminative model, and the two models are trained in a competitive way to get the Nash equilibrium. The competition between the generative model and discriminative model leads to better discrimination. To improve the discrimination performance, fraud samples are added to the loss function of the discriminative model. The proposed classification method does not require oversampling or undersampling of the data samples. The proposed method is applied to the credit card fraud detection based on GAN. Experimental results show that the recall reaches 82.7%.

2 Generative Adversarial Networks

Generative Adversarial Network (GAN) consists of a generative model (generator) that try to generate samples from noise variable without detection and a discriminative model (discriminator) that learns to determine whether a sample is from the real data distribution or generative model. The generative model can be thought of as a counterfeit team, which tries to create a fake currency and uses it without detection. The discriminative model simulates the police team and tries to detect counterfeit currency.

Competition in this game drives the two teams to improve their methods until the counterfeit goods could not be distinguished from the genuine ones [6]. When the training is converged, the discriminator achieves Nash equilibrium, and the discriminator cannot determine whether the sample is from discriminator or real data samples. At the same time, the discriminator also achieves high detection capability. The structure of GAN is shown in Fig. 1. Noise through the generator consists of several neural networks generate samples, and the discriminator determines whether samples are generated samples or real samples through an output value between 0 and 1.

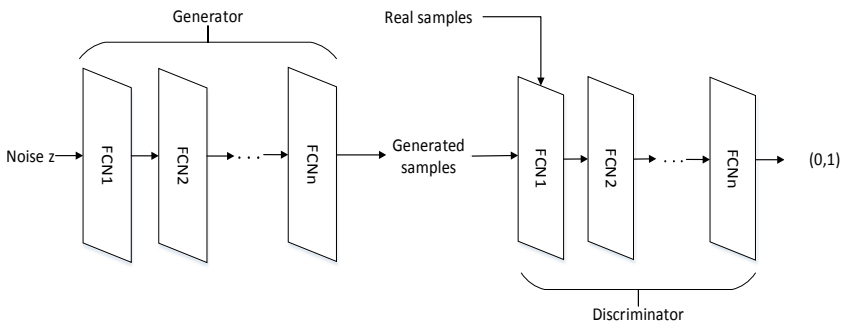


Fig. 1. Generative adversarial network

The fully connected neural networks are used in the generator and discriminator. The loss function is as follows:

$$L = \max_G E_{z \sim p_z} \log D(G(z)) \tag{1}$$

$$L = \max_D E_{x \sim p_r} \log D(x) + E_{z \sim p(z)} \log(1 - D(G(z))) \tag{2}$$

The symbol D denotes the discriminator, G denotes the generator, and p_z represents the distribution of Gaussian noise. The symbol p_r denotes the distribution of real samples.

Recently, many improved GANs have been proposed. Conditional Generative Adversarial Nets (CGAN) is designed by simply feeding the conditional information on both generator and discriminator [7]. Deep Convolutional Generative Adversarial Networks (DCGAN) perfectly applies the network structure of Convolutional Neural Network (CNN) to the framework of GAN [8]. Wasserstein Generative Adversarial Networks (WGAN) uses EM distance to represent the distance between real data distribution and generated data distribution. The loss function of WGAN as follows:

$$L = \max_G E_{z \sim p_z} D(G(z)) \tag{3}$$

$$L = \max_D E_{x \sim p_r} D(x) - E_{z \sim p(z)} D(G(z)) \tag{4}$$

WGAN completely solves the problem of unstable training in GAN, does not need to be careful about the balance in the training of generator and discriminator. WGAN basically solves the problem of mode collapse and ensures the diversity of generated samples [9, 10] (Fig. 2).

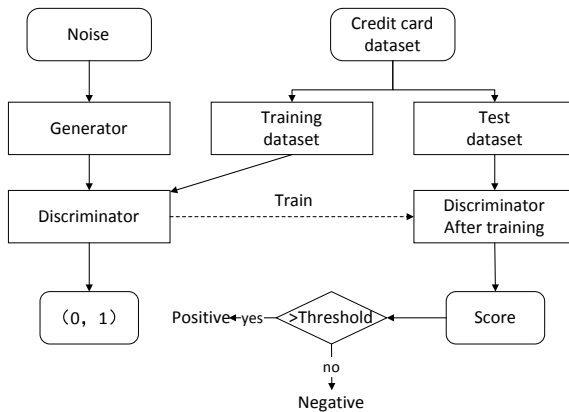


Fig. 2. GAN in classification

3 GAN-Based Imbalanced Classification

3.1 Advantages of GAN

Imbalanced dataset often consists of two parts which are extremely uneven in the number of samples. It makes it difficult to classify imbalanced dataset with common classifiers. Since undersampling may result in loss of data information, and oversampling may change the data distribution. A generator capable of fitting a majority of the sample distribution is needed, and a strong discriminator which can determine whether samples obey the distribution is needed. The proposed method uses GAN to classify the imbalanced dataset.

The deep neural networks have better fitting capabilities than many linear classifiers with strong robustness. The discriminator consists of several fully connected neural networks. The trained discriminator can well distinguish real data distribution from other data distribution. The structure of discriminator is shown in Fig. 3. The neural network consists of four layers of fully connected networks with a sample vector input. Each neural node consists of weight, bias, and activation function. The symbol w means weight, b means bias, and AF means activation function. Similarly, the generator consists of several fully connected neural networks. When the training of generator is converged, the generator can well fit the real data distribution (Fig. 3).

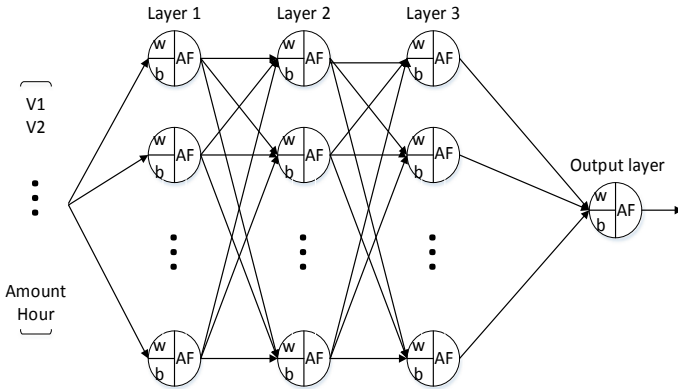


Fig. 3. Discriminator structure

During the training of generative adversarial network, generator and discriminator compete with each other and are trained alternately. After the discriminator training is converged, it can better distinguish between real non-fraud samples and other samples, including generated samples and fraud samples. And generator can generate samples which cannot be distinguished from real non-fraud samples, so that the discriminator can accurately determine whether the samples obey the distribution of real non-fraud samples. Samples which do not obey the distribution the non-fraud samples will be identified as a fraud by the discriminator.

The generator has a strong fitting capability, and the discriminator has a strong discrimination capability. The generator fits the distribution of non-fraud samples, and the discriminator determines whether samples are real non-fraud samples. The proposed method uses the advantages of generator and discriminator and the competition between generator and discriminator to get a better performance of classification on imbalanced dataset.

3.2 Modification in Discriminator

The discriminator may not be sensitive enough to fraud samples. The loss function of discriminator is modified as follows:

$$L = \max_D E_{x \sim p_r} D(x) - E_{z \sim p(z)} D(G(z)) + W * E_{x' \sim p_r'} D(x') \quad (5)$$

The symbol D represents discriminator, the symbol G represents generator, p_r represents the negative samples data distribution, z represents the input noise variable, p_r' represents the positive samples data distribution, and W denotes the weight of the discriminator output while input positive samples. The discriminator is used to detect the positive samples, so we assign W a large value, such as 10. When W takes 0, the discriminator performs well in distinguishing real non-fraud samples and generated non-fraud samples, but the discriminator is expected to determine whether samples are fraud or non-fraud. So, we add the fraud samples to the loss function of discriminator. Trained discriminator will be able to determine whether samples are positive or negative. Samples which does not obey the distribution of non-fraud samples and like fraud samples distribution will be identified as a fraud by the discriminator.

3.3 Threshold Selection

Different thresholds cause different recall and precision. Adjust different thresholds according to the actual situation to get the corresponding recall and precision. The threshold selection method is based on Bayesian algorithm. When C is minimum, we get the optimal threshold. As the formula (6) shows, the symbol C denotes the cost. Different thresholds are chosen to get the minimum C .

$$C = -\lambda * recall - precision \quad (6)$$

λ is a value to balance the recall and precision. Due to credit card fraud issues, we would detect more fraud samples with reasonable precision and would set λ to a large value.

4 Experimental Result

In this paper, the generative adversarial network is used to classify the credit card fraud dataset. Experiments are conducted using WGAN. The dataset contains data submitted by European cardholders using credit cards in September 2013. It contains only the

digital variables that are the result of the PCA conversion which is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The only features that are not converted with PCA are “Time” and “Amount”. The feature “Time” represents the number of seconds between each transaction and the first transaction. And the feature “Amount” is the transaction amount. The feature “Class” is the response variable, which takes a value of 1 if a credit card fraud occurs or 0 otherwise, and fraud samples are defined as positive samples. The dataset has 284807 samples with 30 features, and the proportion of positive and negative samples is 0.173 and 99.827%. Dataset is divided into training set and test set, account for 80 and 20%.

The structure of discriminators is shown in Table 1. The loss function of the discriminator is shown in (5). We add batch normalization to generator to avoid overfitting. For credit card fraud detection, we need high recall rather precision. λ is assigned to 5 in formula (6).

Table 1. Discriminator parameters

Layer	Size	Activation function	Batch normalization
Input	none,30	–	–
Layer1	100,100	ReLu	False
Layer2	100,100	ReLu	False
Layer3	100,100	ReLu	False
Output	100,1	–	False

Figure 4 shows the GAN-based classification method confusion matrix. The confusion matrix shows that the recall reaches 82.7%. The number of negative samples is 56883, and 215 samples of them are classified to positive. It means that there only less than 300 samples contained in more than 50,000 samples are checked, and 82.7% of credit card fraud samples can be detected.

Figure 5 shows the classification result when we set the weight of fraud samples W to 0, and it shows the recall reaches 70.67% with a low precision. It indicates that adding fraud samples in the training of discriminator can improve the classification performance. By setting different values to W , we found that setting W to 10 is the best choice.

Feature engineering of data before the experiment is implemented. To keep the features “Time”, “Amount” and other features consistent in magnitude, the feature “Time” and “Amount” are normalized. Through the random forest classifier, features V1, V2, V3, V4, V5, V6, V7, V9, V10, V11, V12, V14, V16, V17, V18, and V19 are more important than other features. Other features are removed to avoid overfitting. Figure 6 shows the GAN-based confusion matrix with feature engineering. The result shows that the recall reaches 77.3%. The feature engineering proved to be effective, but the performance of classifier with feature engineering is worse than the classifier without feature engineering. This shows that the classifier does not require feature engineering of data.

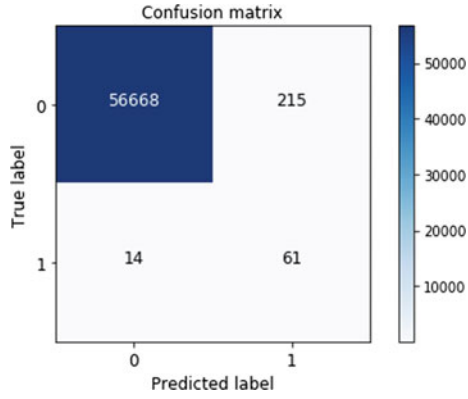


Fig. 4. GAN-based confusion matrix

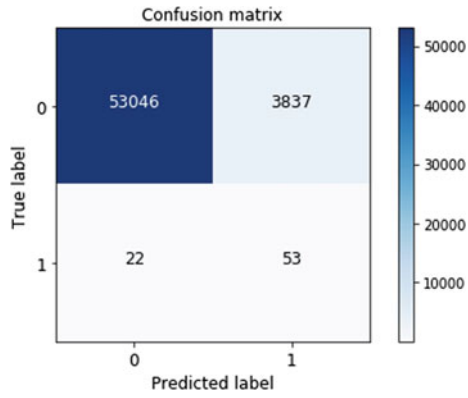


Fig. 5. GAN-based confusion matrix ($W = 0$)

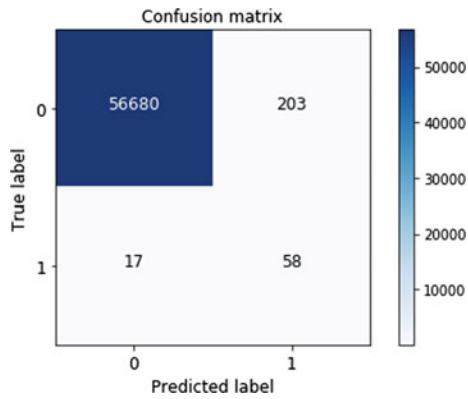


Fig. 6. Confusion matrix with feature engineering

5 Conclusion and Future Work

In this paper, a GAN-based classification method for credit card fraud is proposed. The experiments using modified GAN show that 82.7% of fraud samples can be detected with acceptable precision. The classification method has advantages; it does not require feature engineering and oversampling or undersampling.

This paper demonstrates the effectiveness of the generative adversarial network in the classification of imbalanced datasets, which will perform better through the optimization of the network structure and the fine-tuning of parameters.

Acknowledgements. This work was partially supported by National Natural Science Foundation of China Grants (No. 61701258), Natural Science Foundation of Jiangsu Province Grant (No. BK20170906), Natural Science Foundation of Jiangsu Higher Education Institutions Grant (No. 17KJB510044), Jiangsu Specially Appointed Professor Grant (RK002STP16001), and Innovation and Entrepreneurship of Jiangsu High-level Talent Grant (CZ0010617002).

References

1. Maloof, M.A.: Learning when data sets are imbalanced and when costs are unequal and unknown. In: ICML-2003 Workshop on Learning from Imbalanced Data Sets II (2003)
2. Chawla, N.V., Bowyer, K.W., Hall, L.O., et al.: SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**(1), 321–357 (2002)
3. Batista, G.E.A.P.A., Prati, R.C., Monard, M.C.: A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explorations Newsl.* **6**(1), 20–29 (2004)
4. Hanif, A., Azhar, N.: Resolving class imbalance and feature selection in customer churn dataset. In: International Conference on Frontiers of Information Technology. IEEE, pp. 82–86 (2018)
5. Gaffer, S.M., Yahia, M.E., Ragab, K.: Genetic fuzzy system for intrusion detection: Analysis of improving of multiclass classification accuracy using KDDCup-99 imbalance dataset. In: International Conference on Hybrid Intelligent Systems. IEEE, pp. 318–323 (2013)
6. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., et al.: Generative adversarial nets. In: International Conference on Neural Information Processing Systems. MIT Press, pp. 2672–2680 (2014)
7. Mirza, M., Osindero, S.: Conditional generative adversarial nets. *Comput. Sci.* 2672–2680 (2014)
8. Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. *Comput. Sci.* (2015)
9. Arjovsky, M., Bottou, L.: Towards principled methods for training generative adversarial networks (2017)
10. Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein GAN (2017)



A Novel Improved System Based on CVSS

Wen Tao¹(✉), Zhang Yuqing^{2,3}, Zhang Bo¹, Lei Jing¹,
Yang Yunxiang¹, and Guo Jing¹

¹ China Academy of Electronics and Information Technology, Beijing 100041,
China

{wt_bj_2016, edchang, yyxsdu}@126.com,
leijanet@163.com, 572179714@qq.com

² State Key Laboratory of Integrated Services Networks, Xidian University,
Xi'an 710071, China

zhangyq@nipc.org.cn

³ National Computer Network Intrusion Protection Center, University of
Chinese Academy of Sciences, Beijing 101408, China

Abstract. CVSS is one of the most representative quantitative assessment algorithms for security vulnerabilities. The calculated values of vulnerability harm according to the given index and formula. However, there is still insufficient objectivity of CVSS. Expert System is a kind of expert system for vulnerabilities; the severity of vulnerability can be assessed by experts based on the Expert System. The objectivity of Expert System and CVSS is analyzed, and CVSS is revised based on Expert System. The relationship between the above systems and the CWE, and between above systems and Product is analyzed, respectively, in the process. A new way of using Expert System is put forward based on CWE, namely, CWE Cycle Sorting Algorithm, in order to sort the average of vulnerability severity. Meanwhile, CWE Sort Factor is put forward based on CWE Cycle Sorting Algorithm to modify CVSS. The result is closer to the Expert System in terms of objectivity.

Keywords: CVSS vulnerability · Vulnerability assessment · Expert System · CSA · CSF

1 Introduction

Common Vulnerability Scoring System [1] (CVSS) as a quantitative vulnerability severity assessment system officially released version 1.0 in 2005, version 2.0 was released in 2007 and applied to U.S. National Vulnerability Database [2] (NVD), and version 3.0 was released in 2012. CVSS is a representative quantitative assessment system at present, most of the study of the vulnerability quantitative assessment system is based on the improvement of CVSS [3].

In the severity calculation process of CVSS, CVSS six key metrics of CVSS must be getting first, and then put into the formula to calculate the severity of CVSS, namely, the CVSS Base Score. Quantitative severity assessment system represented by the CVSS has the following advantages:

- **Global Importance:** It is a good global degree in accordance with the unified standard to tens of thousands of vulnerabilities in NVD.
- **Metrics Objectivity:** There are three options in each metric, a total of 18 options of six metrics. Under such a mechanism, different person score difference will be minimized, and repeatability is better.

But the objectivity of CVSS remains to be further improved; in the CVSS, there still exist the following problems:

- **Middle Complexity:** Although each metric can make more accurate, the final severity base score combined by six metrics is hard to do accurate and objective. That is the formula used to integrate with six metrics still not to objectively reflect the harm of vulnerabilities, compared with manual analysis.
- **Metric Subjective Degrees:** The options of {C: the destruction degree of confidentiality; I: the destruction degree of integrity; A: the destruction degree of availability} contains {Completely, Partly, and None}. When options are all {Partly}, the severity of different vulnerability categories is different. But the influence of the CWE cannot be reflected by CVSS.

Expert System scores the severity of vulnerability directly by humans. Experts are invited by an organization (including the workers on security vulnerabilities and technical, academic researchers, etc.); each expert is assigned a number of vulnerabilities to evaluate the severity of these vulnerabilities directly. Holm [4] invited 304 experts to assess 3040 NVD vulnerabilities, each expert was assigned 10 vulnerabilities, and the base score range of these vulnerabilities is in [0.0, 10.0], the same as the CVSS range. Expert System has the following advantages:

- **Authority:** The experts are all the authority of the information security field.
- **Thinking Depth and Score Contrast:** Experts for the 10 vulnerabilities, must be after careful scrutiny, so the severity order of the 10 vulnerabilities is very high credibility. That is although the absolute severity of each vulnerability may have a deviation, the relative severity is reliable.
- **Middle Process Complexity:** For the process of vulnerability scoring by Expert System is directly and without intermediate process.

But because of low efficiency, Expert System is always only used to verify and design quantitative severity system. At the same time, there are some other defects in Expert System, direct usage of Expert System to obtain the base score is not an appropriate way, and it is needed to design a more effective application way according to the characteristics of the Expert System.

Studies found that most vulnerabilities have similarities on some attributes; it is needed to analysis existing vulnerabilities from multiple perspectives, extract the commonness, and classify them. CWE standard summarized the advantages and disadvantages of the existing vulnerability classification scheme, enumerates 1500 vulnerabilities with different types in the real world, classify them using a conceptual framework of the detail. According to own characteristics, NVD further extracts 19 categories of vulnerabilities as the classification criteria of NVD. In some studies, vulnerability classification is direct as a metric of quantitative assessment system, and

do not study the principle of how vulnerability classification influences the severity. In other studies, vulnerability classification is estimated according to the metrics of CVSS.

In this paper, the objectivity of CVSS and Expert System is researched. The influence of vulnerability classification and the software product version to vulnerability severity are mainly studied. Based on Expert System, Revised CVSS is put forward so as to further enhance the objectivity of CVSS. Figure 1 shows the structure of research, first analyzes the fields which can affect the severity, respectively, on CVSS and Expert System, fields mainly contain CWE [5] and Product. Then a new Expert System application solution is put forward and finally applied to revise CVSS.

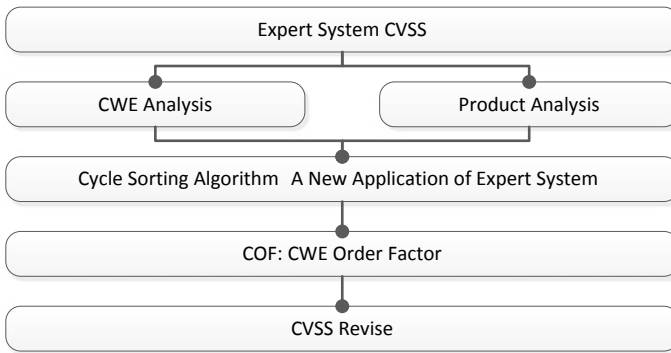


Fig. 1. The structure of study point

Need to note that, in this paper, all the algorithms are based on a premise, on Expert System, for each expert to assess the vulnerabilities, the ordering of severity is very objective. For example, experts assessed vulnerabilities $\{V1, V2, V3\}$, severity values are $\{2.0, 4.0, 6.0\}$, respectively, then the severity order of vulnerability is $V3 > V2 > V1$, this order is very credible. This is because that when an expert assesses vulnerabilities, he must be after repeated deliberation, even if there may be a deviation in terms of absolute severity, But the deviation of the relative severity of these vulnerabilities is indeed very small. Therefore, the idea of this paper is to use this premise and reflect on the assessment of CVSS.

Contributions:

- The influence of CWE to vulnerability severity is analyzed in CVSS and Expert System. Results show that the CWE reflect is more obvious in CVSS than that of Expert System.
- The influence of Product to vulnerability severity is analyzed in CVSS and Expert System. Results show that there is no direct relationship between the severity of the average product value and the usable range of these products that is a wild used product has the same severity with a less used product. it is proved that severity not affected by the Product Field in both CVSS and Expert System.

- A new way of application of Expert System is proposed, namely Cycle Sorting Algorithm, which can obtain a more objective severity sort of CWE class on the basis of the Expert System.
- CWE Sort Factor (CSF) is introduced into CVSS based on Cycle Sorting Algorithm. Using CSF, metrics of CVSS (C, I, A) is corrected, and the revised CVSS severity base score can more objectively reflect the CWE classification features.

2 Related Works

CWE analyzed the existing security vulnerability classification standards and summarized the advantages of various standards; meanwhile, hundreds of vulnerability types were enumerated. NVD extracted 19 types of the most commonly as NVD included classification standard. Currently, CWE has been recognized as the most authority and influence classification standard, the contained categories are considered to be more objective and reasonable.

Holm [4] invited 304 experts to assess 3040 vulnerabilities, each expert assessed 10 vulnerabilities, and the range of vulnerability severity is [0.0, 10.0], the same as the CVSS Base Score. In this paper, the characteristics of CVSS and Expert System are analyzed, as well as the relationship between the CWE and above system, and a revised CVSS system is put forward base on CWE and Expert System.

3 The Objectivity of Expert System and CVSS

In this section, the objectivity of Expert System and CVSS is studied, the factor that could affect the objectivity mainly include CWE and Product. The relationship between CVSS and CWE is studied first, and a question that if CWE classification features can be reflected in CVSS is also studied; meanwhile, the advantages and disadvantages of CVSS and Expert System were analyzed.

3.1 The Influence of Assessment System from CWE

The CVSS and Expert System can produce a classification translation for each CWE; each CWE classification has a certain deviation, comparing to the total average, that CWE classification can influence vulnerability severity score. The same CWE category has a different deviation in Expert System and CVSS, that is to say, Expert System and CVSS are inconsistent. CWE characteristics of the CVSS are more obvious than Expert System. There is a smaller variance and more obvious average deviation in CVSS.

3.2 The Influence of Assessment System from Product

The purpose of this section is to study whether the CVSS and Expert System is affected by the software Vendor and Product, and whether the difference between CVSS and Expert System is associated with the using range of Vendor and Product. We want to make sure that whether experts consider the Vendor and Product affected by the

vulnerability, when they assess vulnerabilities. If the products affected by vulnerability are more widely used and important, the severity of the vulnerability is serious. We can draw the conclusion that the severity score of CVSS and Expert System is not affected by the Vendor and Product.

4 Architecture Designs of Revised CVSS

4.1 Cycle Sort Algorithm

We believe that for the subjectivity of Expert System, Expert System should not be used directly for vulnerability severity assessment. But the order of vulnerability severity is after the expert's careful scrutiny and is objective. So Cycle Sort Algorithm is put forward based on this premise in order to use the results of the Expert System more objectively (see Table 1).

Table 1. Expert marked average

Category	CVE-ID	Expert marked severity	Average
CWE-119	CVE-2009-2832	6.9	7.95
	CVE-2009-2203	9.0	
CWE-189	CVE-2011-1659	4.8	4.8
CWE-20	CVE-2012-0151	9.0	9.0
CWE-200	CVE-2011-2800	6.0	6.0
CWE-310	CVE-2009-4144	7.2	7.2
CWE-399	CVE-2007-5001	4.0	4.0
CWE-94	CVE-2006-7102	6.5	6.5

Calculation process of winning percentage: First, calculate the average score of each category and each expert, respectively; calculate the victory number and the failure number of vulnerabilities assessed by each expert; and calculate the Difference of Victory Rate (DVR) of each category.

4.2 CWE Score Factor

Known in the above discussion, the relative severity is objective and the absolute severity is subjective on Expert System. Instead, the relative severity is not accurate enough and the absolute severity is relatively accurate on CVSS. Meanwhile, the three metrics of CVSS (C, I, A) is valued subjectively actually. C, I, A have three options {None, Partial, Completely}, however, even {Partial, Completely}; actually, the influence is different from C, I, and A (see Table 2).

C, I, and A are revised according to CWE sort ordered by Expert System, such as CWE-200, when C, I, and A are marked as Partial, the values of C and I should be multiply by 0.8; when C, I, and A are marked as Completely, the values of C and I also should be multiply by 0.8 (see Table 3).

Table 2 CVSS severity values

CVE		C	I	A	Base score
CVE-2009-0499	Avg	None	Partial	Partial	6.4
	Num	0	0.275	0.275	
CVE-2007-5213	Avg	Computer	Computer	Computer	9.3
	Num	0.66	0.66	0.66	

Table 3 Revised CVSS severity values

CVE		C	I	A	BaseScore
CVE-2009-0499	Avg	None	Partial	Partial	5.7
	Num	0	$0.275 * 0.75$	$0.275 * 0.75$	
CVE-2007-5213	Avg	Computer	Computer	Computer	8.5
	Num	0.66	$0.66 * 0.7$	$0.66 * 0.7$	

5 Conclusion

With the increasing number of security vulnerabilities, the severity of a vulnerability assessment is becoming more and more important. Currently, CVSS is the most representative security vulnerabilities quantitative assessment algorithm, which can compute the severity according to metrics and formulas. Expert System is a kind of expert assessment system for vulnerabilities, namely, the severity of vulnerability is scored directly by security experts. CWE is currently the most representative vulnerability classification systems. CVSS obtained wide attention and discuss the industry, which mainly focused on the discussion of the objectivity and dispersion of severity score. This paper mainly attempts to discuss and revise the objectivity of CVSS based on Expert System.

The relationship among CVSS, Expert System, and CWE was analyzed. It is more obvious CWE character reflected in CVSS than in Expert System. The influence Product was analyzed in CVSS and Expert System. The results show that it is not true that Product is more wildly, the severity is greater.

Based on the advantages and disadvantages of CVSS and Expert System, a conclusion is that the Expert System for CWE severity ordered more reasonable. Based on the above conclusions, Cycle Sort Algorithm is put forward. A new application for Expert System is put forward, and CWE Sort Factor (CSF) is introduced based on Cycle Sort Algorithm to revise CVSS. Metrics C, I, and A can be revised using CSF. The CWE classification features can be reflected objectively in the revised CVSS.

The next works will focus on two aspects: the relationship between metrics and CWE in CVSS; find an objective standard to test the objectivity of the quantitative severity assessment system.

Acknowledgements. This paper is supported by Beijing NOVA Program (Z181100006218041); the National Key R&D Program of China (2016YFB0800700), the National Natural Science Foundation of China (61572460, 61272481), the Open Project Program of the State Key Laboratory of Information Security (2017-ZD-01), the National Information Security Special Projects of National Development and Reform Commission of China [(2012) 1424], and China 111 Project (B16037).

References

1. CVSS, Common Vulnerability Scoring System. <http://nvd.nist.gov/cvss.cfm>[EB/OL]. (2014-08-01) [2015-08-01]
2. NVD, National Vulnerability Database. <http://nvd.nist.gov>[EB/OL]. (2014-08-01) [2015-08-01]
3. Wang, L.Y., Jajodia, S., Singhal, A., et al.: K-zero day safety a network security metric for measuring the risk of unknown vulnerabilities. *IEEE Trans. Dependable Secure Comput.* **11** (1), 30–44 (2013)
4. Holm, H., Ekstedt, M., Andersson, D.: Empirical analysis of system-level vulnerability metrics through actual attacks. *IEEE Trans. Dependable Secure Comput.* **9**(6), 825–837 (2012)
5. CWE, Common Weakness Enumeration. <http://cwe.mitre.org/>[EB/OL]. 2015-08-01



Efficient Traffic Coordination Strategies at Intersections Using Multiple Collision Sets

Yangan Mo^(✉), Mengqi Wang, Tingting Zhang, and Hongguang Xu

Shenzhen Graduate School, Communication Engineering Research Center,
Harbin Institute of Technology, Shenzhen, People's Republic of China
mo_yangan@163.com, zhangtt@hit.edu.cn

Abstract. Appropriate traffic coordination at intersections where multiple roads merge plays an important role in modern intelligent transportations systems. In this paper, we try to propose an efficient traffic coordination framework using multiple collision sets. Aiming at the essentially non-convex problem, we try to reformulate the original problem into a mixed binary integer quadratic programming one by proper relaxations. Low complexity solutions are also given afterwards. Numeric results show that the traffic throughput at intersections can be significantly improved compared to the existing investigations.

Keywords: Intelligent transportations · Intersection coordination · Traffic optimization · Mixed binary integer programming

1 Introduction

The intersection is a popular transportation topology structure where multiple roads merge. However, the current scheduling methods such as the traffic light and stop sign based methods usually lead to inefficient traffic congestions. Furthermore, most traffic accidents occur at intersections. Based on the investigation in EU and US, nearly 20% of fatal accidents take place around the intersections [1].

During the past few years, considerable new intersection coordination strategies have been proposed. Lin proposed a set-projection algorithm and a three-segment linear speed profile to control the trajectories of vehicles in the buffer area, resulting in a constant-speed driving in the core area [2]. Hafner focused on a two-vehicle collision avoidance at intersections and developed a decentralized control algorithm using V2V communications to decide whether automatic control is needed to prevent a collision [3]. Lioris acclaimed that the platoon of connected vehicles can double the throughput with the proof of three queueing models [4]. Kim proposed an optimal coordination framework based on MPC approach to generate the motion of each vehicle locally [5]. Campos focused on the decentralized solution to the local optimization problems by which low

complexity coordination schemes can be achieved [6]. Onieva presented a heuristic algorithm by which an autonomous vehicle can cooperate with a non-autonomous vehicle, and cross the intersections without any collisions [7].

In this paper, we first propose intersection coordination frameworks based on single collision set (CS) and multiple collision sets (CSs), respectively. We adopt the M condition [8] and mixed binary integer quadratic programming (MBIQP) [9] to achieve the optimal vehicle maneuvers. The simulation experiments state that multiple CSs can reduce approximate 80% of the coordination costs and 60% of the coordination time, comparing to single CS scenario, that leads to significantly higher traffic throughput and lower fuel consumption.

2 System Model

We consider a typical intersection with four roads, as seen in Fig. 1. The coordination system mainly consists of two parts, named buffer area (BA) and core area (CA) respectively. CA is the whole intersection in this model, while BA is the area where vehicles are waiting to enter the intersection. In this paper, we assume that a collision occurs when any two vehicles occupy the CS simultaneously. Through possible trajectories from vehicles on each road, we denote each overlapping part as a collision set (CS). The vehicle is considered as a circle, whose radius is indicated by the location uncertainty ρ . Suppose that there is only one vehicle in the BA each road, and our goal is to find the optimal maneuver for each vehicle to minimize the coordination cost.

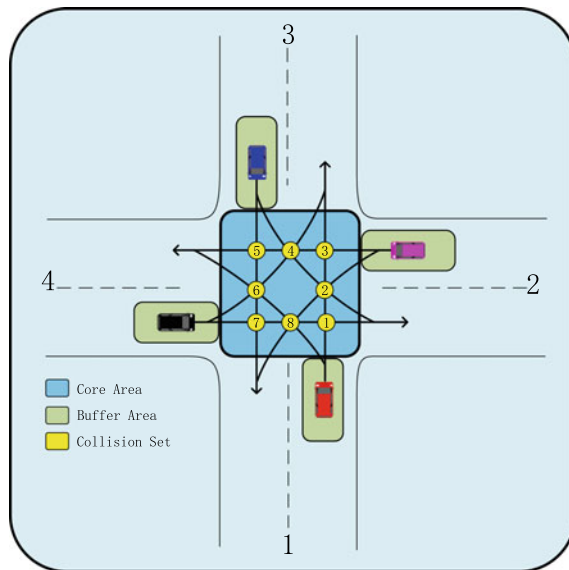


Fig. 1. Illustration of coordination system.

As shown in Fig. 1, the state parameters of i th vehicle can be described by $x_i(t) = [p_i(t), v_i(t)]^T$, where $p_i(t)$ denotes the current position of the vehicle, and $v_i(t)$ denotes the current velocity of the vehicle. Therefore, the motion model of each vehicle can be considered as:

$$x_i(t+1) = Ax_i(t) + Bu_i(t), i \in \{1, 2, 3, 4\} \quad (1)$$

where $t \in \{1, 2, \dots, T\}$ denotes discrete time interval after being sampled at the period of T_s , $u_i(t)$ denotes the current input control on the vehicle, in other words, acceleration and deceleration in our model. The speed and acceleration of vehicles are required as followed, due to practical considerations.

$$v_{i\min} \leq v_i(t) \leq v_{i\max} \quad (2)$$

$$u_{i\min} \leq u_i(t) \leq u_{i\max} \quad (3)$$

We define $v_{i\min} \geq 0$ in (2), which means that backing up is not allowable for vehicles. Additionally, the vehicles are not permitted to stop in the BA, preventing the deadlock problem.

3 Coordination Framework

3.1 Single Collision Set

As shown in Fig. 1, for single CS, we consider the entire CA as one single collision set. As a result, two vehicles can not show up in the CA simultaneously. Suppose L_i and H_i denote the position where vehicle in road i enters and leaves the CA respectively. For simplicity, the intersection is defined equally large for all vehicles [10], so that $H_i - L_i = H_j - L_j, \forall i, j \in \{1, 2, 3, 4\}$. The collision-free constraint is thus given as:

$$[p_i(t), p_j(t)] \notin \varepsilon_i \times \varepsilon_j, \forall t, \forall i \neq j \quad (4)$$

There exists a desired vehicle velocity by which a minimum fuel consumption can be achieved [8]. Thus, the objective function or cost function can be described by

$$A_i(x_i(t), u_i(t)) = (v_{d,i} - v_i(t))^2 Q_i + u_i^2(t-1) R_i \quad (5)$$

where $v_{d,i}$ denotes the desired velocity we define in road i , Q_i and R_i denote the weight of velocity offset and the weight of control respectively. In summary, based on the proposed model, the coordination optimization problem can be described as

$$\begin{aligned} \mathcal{P}_1 : \min. & \sum_{i=1}^4 \sum_{t=1}^T A_i(x_i(t), u_i(t)) \\ \text{s.t.} & (1)-(4) \end{aligned} \quad (6)$$

With the constraint of Eq. (4), only one vehicle can occupy the CS at any time slot. However, the feasible set of constraint (4) is non-convex, which makes the optimization problem NP-hard.

Only one vehicle can enter the intersection in each time slot, one can imagine that the coordination cost and road efficiency are not ideal under such circumstances, just like the simulation results later in this paper. In order to improve efficiency, we propose a coordination strategy based on multiple collision sets.

3.2 Multiple Collision Sets

As shown in Fig. 1, we assume that vehicles make left-turn movement in a manner of curved turn, so that we can get eight CSs according to their trajectories [11]. Table 1 presents the mapping between possible collisions with different CSs. For instance, vehicle1 and vehicle4 may collide in CS ①. Hence, all we have to do is make sure that vehicle1 and vehicle4 do not show up in CS ① simultaneously.

Table 1. Collision sets of four vehicles in eight collision sets

	CS ①	CS ②	CS ③	CS ④	CS ⑤	CS ⑥	CS ⑦	CS ⑧
Possible collisions	Vehicle1	Vehicle1	Vehicle1	Vehicle2	Vehicle2	Vehicle1	Vehicle3	Vehicle1
	Vehicle4	Vehicle2	Vehicle2	Vehicle3	Vehicle3	Vehicle3	Vehicle4	Vehicle2
		Vehicle3		Vehicle4		Vehicle4		Vehicle4

Since the collision-free constraint (4) is a non-convex constraint, the key to solving this optimization problem lies in how to satisfy the Eq. (4). Using the M condition method, the collision-free constraints (4) are implemented by introducing auxiliary binary decision variables [8], $\gamma_i, \delta_i \in \{0, 1\}, \forall i \in \{1, 2, 3, 4\}$ and requiring that

$$H_i - \gamma_i \times M \leq p_i \leq L_i + \delta_i \times M \tag{7}$$

$$\delta_i + \delta_j + \gamma_i + \gamma_j \leq 3 \tag{8}$$

where M is a sufficiently large constant, and p_i denotes the current position of vehicle i . The mapping between binary decision variables and position of vehicle is shown in the following table.

Table 2. The mapping between binary decision variables and position of vehicle

Binary decision variables	Position of vehicle
$\delta = 1, \gamma = 1$	In the CS
$\delta = 1, \gamma = 0$	Already passed the CS
$\delta = 0, \gamma = 1$	Not yet passed the CS

As shown in Table 2, when the binary decision variables δ_i , γ_i are both 1, vehicle i is in the CS. Any two vehicles cannot be in the same CS at the same time, constraint (8) thus should be satisfied. Therefore, the collision-free constraints transform from (4) to (7) and (8). In summary, the original problem thus becomes

$$\begin{aligned} \mathcal{P}_2 : \min. & \sum_{i=1}^4 \sum_{t=1}^T A_i(x_i(t), u_i(t)) \\ \text{s.t.} & (1)-(3), (7), (8) \end{aligned} \quad (9)$$

It should be noticed that this problem is a typical mixed binary integer quadratic programming problem (MBIQP), where the number of binary decision variables increases linearly with the number of CSs. In the eight CS, the coordination algorithm is shown in Algorithm 1.

Algorithm 1 Coordination algorithm at intersections in multiple collision sets

- 1: initialize the vehicles' parameters subject to (2) and (3)
 - 2: **for** each vehicle $i \in [1, 4]$ **do**
 - 3: **for** each sampling slot $j \in [1, T]$ **do**
 - 4: update the vehicles' parameters every slot according to (1)
 - 5: $H_{i,k} - \gamma_{i,j,k} \times M \leq x_{i,j}[1] \leq L_{i,k} + \delta_{i,j,k} \times M$
 - 6: where $L_{i,k}$ and $H_{i,k}$ denote the k^{th} collision set in road i , $k \in \{1, 2, 3\}$
 - 7: **suppose** $1 - 1 = 4$, $4 + 1 = 1$, $4 + 2 = 2$ in the following equations
 - 8: $d_1 = i - 1$, $d_2 = i + 1$, $d_3 = i + 2$
 - 9: $\delta_{i,j,1} + \gamma_{i,j,1} + \delta_{d_1,j,1} + \gamma_{d_1,j,1} \leq 3$, $\delta_{i,j,3} + \gamma_{i,j,3} + \delta_{d_3,j,3} + \gamma_{d_3,j,3} \leq 3$
 - 10: $\delta_{i,j,2} + \gamma_{i,j,2} + \delta_{d_2,j,2} + \gamma_{d_2,j,2} \leq 3$, $\delta_{i,j,2} + \gamma_{i,j,2} + \delta_{d_3,j,2} + \gamma_{d_3,j,2} \leq 3$
 - 11: **end for**
 - 12: **end for**
 - 13: solve the coordination optimization problem \mathcal{P}_2 , get the optimal trajectory for each vehicle, calculate the coordination cost and time
-

In the multiple CSs model, we specify the turning trajectories of vehicles in a curved manner and get the overlapping areas as eight CSs. The model should be more and more realistic while the number of collision sets gets larger and larger. However, the computational complexity increases exponentially with the increase of the collision sets.

4 Numerical Results

4.1 Simulation Settings

In simulation, for each vehicle, the initial data are given by Table 3 and the sampling period is specified as $T_s = 0.01$ (s).

Additionally, $P_0 = -110$ (m), $v_d = 40$ (m/s), $v_{\min} = 0.1$ (m/s), $v_{\max} = 80$ (m/s), where P_0 denotes the starting point, v_0 denotes initial velocity, v_d denotes

desired velocity, v_{\min} denotes minimum velocity, v_{\max} denotes maximum velocity, Q_i denotes velocity offset weight, R_i denotes control weight, u_{\min} denotes minimum control, u_{\max} denotes maximum control, $i \in \{1, 2, 3, 4\}$.

4.2 Coordination Realizations

As shown in Fig. 2, the trajectories of four vehicles passing the intersection are in the order of 2, 4, 1, 3. The colored triangles in the figure represent the collision sets for each road, for example, in road1, CS ①, CS ② and CS ③ are represented by green, gray and brown triangle respectively. The trajectories of four vehicles in multiple collision sets are more compact and tight than single collision set, indicating that the coordination time becomes shorter.

Table 3. Collision sets of four vehicles in eight collision sets

	Q_i	R_i	v_0 (m/s)	u_{\min} (m/s ²)	u_{\max} (m/s ²)
Min	0	0	30	-10	0
Max	10	10	40	0	10

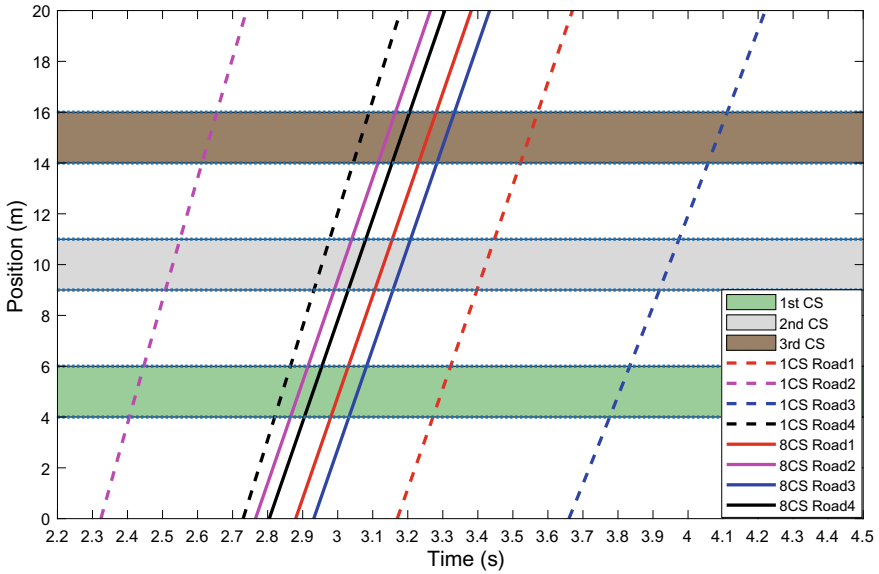


Fig. 2. Comparisons of four vehicle trajectories between 1CS and 8CSs.

As shown in Fig. 3a, in order to minimize the coordination cost, each vehicle will reach the desired velocity in the end. However, vehicles in 8CSs reach the desired velocity in a shorter interval, and the vehicle velocity in 8CSs are

smoother than that in single CS, which provides drivers and passengers a comfortable trip.

Figure 3b illustrates that in order to minimize the coordination cost, each vehicle will reach zero control in the end. However, vehicles in 8CSs reach zero in a shorter interval, and the control in 8CSs are smoother than that in single CS, which indicates that sudden acceleration and deceleration will not happen.

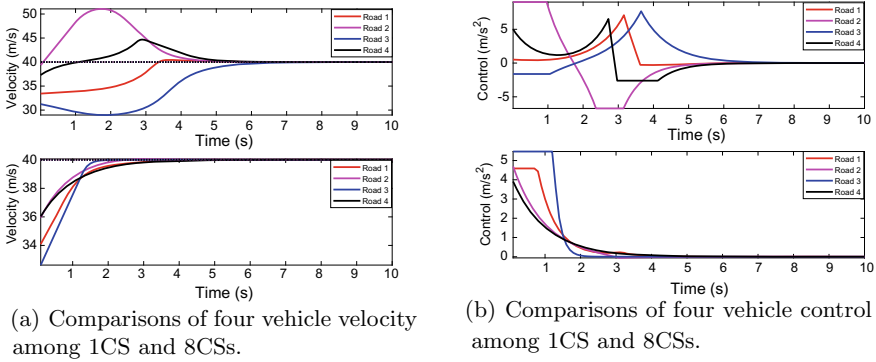


Fig. 3. Comparisons of four vehicle trajectories.

4.3 Statistical Analysis

As shown in Fig. 4a and Table 4, compared with single CS, the coordination cost of 8CSs decreases over 80%, indicating that collaborative vehicles can pass the intersections with fewer coordination, leading to a lower fuel consumption.

As illustrated in Fig. 4b and Table 4, compared with single CS, the coordination time of 8CSs decreases nearly 60%, indicating that vehicles can pass the intersection in a more efficient way.

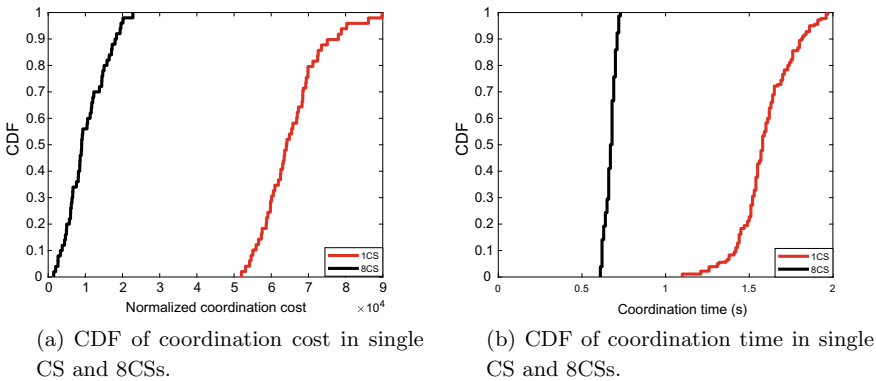


Fig. 4. Comparisons of coordination cost and coordination time.

Table 4. Average coordination cost and time in different scenarios

	1CS	8CSs
Normalized average coordination cost	1	0.156
Average coordination time(s)	1.651	0.676

5 Conclusion

A coordination system at intersections based on single and multiple collision sets is proposed in this paper. Using MBIQP, we transform a NP-hard constraint into a low complexity and easily solved problem. Numerical results show that: (i) In the single CS circumstance, the coordination system is inefficient cause any two vehicles can not be in the intersection at the same time. It leads to a large cost and a long interval. (ii) In eight CSs circumstance, vehicles in this coordination system take curved-turn in turn-making. The cost is smaller and the coordination time interval is shorter compared to single collision set. (iii) The velocity and control in eight CSs change smoothly while passing the intersection, resulting in a better driving experience and lower fuel consumption.

In the future work, we will consider the dynamics coordination framework, consisting of arriving vehicles and leaving vehicles all the time. Additionally, the accommodation of human-driven vehicles and automatic vehicles will be added in the coordination framework.

Acknowledgements. This paper was supported by the Natural Science Foundation of China (under Grant No. 91638204 and 61771159), Guangdong Natural Science Foundation under Grant No. 2017A030313392, Shenzhen Fundamental Research Project (under Grant No. JCYJ2017081115369780 and JCYJ20160318094236224).

References

1. 2010 Motor vehicle crashes: Overview. In: Traffic Safety Facts—Research Note (2012)
2. Lin, P., Liu, J., Jin, P.J.: Autonomous vehicle-intersection coordination method in a connected vehicle environment. *IEEE Trans. Intell. Transp. Syst. Mag.* pp. 37–47 (2017)
3. Hafner, M.R., Cunningham, D., Caminiti, L.: Cooperative collision avoidance at intersections: algorithms and experiments. *IEEE Trans. Intell. Transp. Syst.* 1162–1175 (2013)
4. Lioris, J., Pedarsani, R., Tascikaraoglu, F.Y.: Platoons of connected vehicles can double throughput in urban roads. *Transp. Res. Part C*, 292–305 (2017)
5. Kim, K.D., Kumar, P.R.: An MPC-based approach to provable system-wide safety and liveness of autonomous ground traffic. *Autom. Control IEEE Trans.* 3341–3356 (2014)
6. Campos, G.R., Falcone, P., Sjoberg, J.: Autonomous cooperative driving: a velocity-based negotiation approach for intersection crossing. In: International IEEE Conference on Intelligent Transportation Systems, pp. 1456–1461 (2013)

7. Onieva, E., Milanés, V., Villagrà, J.: Genetic optimization of a vehicle fuzzy decision system for intersections. *Expert Syst. Appl.* 13148–13157 (2012)
8. Hult, R., Campos, G.R., Falcone, P.: An approximate solution to the optimal coordination problem for autonomous vehicles at intersections. In: *American Control Conference*, pp. 763–768 (2015)
9. Bienstock, D.: Computational study of a family of mixed-integer quadratic programming problems. In: *International IPCO Conference on Integer Programming and Combinatorial Optimization*. Springer, pp. 80–94 (1995)
10. Hult, R., Campos, G.R., Steinmetz, E.: Coordination of cooperative autonomous vehicles: toward safer and more efficient road transportation. *IEEE Signal Process. Mag.* 74–84 (2016)
11. Kamal, M.A.S., Imura, J.I., Hayakawa, T.: A vehicle-intersection coordination scheme for smooth flows of traffic without using traffic lights. *IEEE Trans. Intell. Transp. Syst.* 1136–1147 (2015)



Key-Controlled PEG-LDPC Algorithm Design and Its Application in Secure Communication

Zhiping Shi¹(✉), Shujun Zhang¹, Fengcheng Lyu¹, Fan Bu¹,
Hongxia Sun², and Qian Zhang³

¹ National Key Lab of Science and Technology on Communications, UESTC,
Chengdu 611731, China

szp@uestc.edu.cn, zhang_shujun_cool@126.com

² Southwest Jiaotong University, Chengdu 610031, China

³ Public Security Fire Force College, Kunming 650208, China

Abstract. Because of the security of the symmetric cryptosystem, a secure communication scheme based on PEG-LDPC codes has been put forward in this paper. The secure and reliable integrated communication is based on hiding and changing generator matrix randomly and simultaneously under the premise of ensuring error correction ability of LDPC codes. The performance of LDPC codes constructed by PEG algorithm is excellent, and this algorithm can construct an LDPC code with arbitrary code length and code rate.

Keywords: Secure and reliable communication · LDPC code · PEG algorithm

1 Introduction

The problem of decoding general linear block codes in error correction codes is proved to be a difficult NPC problem in 1978 [1]. Scholars in the field of cryptography generally believe that quantum computers cannot overcome NPC difficulties, so these kinds of cryptographic systems based on error corrections codes can counter quantum attacks. Based on this theory, McEliece proposed a public key cryptosystem named M cryptosystem [2]. In wireless communication systems, because of the high speed and the ability of handling large quantities data of symmetric encryption system, scholars have proposed many schemes about symmetric encryption system. Xinmei proposed a crypto-error correction system combining public encryption and error correction—M_s public key system [3]. Rao proposed M symmetric cryptosystem [4]. Wenyan proposed a McEliece symmetric encryption scheme based on irregular LDPC codes [5].

In paper [6], the author presents a communication scheme which uses RDF algorithm to generate a large number of equal-performance LDPC check matrices and a linear congruential method to control synchronous generation based on LDPC code. For the excellent error-correcting performance and the flexibility of code length and code rate of LDPC codes constructed by PEG algorithm, in this paper, a K-PEG algorithm which is combined with the idea of secure communication to construct a mass-equivalent LDPC code using key sequence control is proposed.

2 The Secure and Reliable Communication Scheme

According to the proposed transmission scheme in paper [6], suppose that the parameters H, S, P should be known by both transmitter and receiver, where H is a $(n - k) \times n$ parity check matrix for an LDPC code, G is generator matrix corresponding to H whose size is $k \times n$, S is a $k \times k$ nonsingular matrix, and P is a $n \times n$ permutation matrix (i.e., $P^T = P^{-1}$). The scheme is as follows:

Transmitter:

1. Compute $c = m \cdot S \cdot G \cdot P$.
2. Transmit c through the noisy channel.

Receiver:

Received vector $r = c + n$, and n denotes the noise in the channel.

1. Compute $r' = rP^{-1} = rP^T = c \cdot P^T + n \cdot P^T = m \cdot S \cdot G + n'$.
2. Decode r' , and recover $u' = m \cdot S$ with the matrix G .
3. Apply S^{-1} to u' and obtain $m = u' \cdot S^{-1} = m \cdot S \cdot S^{-1}$.

The difference between this scheme and paper [7] is that matrixes S, P, H , and G are invariable in cryptosystems. These matrices are randomly changing by synchronous controlling. This scheme mainly discusses the equivalent matrices (parity check matrix H) designed by using PEG algorithm. The design of matrices based on PEG algorithm is better than using RDF in error correction performance.

3 Key-Controlled PEG Algorithm

The PEG (progress edge growth) algorithm is an effective method for constructing a short and medium-length LDPC code with excellent construction performance. The PEG algorithm increases the edges between symbol nodes and checks nodes one by one while keeping the ring length of the Tanner graph as large as possible [8]. Due to the high coding complexity of the original PEG algorithm, the literature [9] proposed an improved PEG algorithm, which reduces the coding complexity when construct the check matrix by using lower triangles.

3.1 The Generation of Key

Legitimate communication parties use TDD (Time Division Duplexing) signal transmission methods to detect the characteristics of wireless channels. Based on the feature of fast time-varying of wireless channels, both legitimate parties can obtain a random key with "noise" properties at the same time. Due to the reciprocity of the wireless channel, the characteristics of the channels detected by both parties of the communication will show a high degree of correlation and then both parties use quantitative negotiation and other means to extract a completely consistent key sequence.

Due to the presence of fast vacancies in the wireless channel, the legitimate communication parties can easily keep more than half a wavelength away from

potential attackers to ensure the physical security of key information when they detect the characters of channel information. Therefore, the legitimate communication parties can use the characteristics of the wireless channel to convert the wireless channel characteristic information into an encryption key. The encryption key used in this paper is obtained through the wireless channel characteristic information.

3.2 The Description of K-PEG Algorithm

In the K-PEG algorithm, let n be the number of symbol node, m be the number of check node; define $D_s = \{d_{s_1}, d_{s_2}, \dots, d_{s_n}\}$, where D_s is the degree sequence of symbol node. For the i -th symbol node of the Tanner graph, it is necessary to add d_{s_i} edges to connect it with d_{s_i} check nodes, and the selection scheme of the d_{s_i} edges are related to the structure of the Tanner graph that has been constructed.

The following are the specific steps of the LDPC check matrix H generation scheme based on the K-PEG algorithm:

1. Convert the binary key K sequence into the corresponding decimal integer K_d .
2. In the condition of giving degree distribution, we need to use the key to determine the position of the check node connected to the j th edge of the i th symbol node. The element $c_{i,j}$ in the key-location collection $L_{i,j}$ corresponds to

$$\text{index}_{i,j} \equiv (K_d \bmod X_{i,j}) + 1 \quad (1)$$

$$c_{i,j} = L_{i,j}(\text{index}_{i,j}) \quad (2)$$

$L_{i,j}$ denotes a position set of the check node which the j th edge of the i th symbol node is connected to. $X_{i,j}$ represents the number of check nodes that the j th edge of i th symbol node can connect to. And $X_{i,j}$ equals to length ($L_{i,j}$). $\text{Index}_{i,j}$ represents a certain index of the position set $L_{i,j}$. $c_{i,j}$ denotes the position value in $L_{i,j}$ that the $\text{Index}_{i,j}$ pointed to. For example, $c_{5,3} = 30$ represents the third edge of the 5th symbol node connects the 30th check node. And it means that the third nonzero element of the 5th symbol node should be placed on the 30th row of the 5th column of the matrix H.

To ensure that there are no 4-rings in the matrix H, a lower triangular structure can be formed with the degree of the last m symbol nodes not exceeding 2 (this lower triangular structure is usually a dual-diagonal structure). The number of check nodes connected to symbol nodes is always 1.

The above steps are performed separately for each updated key K. In each Tanner graph, an LDPC code check matrix H corresponding to the key K can be obtained after all edges of n symbol nodes are added.

We can make sure that different keys generate different Tanner graphs to guarantee that the key space is not reduced during the conversion from the key sequence to the LDPC code check matrix in this algorithm.

4 Reliability Analysis of the Proposed K-PEG Algorithm

We analyze the anti-jamming ability of this scheme in signal transmission by analyzing the error correction ability of H constructed by K-PEG and comparing it with the one constructed by the improved PEG algorithm. The two algorithms are implemented in the program on a simulation platform. The matrix H of the LDPC code is constructed with the same parameters, and the corresponding generation matrix G for encoding is generated. In the simulation, we simulate the matrix hopping strategy of the secure communication scheme, and repeatedly perform key extraction at a certain frequency and synchronously generate a new matrix of encryption and coding.

The coding and decoding experiments are performed in AWGN channel and BPSK modulation mode, using BP iterative decoding algorithm with the maximum number of iterations of BP decoding set to 50. First, we simulated under a short code length and set the code length with 1024 and code rate with $1/2$. The degree distribution used in the simulation is $0.38354x + 0.04237x^2 + 0.57409x^3$, and the degree distribution has better performance. In the experiment, a total of 10,000 frames of randomized messages were generated for coding and decoding. By using the K-PEG algorithm, a binary key sequence with the same length as the code length is regenerated every 10 frames, and a corresponding check matrix H is reacquired according to the key sequence, so that a total of 1000 matrixes of H are generated.

Comparison of the performance curves of LDPC codes with a code length of 100 under both methods shows in Fig. 1. There is almost no difference in performance between K-PEG algorithm and the improved PEG algorithm in the ability of decoding.

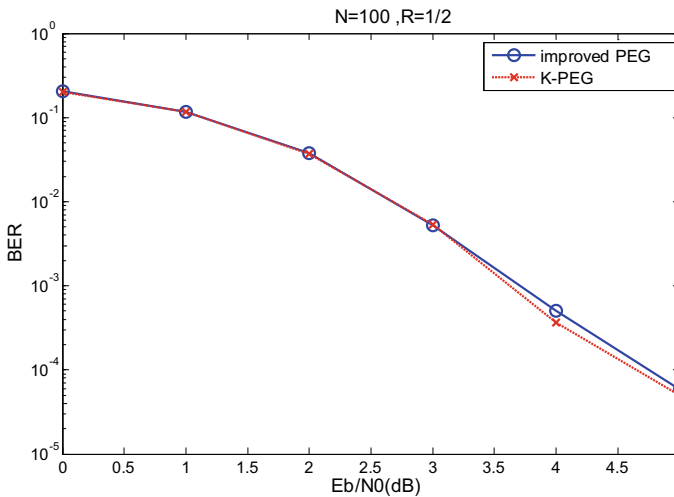


Fig. 1. Comparison of K-PEG algorithm and improved PEG algorithm in short-length code

In order to compare the performance of the proposed K-PEG algorithm and the original RDF scheme in the paper [6] under the same code length and bit rate, the two

schemes are simulated and the simulation result is shown in Fig. 2. From this figure, the proposed K-PEG algorithm is better than RDF in the performance of decoding under the same parameters. The performance improvement is about 1 dB in case of 10^{-4} bit error rate.

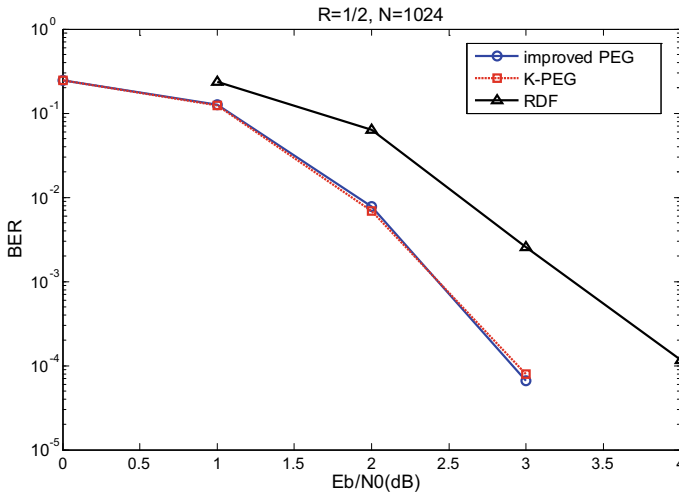


Fig. 2. Comparison of simulation of K-PEG algorithm and improved PEG algorithm in medium-long code with RDF algorithm

It can be shown that mapping the random binary key sequence to the parity check matrix H generated by the PEG algorithm does not substantially affect the decoding performance of the LDPC code. The simulation results also fully demonstrate that this massive matrix transition scheme does not reduce the reliability while improving the security of the communication system.

5 Security Analysis of the Proposed K-PEG Algorithm

Let X_i be the number of different schemes to establish edges denoted by the degree of the i th symbol node. Let $X_i(j)$ be the number of check codes which be collected by the j th edge ($j \in (1, \dots, d_{s_i})$) of the i th symbol code. All optional position indices of the j th edge make up the position set $L_{i,j}(X_i(j) = \text{length}(L_{i,j}))$.

It is easy to know that the key space is 2^n while the length is n . For any two different keys K_{d1} and $K_{d2}(K_{d1} > K_{d2})$ with the same sequence of index,

$$\forall i, j, K_{d1} \bmod X_i(j) \equiv K_{d2} \bmod X_i(j) \quad (3)$$

The number of sequences of the index is

$$Y = \min(K_{d1}-K_{d2}) = [X_1(1), X_1(2), \dots, X_n(d_{s_n})] \tag{4}$$

$[X_1(1), X_1(2), \dots, X_n(d_{s_n})]$ represents the least common multiple of all elements in $\{X_1(1), X_1(2), \dots, X_n(d_{s_n})\}$.

$\forall i, j, K_{d1} \bmod X_i(j) \equiv K_{d2} \bmod X_i(j)$, there must be $K_{d1}-K_{d2} \equiv 0 \bmod X_i(j)$, that is, all elements in the $\{X_1(1), X_1(2), \dots, X_n(d_{s_n})\}$ are factors of $K_{d1}-K_{d2}$. Calculating the minimum value of $K_{d1}-K_{d2}$ equals to calculating the least common multiple of all elements in $\{X_1(1), X_1(2), \dots, X_n(d_{s_n})\}$.

There is a situation that the sequence of position index is different with the same type of connection. The number of placement schemes is equal to the number of different arrangements of the last nonzero elements in the i -th column of the check matrix H. For example, when $d_{s_{100}}$ equals 2, $c_{100,1} = 5, c_{100,2} = 7$ with their position index of $\text{index}_{100,1} = 5, \text{index}_{100,2} = 6$ and $c_{100,1} = 7, c_{100,2} = 5$ with position index of $\text{index}_{100,1} = 7, \text{index}_{100,2} = 5$ is different in index, but the position distribution of “1” in the 100th column of the check matrix H is the same. There are $d_{s_i}!$ repeats for each type of connection for each symbol node. X_{total}^k represents the number of different H constructed by K-PEG algorithm.

$$X_{total}^k \geq \frac{Y}{\prod_{i=m+1}^n (d_{s_i}!)} = \frac{[X_1(1), X_1(2), \dots, X_n(d_{s_n})]}{\prod_{i=m+1}^n (d_{s_i}!)} \tag{5}$$

The simulation uses the code rate of 1/2, and the degree distribution is

$$0.38354 * x + 0.04237 * x.^2 + 0.57409 * x.^3 \tag{6}$$

and matrix H with a code length of 100–1100 is randomly constructed. H is constructed 1000 times for each code length, and the number of different H-matrixes finally obtained is the average of 1000 times configuration. Simulation results are shown in Fig. 3.

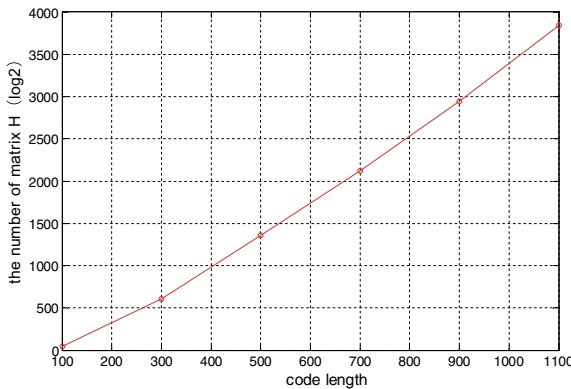


Fig. 3. The approximate number of check matrices H under different code lengths for the algorithm

With the above conclusions, we can determine the size of the key space based on the security requirements of the communication system. According to the simulation results, it is recommended that the length l_k of the key sequence is increasing with the increasing of code length. We recommend that the key length should be between $2n$ and $3n$. According to the simulation, the k-PEG algorithm can obtain a large number of equivalent matrices that guarantee the security of the communication scheme.

6 Conclusion

This paper proposes a K-PEG algorithm for constructing parity check matrixes of LDPC codes based on key sequences, analyzes the calculation method for the number of different parity check matrixes that the algorithm could construct. And we get the approximation of the number of different parity check matrixes in different lengths through simulation. Simulations show that K-PEG algorithm can generate a large number of LDPC code equivalent matrices. Under the same code length and code rate, the error correction performance of these LDPC matrices is better than that of QC-LDPC codes generated by the original RDF scheme. Therefore, the proposed K-PEG-LDPC code in this paper is a better choice for the proposed secure and reliable communication scheme.

Acknowledgement. This work was in part supported by the Sichuan key research and development project under grand 19ZDYF0489.

References

1. Lobstein, A., Cohen, G.: On the inherent intractability of certain coding problems. *Rairo Inform. Théor. Appl* (1), 25–32 (1987)
2. McEliece, R.J.: A public-key cryptosystem based on algebraic coding theory. *Deep. Space Netw. Prog. Rep.* **44**, 114–116 (1978)
3. Wang, X.M.: Generalization of M public key system and analysis of its performance on noisy. *Acta Electron. Sin.* (4), 84–91 (1986)
4. Rao, T.N.R.: Joint encryption and error correction schemes. In: *Proceedings of the 11th annual international symposium on Computer Architecture*, Ann Arbor, Mich, vol. 12 no. 3, pp. 240–241 (1984)
5. Leng, W.Y.: Research on McEliece cryptosystem based on LDPC codes. *Beijing Univ. Posts Telecommun.* 1–6 (2011)
6. Shi, Z.P.: Research on secure and reliable communication method based on LDPC codes. *Univ. Electron. Sci. Technol. China* (5), 641–647 (2017)
7. Jiang, D.S., Jin, L.J.: Research on synchronization technique for a high-speed FH communication system. *J. Univ. Electron. Sci. Technol. China* **1**(34), 48–52 (2005)
8. Zhao, D.F.: An improved PEG construction algorithm for LDPC codes. *Appl. Technol.* **35**(8), 9–11 (2008)
9. Richardson, T.J., Urbanke, R.L.: Efficient encoding of low-density parity-check codes. *IEEE Trans. Inf. Theory* **47**(2), 638–656 (2002)



User Relationship Privacy Protection on Trajectory Data

Zi Yang, Mingda Yang, and Bo Ning^(✉)

DaLian Maritime University, Dalian, Liaoning, China
ningbo@dlmu.edu.cn

Abstract. Various mobile devices facilitate the users' life, but the issues are brought into privacy focus by the individuals. This paper aims at the protection of intimate relationships among users. We consider the intimacy of user relationships based on similar sub-trajectories between the users. Then, we propose a k_{mn} -anonymity protection model. We generalize from two aspects: location and time. The first is location generalization. The range that user pass within the time that the location point stays is the generalization region, and the corresponding location point in the region's trajectory is represented by the generalization region. When the location generalization is not enough to satisfy the k_{mn} -anonymity, then we use time generalization. Finally, the performance of our algorithm is evaluated by the experiment and the validity of our algorithm is verified.

Keywords: Relationship protection · Trajectory · Intimate relationship · k_{mn} -anonymity

1 Introduction

The applications based on mobile devices bring great convenience to people, but they also bring challenges to users' privacy issues [1]. The users' trajectory data contain important personal information. The attacker can infer the home address, company, and other preferences from the users' trajectory, and it is easy for attacker to know the users' relationships if they want. The topic of re-identify attack has been widely studied in previous works. There are several commonly used privacy protection models, which including that the k -anonymity, the l -diversity, the t -closeness, and the differential privacy, respectively. The k -anonymity means that each user is indistinguishable from at least $k-1$ other users [2]. The probability of discovering a target user is $1/k$; the basic idea of the l -diversity is to ensure that at least l distinct value exists for the sensitive attributes of each equivalence group [3]; the t -closeness method requires that the difference between the distribution of sensitive attribute values and the distribution of the original data does not exceed t [4]. The differential privacy [5, 6] has strict theoretical guarantees. It does not care about the attacker's background knowledge.

In this article, we propose an algorithm to protect the privacy of intimate relationships between users. The protection model used in this paper is based on the k -anonymous model. We call it k_{mn} -anonymity. It is still essentially k -anonymity, where n is the number of user relationships that need to be protected and m is the number of people whose trajectories are similar. This model can protect the relationships among multiple users. Compared with k -anonymity, k_{mn} -anonymity can better reflect the utility of data, and it can also improve the efficiency of protection. To achieve k_{mn} -anonymity, we use the method of generalization. We use the edit distance to measure data utilities.

Paper [7] introduces the technique of Ambiguity in Social Network data (ASN) based on anatomy, which specifies how to publish social network data. Paper [8] presents an algorithm of location privacy protection which is designed to calculate the probability of relationship privacy risk, however, they did not elaborate on the steps of protection. Paper [9] proposes a new privacy metric p -confidentiality that ensures location diversity by bounding the probability of a user visiting a sensitive location with the p input parameter. In paper [10], they use machine learning techniques to classify user pairs as friends and non-friends.

In this paper, first of all, we present the k_{mn} -anonymity protection model to protect the intimate relationship among the users. According to the k_{mn} -anonymity protection model, we find the intimacy relationship that needs protection. Then we use global generalization method to protect them. Finally, we conducted an experimental verification and the experimental results show that our algorithm is useful.

Contribution is as follows:

- (1) We present a k_{mn} -anonymity model to protect intimate relationship among users.
- (2) We use a generalization approach method to protect the intimacy relationship that needs protection.
- (3) We evaluate the algorithm on real data sets. The results show that the proposed algorithm is feasible and efficient.

This paper is organized as follows:

In the Sect. 2, we give the relevant definitions; in Sect. 3, we design the corresponding algorithm to protect the intimate relationship that need protection. In Sect. 4, the experimental verification, and evaluation are performed under the real data set. At last, we come to the conclusion.

2 Related Definition

Definition 1. (*Trajectory*) The trajectory is defined as follows:

$$tr_i = p_1^i(t_{1.ar}^i \sim t_{1.le}^i) \rightarrow p_2^i(t_{2.ar}^i \sim t_{2.le}^i) \rightarrow \dots \rightarrow p_z^i(t_{z.ar}^i \sim t_{z.le}^i)$$

The location point p contains the latitude and longitude of the location. The z represents the total number of locations in the user i 's trajectory. The $t_{k.ar}^i$ represents the time to arrive of user i in the k th location. Similarly, the $t_{k.le}^i$ represents the time to leave of user i in the k th location

Definition 2. (*Trajectory Set*) Trajectory data set $Tr = [tr_1, tr_2, \dots, tr_n]$, the tr_i defines the original trajectory of the user i , where the n represents the number of users.

Definition 3. (*Similar Sub-trajectory*) The trajectory consisting of the following three constrained position points is called the similar sub-trajectory, We use Sub_{ij} to represent it. The implication is that two users, i and j , stay in the same location for a period of time, and there is a similar transfer time to the next same location.

- (1) $\forall 1 \leq u \leq z, \forall 1 \leq v \leq z', p_u^i = p_v^j$;
- (2) $Cov_{uv} = |\min(p_{k.le}^i, p_{k.le}^j) - \max(p_{k.ar}^i, p_{k.ar}^j)| > 0$, it refers to the time covered by two users at the same location, that is, when two users stay together at one location;
- (3) $\forall 1 \leq u \leq z, \forall 1 \leq v \leq z', \Delta t_u^i = |t_{u.ar}^i - t_{u'.le}^i|, \Delta t_v^j = |t_{v.ar}^j - t_{v'.le}^j|, |\Delta t_u^i - \Delta t_v^j| \leq T_{time}$.

u' and v' represent the previous locations of u and v , respectively, and T_{time} is a preset threshold for the transition time from one place to another.

Definition 4. (*Anonymous model: k_{mn} -anonymity*)

$$k_{mn} = \frac{1}{C_m^n}, m = \operatorname{argmin}(C_m^n \geq k). \quad (1)$$

The k_{mn} -anonymity of this article is a variant of k -anonymity. It is to protect the privacy of the relationship among users. m refers to the total number of similar trajectories, n refers to the need to protect the relationships among n users, that is, finding n similar trajectories from m . n is the user's own designation and k is the privacy requirement that the user wants to achieve.

For example, suppose the user's privacy requirement is $k = 1/3$, and the relationship between two users needs to be protected, $n = 2$. Then the smallest m that satisfies the condition is equal to 3. That is, we need to find three trajectories that are related to each other. Then the probability of finding the target relationship is $k_{mn} = \frac{1}{C_3^2} = \frac{1}{3}$.

3 Global Generalization

In this section, we propose specific protection measures for the relationships that need to be protected. We mainly use generalization methods. We map the locations to the actual geographic location network and operate the user's trajectory in the geographic location network to achieve the protection of the users' intimate relationships. The meaning of global generalization is to generalize multiple locations into a whole area, and the corresponding locations of the trajectories that pass through the area are replaced by area. The following are the specific steps.

Definition 5. (*Neighbor*) Taking the location before the location in the similar sub-trajectory of the relationship r_{ij} as the center of a circle, taking the distance from the previous location to the location in the similar sub-trajectory as the radius to make a circle 1. Similarly, doing the same operation for the latter position of this location in the similar sub-trajectory to make circle 2. Then the intersection of circle 1 and circle 2 is called the neighbor of the location.

For example, in Fig. 1, it is a neighbor graph. The q_1 is the neighbor of A in the sub-similar trajectory about user 1 and user 2. In like manner, q_2 is the neighbor of B , q_3 is the neighbor of C .

The first step is finding neighbors at each location in the similar sub-trajectory;

The second step is global generalization. First, we generalize the location and then we generalize the time. If the k_{mn} -anonymity has been met after the location generalization, the time generalization is no longer continued. If not, the time generalization is continued.

We denote each location in the similar sub-trajectory combined with all neighboring locations of this location as a generalized region q . The corresponding location in the trajectory passing through the area q is represented by q . If the number of all trajectories passing through the all generalized region $[q_1, q_2, q_3, \dots]$ of the similar sub-trajectory satisfies k_{mn} -anonymity. The time generalization is not continued. On the contrary, generalizing time as the maximum time range for the corresponding location in the generalization region. If calculating intimacy still does not satisfy k_{mn} -anonymity, then the second step is repeated based on the trajectory of the newly added region q , the neighbor is repeatedly searched, and the generalized region is expanded until relationships meets k_{mn} -anonymity. The example is shown in the figure below. The pseudo code is as follows Algorithm 1.

The example data is as follows.

$tr_1: A(7:00 \sim 9:00) \rightarrow B(11:00 \sim 13:00) \rightarrow C(15:00 \sim 17:00) \rightarrow D(19:00 \sim 21:00)$

$tr_2: A(8:00 \sim 10:00) \rightarrow B(12:00 \sim 14:00) \rightarrow C(15:00 \sim 17:00)$

$tr_3: N(7:00 \sim 9:00) \rightarrow O(11:00 \sim 13:00) \rightarrow P(15:00 \sim 18:00)$

$tr_4: Q(7:00 \sim 9:00) \rightarrow U(11:00 \sim 13:00) \rightarrow R(15:00 \sim 18:00)$

First of all, in Fig. 1, ABC is a similar sub-trajectory of Sub_{12} . We have got the neighbors of Sub_{12} , suppose that $k = 1/3$, r_{12} is the relationship that need protection. Second, A 's neighbor area is represented by q_1 , which contains E, H ; B 's neighbor area is represented by q_2 , which contains F, I ; C 's neighbor area is represented by q_3 , which contains G, J . According to the global generalization rule, the corresponding location in the trajectory that passing through the generalized region is represented by a generalized region. From Fig. 1, we can see that tr_3 and tr_4 pass through the generalization region, we should change the location of original trajectory to the generalization region. After the calculation, the relationship satisfies k_{mn} -anonymous and no further time generalization. So, the final trajectory set is:

$tr_1: q_1 (7 : 00 \sim 9 : 00) \rightarrow q_2 (11 : 00 \sim 13 : 00) \rightarrow q_3 (15 : 00 \sim 17 : 00) \rightarrow D (19 : 00 \sim 21 : 00)$
 $tr_2: q_1 (8 : 00 \sim 10 : 00) \rightarrow q_2 (12 : 00 \sim 14 : 00) \rightarrow q_3 (15 : 00 \sim 17 : 00)$
 $tr_3: q_1 (7 : 00 \sim 9 : 00) \rightarrow q_2 (11 : 00 \sim 13 : 00) \rightarrow q_3 (15 : 00 \sim 18 : 00)$
 $tr_4: q_1 (7 : 00 \sim 9 : 00) \rightarrow q_2 (11 : 00 \sim 13 : 00) \rightarrow q_3 (15 : 00 \sim 18 : 00)$

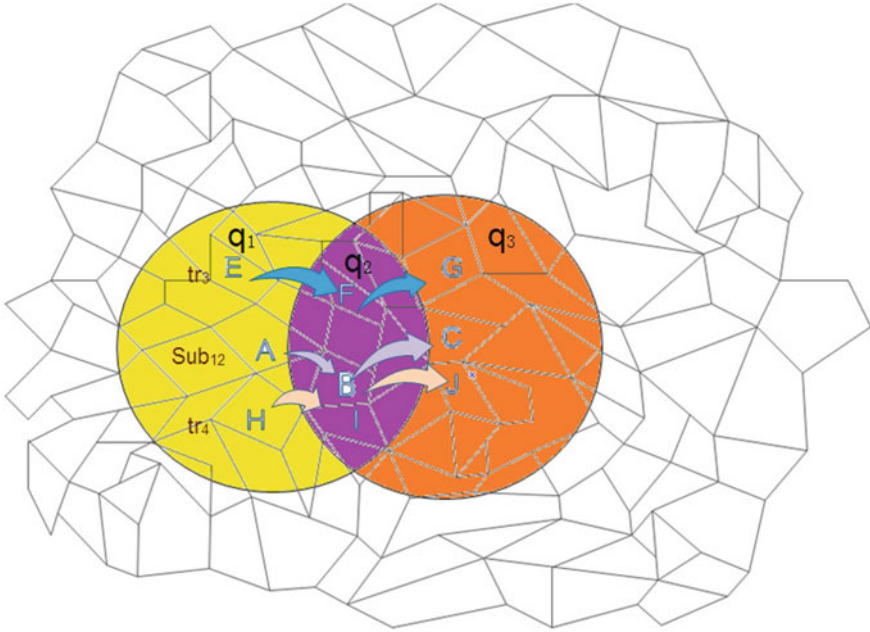


Fig. 1. Neighbor graph

4 Experiment

4.1 Experimental Setup

The algorithms were implemented in JDK 1.7. The trajectory data set in BeiJing, it has 1042 trajectories of 20 users. Attributes include user name, date, location name, latitude and longitude, time of arrival and leave etc. We test two groups of experiments, it includes the quality of the algorithm and the runtime of the algorithm. We use the edit distance to measure the usefulness of the data. Edit distance is a measure of string changes.

4.2 Performance Study

The Quality of the Two Protection Algorithms with the k and n We test the quality of the protection algorithms with the k and n in Fig. 2. When

Algorithm 1: *Global generalization*

Data: Trajectory data set: Tr, k, n
Result: Modified trajectory data set Tr'
begin
 Calculating m according to n and k ;
 for $tr_i \in TR$ **do**
 $R_need_protect[] =$ *The trajectories does not satisfy the*
 k_{mn} - anonymous;
 for $R_1 \in R_need_protect$ **do**
 similar sub - trajectory = Finding similar sub - trajectory of R_1 ;
 while *Relationship does not meet k_{mn} -anonymity do*
 $q =$ *finding neighbors of each location from the similar*
 sub - trajectory ;
 $tr[]$ *is trajectory set of passing q ;*
 $tr' =$ *generalizing the locations of trajectories passing*
 through area q ;
 $sim =$ *Calculating the intimacy of the tr' and similar*
 sub - trajectory ;
 if *sim satisfied the k_{mn} -anonymity then*
 $Tr' =$ *Trajectories after generalization; return Tr' ;*
 else
 generalizing the time of tr' and similar sub - trajectory;
 $sim =$ *Calculating the intimacy of the tr' and similar*
 sub - trajectory ;
 if *sim not satisfied the k_{mn} -anonymity then*
 $Tr' =$ *Trajectories after generalization; return Tr' ;*
 else
 $Tr' =$ *Repeatedly generalize location and time;*
 end if
 end while
 end for
 end for
return Tr' .

$n = 2$ or $n = 3$, the algorithm is ladder-like because our k_{mn} algorithm can satisfy multiple k conditions. When $n = 3$, the quality of the algorithm is greater than when $n = 2$. That is because protecting the relationship among three individuals generally makes the original data change more significant than the two individuals.

Time Efficiency We tested the time efficiency changes of k and n with the protection algorithm in Fig. 3. When $n = 2$ or $n = 3$, the algorithm is ladder-like, and the reason is consistent with the cause of quality. When $n = 3$, the time efficiency of the two algorithms is greater than $n = 2$, because protecting the relationship of three individuals generally take more time.

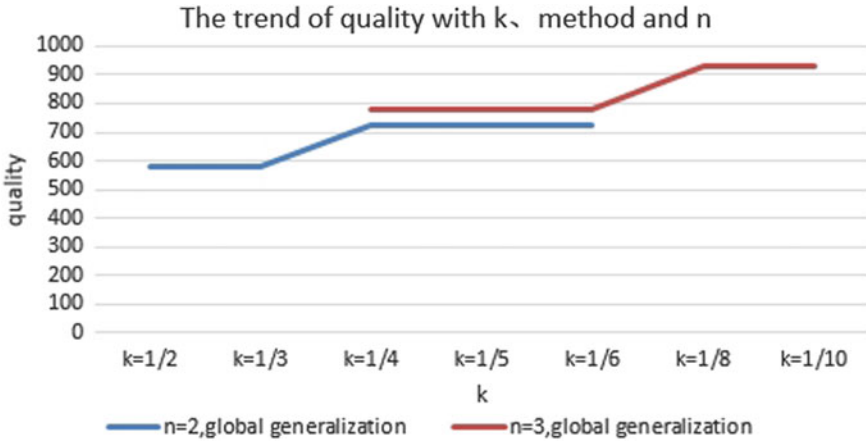


Fig. 2. The quality of the protection algorithm with the k and n

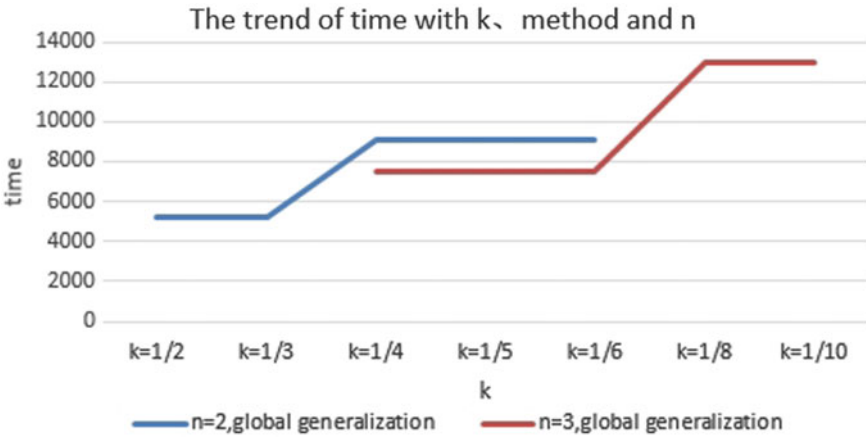


Fig. 3. Time efficiency

5 Conclusions

In this paper, we are mainly about protection of intimate relationships. For the protection of relationships, we have adopted the global generalization. The idea of global generalization is to use generalized regions to represent locations in the trajectories that go through the regions. In the experimental part, we evaluate the effectiveness of the experiment on real data. The final experimental results show that our algorithm is effective.

Acknowledgments. The research described in this paper was supported by the National Natural Science Foundation of China (U1401256), the National Natural Science Foundation of Liaoning province (201602094) and the Fundamental Research Funds for the Central Universities (3132018191).

References

1. Xindi, M.A., Hui, L.I., Jianfeng, M.A., et al.: APPLLET: a privacy-preserving framework for location-aware recommender system. *Sci. China (Information Sciences)* **60**(9), 092101 (2017)
2. Sweeney, L.: k-anonymity: a model for protecting privacy. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **10**(5), 557–570 (2002)
3. Machanavajjhala, A., et al.: L-diversity:privacy beyond k-anonymity. In: Proceedings of 22nd International Conference on Data Engineering (ICDE'06), Atlanta, Georgia, USA, pp. 24–36 (2006)
4. Li, N., Li, T., Venkatasubraman, S.: t-Closeness: privacy beyond k-anonymity and l-diversity. In: IEEE, International Conference on Data Engineering, pp. 106–115. IEEE (2007)
5. Kellaris, G., Papadopoulos, S., Xiao, X., et al.: Differentially private event sequences over infinite streams. *Proc. Vldb Endow.* **7**(12), 1155–1166 (2014)
6. Nergiz, E., Atzori, M., et al.: Towards trajectory anonymization: a generalization-based approach. *Trans. Data Priv.* **2**(1), 47–75 (2009)
7. Rajaei, M., Haghjoo, M.S., Miyaneh, E.K.: Ambiguity in social network data for presence, sensitive-attribute, degree and relationship privacy protection. *Plos One* **10**(6) (2015)
8. Feng, Z., Tan, H., Shen, H.: Relationship privacy protection for mobile social network. In: International Conference on Advanced Cloud and Big Data, pp. 215–220. IEEE (2017)
9. Cicek, A.E., Nergiz, M.E., Saygin, Y.: Ensuring location diversity in privacy-preserving spatio-temporal data publishing. *VLDB J.* **23**, 609–625 (2014)
10. Cranshaw, J., Toch, E., Hong, J., Kittur, A., Sadeh, N.: Bridging the gap between physical location and online social networks. In: Proceedings of the 12th ACM International Conference on Ubiquitous Computing, pp. 119–128. ACM (2010)



Research on Modulation Algorithm Based on Physical Layer Encryption

Xiang Li^(✉) and Yueyong Zhang

College of Computer and Information Engineering, Tianjin Normal University,
Tianjin 300387, China

xiangli_tju@yahoo.com, 284480150@qq.com

Abstract. Communication has become a very important part of people's work and life. Wireless communication technology has the characteristics of low cost, good scalability, and easy use, but it is vulnerable to eavesdropping and other security threats, and eavesdroppers illegally receive through wireless channels. Traditional wireless communication system security methods need to be implemented at the upper layer through authentication and cryptographic techniques. In recent years, as an important complement to traditional security mechanisms, information theory security principles, transmission signal security technologies, spread spectrum and frequency hopping encryption technologies, and channel coding Research on physical layer security technologies such as encryption technology and modulation encryption can effectively protect information during wireless transmission and prevent eavesdroppers from illegally receiving information.

Keywords: Wireless communication · Communication security · Information theory security · Physical layer encryption · Modulation

CLC number: TP338.8

Document Code: A

Wireless communication technology is a communication technology that uses electromagnetic signals to exchange information. In recent years, it has developed the fastest and most widely used in the field of information and communication. Wireless communication technology has the advantages of low cost, no need to set up physical lines, no laying of cables, no restrictions on the industrial environment, strong ability to resist environmental changes, easy fault diagnosis, remote maintenance and maintenance, and the advantage of no need for extended expansion cabling [1]. However, for wireless communication systems involving national security, military information, diplomatic, and trade secrets, the security of communications is even more important than the transmission performance.

The physical layer security of wireless communication takes information theory security as the starting point, uses the signal format of wireless communication and the physical characteristics of wireless channels, and combines cryptography to encrypt the specific signal format of the transmitted information and even achieve cross-layer security design [2]. In practical applications, the physical layer security technology has

the advantages of easy implementation and easy maintenance. At the same time, the wide application of multi-carrier, multi-antenna, and new channel coding technologies provides a broad space for physical layer security technology research.

1 Wireless Communication Physical Layer Encryption System Model

The traditional encryption mechanism is divided into two mechanisms: private key (symmetric) and public key (asymmetric) password. The wireless communication system adopting the traditional encryption mechanism completes the encryption operation outside the physical layer, and the ciphertext information is transmitted to the physical layer for wireless transmission. The information is converted into a form suitable for wireless channel transmission. The transmitting end implements functions such as channel coding, digital modulation, baseband post-processing, and radio frequency modulation and transmits it to the receiving end. The receiving end performs radio frequency demodulation and synchronization on the received signal. Channel estimation, digital demodulation, channel decoding, and other operations achieve signal recovery [3]. Therefore, if the physical layer is not reliable, the system security of the traditional encryption mechanism will be reduced. Physical layer security technology is a necessary supplement to the traditional encryption mechanism to

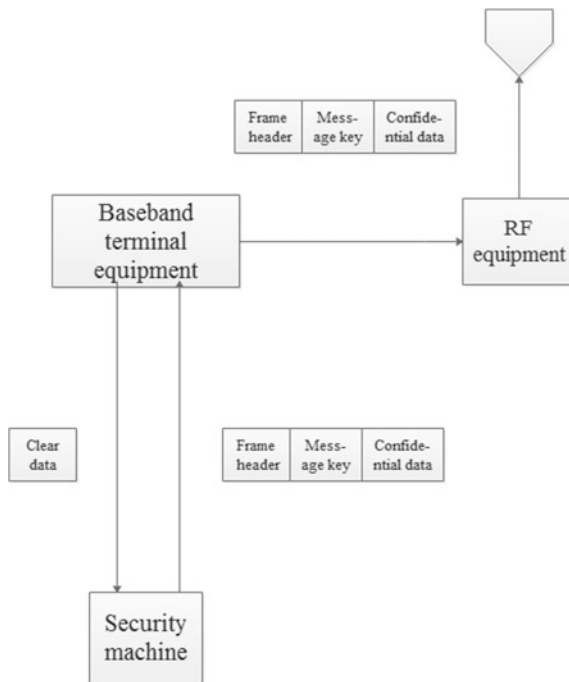


Fig. 1. Communication information encryption system model

improve the protection of ciphertext in wireless transmission, prevent eavesdroppers from illegally receiving information, increase the difficulty of eavesdropping and deciphering encrypted information, and enhance the security of wireless communication.

In communication systems, information can be segmented into fields of the same length and preceded by control information such as packet numbers to form packets of equal length [4, 5]. Therefore, the data packet as the basic encryption unit, each packet encryption, encryption scheme as shown in Fig. 1. At the sending end, the data terminal sends the encrypted plain text to the security device in the form of a packet. The security device encrypts the plain text to form a ciphertext packet and sends it to the terminal device, and sends the data packet in a frame format through the radio frequency device. After receiving the frame, the receiving end decapsulates the frame, and the terminal device sends the ciphertext packet to the security device. The security device decrypts the ciphertext packet to form a plain text packet and sends it back to the terminal device to complete the decryption process.

2 Channel Encryption Technology

Information theory security puts forward the feasibility of implementing confidentiality transmission under the multichannel model. Both parties can achieve confidential communication without eavesdropping through cryptographic techniques. At present, based on the theoretical research category, the theory of information theory security, the study of the channel The model needs to be defined or assumed (the communication channel is better than the eavesdropping channel, and the channel information is known accurately) [6]. It provides a new idea. In practical applications, physical layer security technology is still needed to enhance the security of wireless communication (Fig. 2).

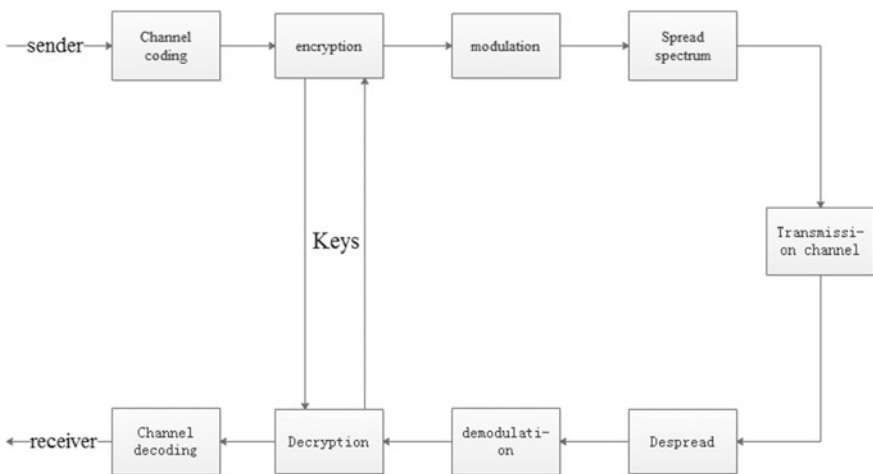


Fig. 2. Channel encryption processing block diagram

The characteristics of channel encryption technology are mainly reflected in: the interceptor intercepts the sender's signal from the transmission channel. If there is no corresponding key, after the spreading code estimation, modulation pattern recognition, and parameter estimation, noncooperative despreading and solution cannot be used. The way to decrypt it. The signal intercepted by the listener is affected by both the channel error and the key deviation. Compared with the conventional key communication, the intercepted signal is affected only by a factor of the subsequent key deviation, which has higher security.

3 Spread Spectrum Frequency Hopping Encryption

Spread spectrum encryption and frequency hopping encryption are currently the most widely used physical layer encryption technologies. Among them, spread spectrum encryption uses high-frequency pseudo-random sequences for spreading modulation, demodulation, and signal spectrum spread [7, 8]. The direct-sequence spread spectrum adopts the M-sequence generated by the linear feedback shift register as the spread spectrum code sequence. At the same time, a high-intensity cryptographic algorithm is used to generate a complex spread spectrum code sequence and is mapped into a pseudo-random high-frequency spread spectrum code sequence according to certain rules., which completes the encryption while spreading.

The spread spectrum signal has excellent characteristics such as low probability of acquisition, low detection probability, and strong anti-interference ability. The spread spectrum encryption, decryption, and despreading processing are shown in Fig. 3. The spread spectrum and the encryption link are combined at the transmitting end, and the long period spreading code is combined. A high-speed pseudo-code sequence is superimposed on the sequence, and the despreading is performed at the receiving end using a key known to both parties of the communication, to obtain a subsequent modulated signal, and then to perform demodulation and channel decoding operations.

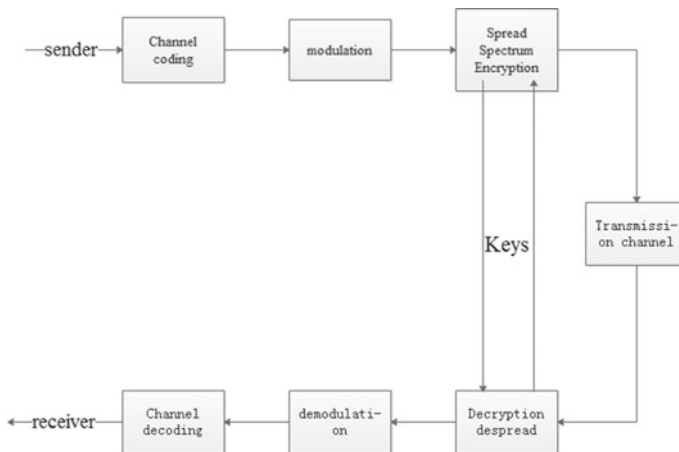


Fig. 3. Spread spectrum despreading processing block diagram

Frequency hopping encryption uses pseudo-random sequences to control the carrier frequency hopping time and duration to achieve the pseudo-randomness of the frequency hopping law. Frequency hopping encryption mainly relies on the frequency hopping rate and the randomness of frequency hopping to enhance electronic opposition and countermeasures. The threat of tracking interference ensures the reliability of communications.

4 Phase Rotation and Amplitude Modulation Modulation Method Encryption Technology

Modulation method encryption technology is an emerging physical layer encryption technology. The phase modulation and amplitude modulation combined modulation encryption algorithm is shown in Fig. 4. In the baseband digital modulation, phase rotation and amplitude adjustment are performed on the constellation mapping part to change the constellation diagram. Distribution, so that illegal users cannot determine the modulation method [9, 10]. At the same time, using the physical irreversibility of the wireless channel, the information bits are mapped in the form of a constellation symbol, and the weak human noise is superimposed on the constellation that has been rotated randomly at the transmitting end to further enhance the physical layer security.

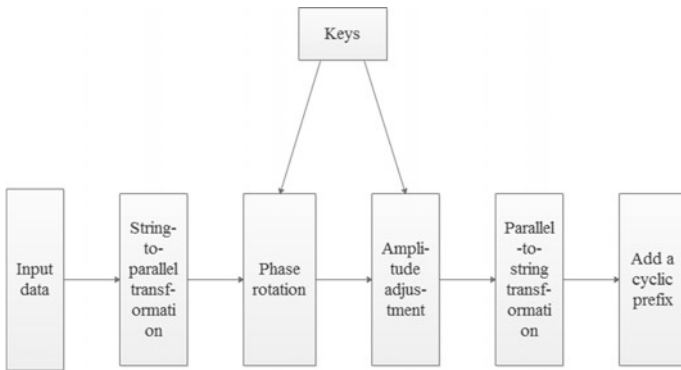


Fig. 4. Phase rotation and amplitude modulation method encryption technology

4.1 Data Preprocessing

The input binary serial data is converted into parallel data and allocated to N subcarriers. It is assumed that each subcarrier is assigned data of K (K depends on modulation) bits, and then the serial data is converted into

$$\begin{pmatrix} a_1^1 & a_2^1 & \cdots & a_k^1 \\ a_1^2 & a_2^2 & \cdots & a_k^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^N & a_2^N & \cdots & a_k^N \end{pmatrix} \tag{1}$$

The a_i^j value is binary 0, 1, the superscript j indicates that the serial data is allocated on the j -th subcarrier, the subscript i indicates the i -th data on the j -th subcarrier, and the data is serially converted and then transmitted in columns:

$$\begin{pmatrix} a_1^1 \\ a_1^2 \\ \vdots \\ a_1^N \end{pmatrix}, \begin{pmatrix} a_2^1 \\ a_2^2 \\ \vdots \\ a_2^N \end{pmatrix}, \dots, \begin{pmatrix} a_k^1 \\ a_k^2 \\ \vdots \\ a_k^N \end{pmatrix} \tag{2}$$

4.2 Key Sequence Processing

Depends on the initial conditions using the generated encryption series, setting $a = a_0a_1a_2 \dots a_i \dots a_j \dots$ P(a) periodic sequence in binary domain F, called the $C_a(x)$ autocorrelation function of the periodic sequence $a = a_0a_1a_2 \dots a_i \dots a_j \dots$

$$C_a(\tau) = \frac{1}{p(a)} \sum_{j=0}^{p(a)-1} (-1)^{a_j + a_{j+\tau}} \quad 0 \leq \tau \leq p(a) - 1 \tag{3}$$

For the average of the statistical eigenvalues of any series:

$$\bar{X} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} X_i = \int_0^1 X_p(X) dX = 0 \tag{4}$$

Two initial values for any sequence-independent sequence correlation function:

$$\begin{aligned} cor(x_1, x_2) &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} (X_i - \bar{X})(Y_i - \bar{Y}) \\ &= \int_0^1 \int_0^1 \rho(X)\rho(Y)(X_i - \bar{X})(Y_i - \bar{Y}) dXdY \\ &= \sigma(x_1 - x_2) \end{aligned} \tag{5}$$

4.3 Phase Rotation

After the constellation mapping, the input binary parallel data is mapped to a complex number, and its angle is added to the rotation angle generated by the key, which completes the angle encryption of the data.

$$m'' = m + m' \tag{6}$$

Encrypted phase m'' , Is the phase mapped by the parallel data constellation m Generated by a rotation phase key generated m' by a chaotic sequence generator.

4.4 Amplitude Adjustment

The amplitude adjustment is to multiply the phase rotated data by the amplitude key:

$$m'' = m * m' \tag{7}$$

The encrypted amplitude m'' is generated by combining the parallel data amplitude m and the amplitude adjustment parameters generated by the hash sequence generator.

5 Simulation Analysis of Influence of Encryption Algorithm on System Performance

Because of the huge number of keys, it illustrates the algorithm’s security, but the need for simulation experiments to determine its encryption effect, the impact on the system performance, the algorithm anti-attack performance were simulated experiments, and through the system before and after the encryption code error. The analysis proves that the encryption algorithm does not destroy the performance of the system.

Figure 5 is the change curve of the error before and after encryption under different SNRs before and after the modulation encryption. In the experiment, an average of 100 units was used for each SNR as the BER at the SNR. From Figs. 5 and 6, it can be seen that the error code curve of the demodulated data after the algorithm basically coincides

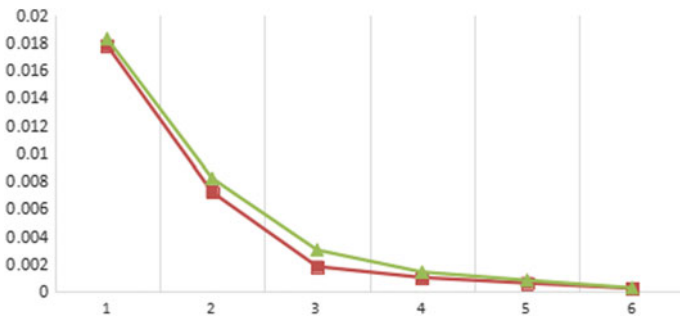


Fig. 5. Comparison of error codes before and after modulation and encryption

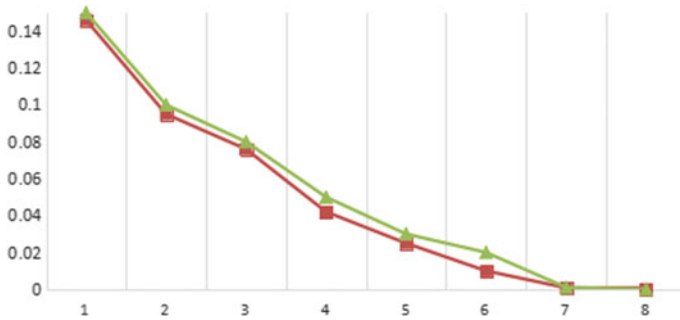


Fig. 6. Comparison of error codes before and after 8PSK encryption

with the error curve of the original modulation method. The modulation and encryption algorithm has no effect on the inherent performance of the system and is a lossless process.

6 Conclusion

The physical layer security technology of wireless communication covers a wide range of technical means, and plays an important role in enhancing the security of the system. Random changes in system configuration and dynamic coordination of internal resources of the system all require encryption. It can be seen that the dynamic target defense has proposed a long-term development goal for the cross-layer security design of wireless communications, and this will also create a new situation for the development of communications security technology. Based on the modulation protection algorithm of the physical layer encryption, its performance is analyzed from the theoretical and experimental perspectives. The theoretical analysis of the security of the key and the ability to resist the attack was performed, and corresponding simulation experiments were designed.

References

1. Sun, K., Wu, D.: MPC-based delay-aware fountain codes for live video streaming. In: IEEE International Conference on Communications, ICC (2016)
2. Lu, Z., Lin, X.: A data privacy-oriented multi-parities location collect scheme in location based services. In: Proceedings of 2009 the Fourth International Conference on Computer Sciences and Convergence Information Technology. 964–969 (2009)
3. Liu, X., Li, Z.: Information-Centric Mobile ad hoc Networks and Content Routing: A Survey. Article in Press (2016)
4. Sun, K., Zhang, H., Wu, D., Zhuang, H.: MPC-based delay-aware fountain codes for real-time video communication. IEEE Int. Things J. (IoT) (2016)
5. Li, Z., Chen, Y.T., Shi, H., Liu, K.: NDN-GSM-R: a novel high-speed railway communication system via named data networking. EURASIP J. Wirel. Commun. Netw. (2016)

6. Sun, K., Wu, D.: Video rate control strategies for cloud gaming. *J. Vis. Commun. Image Represent.* (2013)
7. Li, Z., Liu, K.: MaPIT: an enhanced pending interest table for NDN with mapping bloom filter. In: *IEEE Communications Letters*, vol. 18, no. 11, Nov 2014
8. Sun, K., Yan, B., Gharavi, H.: Low complexity content-aware image retar-geting. In: *IEEE International Conference on Image Processing 2012 (ICIP'2012)* (2012)
9. Li, Z., Song, L., Shi, H.: Approaching the Capacity of K-user MIMO Interference Channel with Interference Counteraction Scheme. *Article in Press* (2016)
10. Sun, K., Yan, B.: Efficient P-frame complexity estimation for frame layer rate control of H.264/AVC. In: *IEEE International Conference on Image Processing 2011 (ICIP'2011)* (2011)



Threat Modeling for Cyber Range: An Ontology-Based Approach

Lei Gong^(✉) and Yu Tian

China Academy of Electronics and Information Technology, Beijing 100041, China
bracey@qq.com

Abstract. Cyber Range has become a very important means to support tasks such as network security technology validation, network weapon testing, training of network attack and defense and network risk assessment. However, Cyber Ranger faces many security threats from internal and external environments. In order to establish an adaptive security protection system, threat modeling is needed to analyze potential threats and provide security solutions. In this paper, we present a novel threat modeling method for Cyber Range. Based on ontology and knowledge graph, our research focuses on the design of threat ontology, knowledge base, and unified description specification. Typical cases are given to demonstrate our approach. This study could serve as groundwork for further Cyber Range researches including security architecture, situation awareness and intelligent decision-making.

Keywords: Cyber range · Threat modeling · Ontology · Knowledge graph · Threat knowledge extraction

1 Introduction

With the rapid development of information technology, network attack and defense has become the main content of national confrontation between various countries. Network environment is changing from Internet to ubiquitous network space, while attack method is developing from single mode to Advanced Persistent Threat. *Cyber Range* is an important infrastructure for government and military departments, which can significantly improve stability, security, and performance of critical systems and infrastructures. Many countries in the world pay great attention to the construction of Cyber Range, which has become a very important means to support tasks such as network security technology validation, network weapon testing, training of network attack and defense and network risk assessment.

Compared with other systems, Cyber Range is more complex and changeable, which not only satisfies the need of network attack and defense experiments, but also ensures its own safety. The threats come from external environments as well as internal elements and other Cyber Ranges. So it needs to establish a

© Springer Nature Singapore Pte Ltd. 2020

Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 517,

https://doi.org/10.1007/978-981-13-6508-9_128

dynamic security protection system, which can adjust security policies automatically along with the changes of external conditions and internal states. *Threat Modeling* is the foundation of security architecture design. Once threat model of Cyber Range is established, we can identify as many known threats as possible, and contribute to analysis and discovery of unknown threats.

In the literature, *Threat* is defined as the potential event can lead to undesired results, such as Information Disclosure or Deny of Service. *Threat Modeling* is defined as the process of recognizing, presenting, modeling and reducing potential threats that system may encounter with. It succinctly describes the methods attackers exploit to hack systems and can be used in whole life cycle of the system.

Starting with the requirements of threat modeling, a novel threat modeling method is derived using ontology and knowledge graph. Our research focuses on design of threat ontology, knowledge base, and unified description specification. Remember that the approach presented in this paper is capable of dynamic updating and attribute extension. This research could support security protection of Cyber Range properties. Furthermore, it could serve as groundwork for further Cyber Range researches including security architecture, situation awareness, and intelligent decision-making.

The rest of this paper is organized as follows. First, in Sect. 2, related works are introduced. A treat modeling method based on ontology is proposed in Sect. 3. Then, in Sect. 4 we design an innovative knowledge base of threats based on knowledge graph and set up an unified description specification. Some cases are presented in Sect. 5 to verify our proposed approach. Finally, Sect. 6 draws conclusion.

2 Related Works

STRIDE model [1] has been the primary method of threat modeling, which is based on the security attributes of systems. The model divides threats into Spoofing, Tampering, Repudiation, Information disclosure, Denial of service, and Elevation of privilege to ensure the system has security attributes including authentication, confidentiality, non-repudiation, integrity, availability, and authorization. The advantage of the STRIDE model is that it inherently suppresses threats based on security attributes. However, the classification of threats by STRIDE model is one-dimensional. The system may be attacked by a certain kind of threat, but it is not sure at which level the attack occurs. Lack of consideration for the OSI model, the STRIDE model is difficult to evaluate the effectiveness of the attack and make the choice of defense strategies.

The tree-structured models applied to the field of security analysis include attack tree, fault tree, and threat tree. Attack tree [2] is a formal description of attack behavior in network based on tree structure. Fault tree [3] is also called accident tree, which is an effective tool for analyzing the safety and reliability of complex and large-scale systems. Threat tree model is a graphical model that describes threats to a system, such as malicious damages to the availability, integrity, or confidentiality of the system.

There are two types of nodes in the privilege graph model [4], the attacker node and the target node. The attacker node represents an attacker, and the edge between the attacker node and the target node represents the set of permissions that the attacker originally possessed. If there is a reachable path between the attacker node and the target node, it indicates that there is a potential security threat.

The attack graph can be used in the form of directed graph, which can describe the network relationship, service relationship and vulnerability utilization relationship between hosts. Due to the different definitions of nodes and edges, researchers proposed state attack graph model [5] and attribute attack graph model [6]. The difference between them is that in attribute attack graph, the execution of a certain attack will only cause the change of adjacent nodes, while in state attack graph it will cause the change of global state, thus avoiding the problem of state space explosion.

Attack tree, privilege graph, and attack graph are three popular threat modeling methods in the academic community. They all successfully prove that within their own predefined scope, analysis based on trees or graphs is feasible. However, the defect of the existing threat models based on trees or graphs is that it does not have a complete construction of threat ontology and property ontology. For example, privilege graph only pays attention to privileges and users, attack tree and attack graph focus only on hosts, vulnerabilities, rights and networks. The lack of ontology classification makes it difficult to apply above theories to real-world scenarios.

In summary, existing threat modeling methods have their specific application scenarios, and cannot completely meet the business needs of Cyber Range. Therefore, in the process of threat modeling, it is necessary to absorb the granularity of description in various standards, and conduct more rigorous hierarchical classification of attack methods, properties, vulnerabilities and defense methods. At the same time, the relationship between ontologies needs to be considered.

3 Treat Modeling Method Based on Ontology

Threat Modeling for Cyber Range is the procedure of researching potential threats, analyzing necessary factors to generate threats, abstracting and correlating threat factors with attack patterns and expert knowledge, and designing an overall model to identify known threats and guide unknown threats discovery. Traditional threat modeling only analyzes one aspect of the closed loop of security, for example, association analysis based on massive alarm logs and analysis of vulnerability and attack path based on attack graph. Several issues must be solved in threat modeling. That is, identification and classification of complex and diverse properties, construction of overall threat view from the property perspective, extraction of threat knowledge, unified and standardized description of threats, construction of threat knowledge base. In order to solve these problems and make Cyber Range automatically identify threats and carry out adaptive security protection, we propose a threat modeling method based on ontology.

3.1 Ontology Modeling

Ontology [7] is a branch of philosophy that mainly studies the nature of existence. In computer and related fields, ontology refers to the theory of applying the basic method of ontology, that is, concept analysis and modeling, to abstract the entities in the real world into a set of concepts and relations between concepts. Ontology modeling is an engineering activity for acquiring and expressing domain knowledge, which can obtain and formally express concepts in the domain, the relationships between concepts, and domain-related constraints. The objective of ontology modeling approach is to provide an efficient and easy-to-use domain ontology acquisition and expression method for ontology modeling activities so that a high-quality domain ontology can be constructed with lower cost or time.

3.2 Threat Meta-Model

The threat meta-model contains five ontologies, including *Threat Actor*, *Attack Method*, *Property*, *Vulnerability* and *Defense Method*. Threat actor ontology describes an attacker or a threat, which may pose a threat to the target system in one or more layers of OSI model. Property ontology has an OSI attribute that indicates at which level service is provided. It does not distinguish between an attacker property and a target property. The distinction between them depends only on whether or not the attacker controls it. Attack method ontology describes a set of attack means and attack tools used by the attacker for the vulnerabilities of target properties. Vulnerability ontology represents the vulnerabilities of a property as a result of different circumstances and configurations, which can be exploited when there is a reachable path between attacker property and target property. Defense method ontology includes defense resources and defense strategies. There are corresponding defense methods against different properties, vulnerabilities, and attack methods. There is an association between ontologies, which is derived from the common attributes or similar attributes.

3.3 Threat Ontology of Cyber Range

According to the fundamental theory of ontology, the five ontologies of threat can be extended by inheritance. For example, vulnerability and configuration instances can inherit the basic attributes from the vulnerability ontology and extend independently. Threat ontology can identify and classify properties according to actual environment, apply attack graph analysis technology to extract relationships between threat behavior patterns and properties, and use unified threat information to describe the threat view of Cyber Range. At the same time, attack actor, attack method, property, vulnerability, and defense method libraries are established based on threat knowledge, which can support situation awareness and intelligent decision-making. Threat ontology of Cyber Range is shown in Fig. 1.

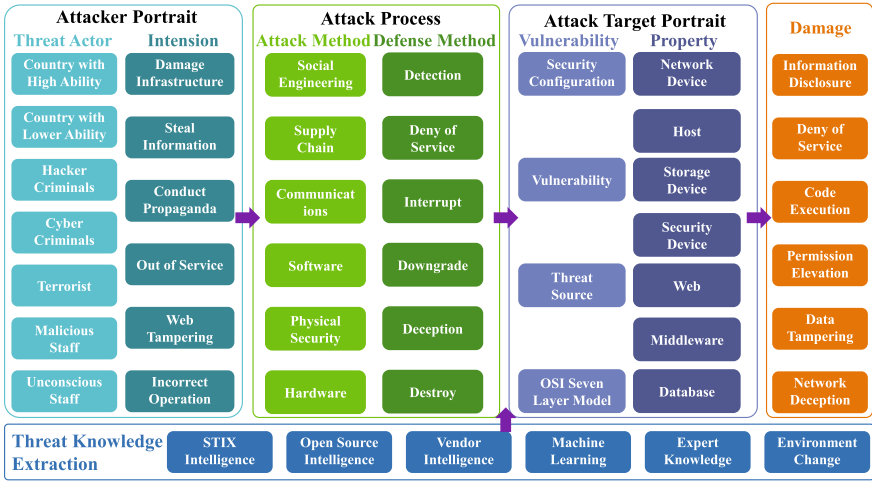


Fig. 1. Threat ontology of cyber range.

4 Threat Knowledge Base Based on Knowledge Graph

Through the research of ontology classification, ontology correlation, and ontology attributes, the design of threat meta-model is basically completed. The next problem to be solved is how to extract corresponding knowledge ontology from massive external data, how to organize the storage structure of threat knowledge and how to use threat knowledge more effectively. In this study, *Knowledge Graph* is used to construct the underlying threat data, that is, threat knowledge base. By means of knowledge graph, the massive external knowledge base is mapped to threat ontology space, which can formally define the relationships between ontologies, can describe, store and calculate instances of the ontology, and can also explore attackers' relationships, attack methods and potential hazards from massive threat knowledge.

4.1 Threat Knowledge Extraction

Threat knowledge extraction is the foundation of knowledge base design, including not only the extraction of entities of five ontologies and related attributes, but also the extraction of relationships between ontologies. There are numerous formats in the vast amount of current security data, semi-structured, unstructured, and structured. It is necessary to use a variety of technologies to extract knowledge, including ETL, wrapper, and semantic extraction. Through knowledge extraction, we have extracted information elements such as entities, relationships, and attributes from the original corpus. The next step is knowledge fusion, which eliminates the ambiguity between entity reference items and entity objects and obtains a series of basic fact expressions.

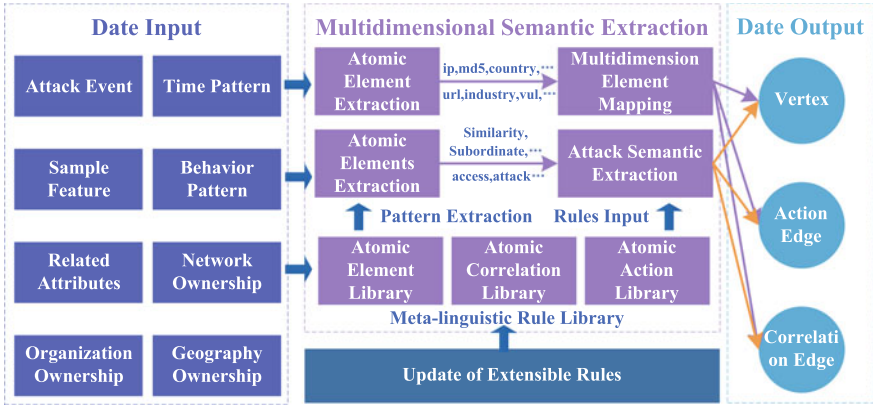


Fig. 2. Construction of knowledge graph.

4.2 Construction of Knowledge Graph

As shown in Fig. 2, attack elements of different sources, different dimensions and different stages are mapped to the vertexes, action edges and correlation edges in knowledge graph to form a more compact data structure, which can solve the fragmentation and organization problem of large-scale attack data. Meanwhile, an extensible attack meta-linguistic rule library is proposed, which can extract data elements and related relationships for hostile organization, attack manipulation, attack process, target property, and countermeasure. The rule organization model based on ontology can accommodate to the incremental fusion of language and syntax, then the rules can be extended and updated, and the attack meta-data can be managed and maintained. The top-level description of knowledge graph constructed is shown as follows.

$$V = (vid, vtype, name, attr_1, \dots, attr_n) \tag{1}$$

$$E = (eid, etype, name, vid_{src}, vid_{dst}, attr_1, \dots, attr_n) \tag{2}$$

V is the vertex element of knowledge graph, all ontologies are inherited from V, and ontology is classified by attributes or ontology inheritance. For example, both threat actor and property ontologies inherit from V. In property classification, the operating system is a secondary classification. The mobile operating system and the host operating system are classified into three levels and both inherit the operating system ontology. E is an edge element that represents the relationship between ontologies. The ontologies are connected through predefined rules, for example, one or more attributes are the same or similar.

4.3 Unified Description Specification of Threats

Current popular information sharing standards such as STIX [8] and OpenIOC [9] have defects such as incomplete ontology classification, incomplete ontology

relationship, and lack of ontology level. It is difficult to directly apply to the description of threat modeling, but their description and organization way of threat ontology should be used for reference. Referring to STIX, OpenIOC, and other standards, this study uses the XML language as the underlying description language of threat modeling, which can describe the five threat ontologies and related attributes, the relationships between threat ontologies, and finally form an overall description of threat event.

5 Case Analysis

5.1 Threat Analysis Based on Attack Path

As Fig. 3 depicted, a threat model based on knowledge graph is established. IP1 is an attacker and IP2 is an IP address in the DMZ area. The attack path of attacker is to use IP1 to infiltrate IP2 to get administrator’s permission, and then penetrate the database server on IP3. The attacker uses the vulnerabilities to construct corresponding data message to infiltrate and finally obtain the administrator rights of IP3. The analysis shows that it will be harmful to the corresponding personnel data after IP3 being captured.

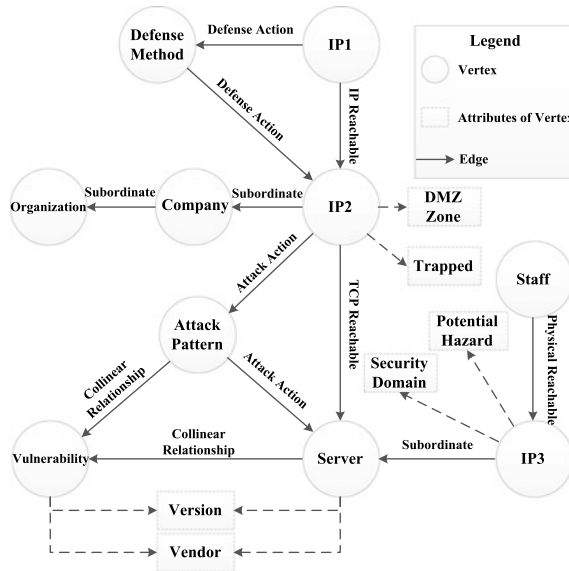


Fig. 3. Example of application of threat knowledge graph.

5.2 Threat Analysis Based on Threat and Defense View

Here is another example of threat modeling scenarios, where a country with high ability steals the database information through Eternalblue virus. In the process of building a threat view, attack method related to the threat actor is detected

first, then attack information is extracted from attack method library, and then attack method is connected in series to form a threat view based on kill-chain model. A similar approach can also be used to build a defense view. By constructing the threat view and defense view, vulnerabilities information of properties can be exported, including incorrect security configuration, vulnerability, lack of security conscious, and abnormality of security audit.

6 Conclusions

Starting with the necessity of hierarchical modeling and multi-ontology association analysis, this paper analyzes the advantages and defects of current modeling methods, and carries out research on ontology-based threat modeling, threat ontology design, knowledge base design based on knowledge graph, and threat unified description specification, which lays a solid foundation for situation awareness, threat analysis, intelligent decision-making and linkage response of Cyber Range.

This paper constructs an ontology-based threat modeling system, and realizes the design of knowledge base based on knowledge graph, which breaks through the key problems of the completeness of ontology classification, the hierarchy of ontology and the application of knowledge base. Meanwhile, this paper solves the problem of the uniqueness and unified classification of property, the prediction of starting point and attack path, and the dynamic updating of knowledge base. It is a methodological and practical innovation in the field of threat analysis.

References

1. Hernan, S., Lambert, S., Ostwald, T., et al.: Threat modeling-uncover security design flaws using the stride approach. *MSDN Mag.-Louisville* 68–75 (2006)
2. Schneier, B.: Attack trees—modeling security threats. *Dr. Dobbs's J.* 24 (1999)
3. Ericson, C.A.: Fault tree analysis. *Hazard Anal. Tech. Syst. Saf.* 183–221 (2005)
4. Ortalo, R., Deswarte, Y., Kaaniche, M.: Experimenting with quantitative evaluation tools for monitoring operational security. *IEEE Trans. Softw. Eng.* **25**(5), 633–650 (1999)
5. Phillips, C., Swiler, L.P.: *A Graph-Based System for Network-Vulnerability Analysis*, 1998. ACM (1998)
6. Jajodia, S., Noel, S.: Topological vulnerability analysis. In: Jajodia, S., Liu, P., Swarup, V., Wang, C. (eds.) *Cyber Situational Awareness. Advances in Information Security*, vol. 46. Springer, Boston, MA (2010)
7. Moore, A.P., Kennedy, K.A., Dover, T.J.: Introduction to the special issue on insider threat modeling and simulation. *Comput. Math. Organ. Theory* **22**(3), 1–12 (2016)
8. Barnum, S.: Standardizing cyber threat intelligence information with the structured threat information eXpression (STIX). MITRE Corp. **11**, 1–22 (2012)
9. Gibb, W.: Nettraveler in openioc format. *FireEye* (2013)



An Analysis of a New Detection Method for Spear Phishing Attack

Yaping Chi^{1,2}(✉), Zhiting Ling^{1,2}, Xuejing Ba¹, and Shuhao Li¹

¹ Department of Communication Engineering, Beijing Electronic Science and Technology Institute, Beijing 100070, China
{guaiguaigol23, chiyp_besti}@163.com

² Key Laboratory of Network Assessment Technology, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China

Abstract. A new method to detect credential spear phishing attack for the network is introduced in the conference of 26th USENIX Security Symposium. First, on the basis of the researching for the processes and the principles of spear phishing attack, and the overall structure of its detector, the Directed Anomaly Scoring technology is analyzed in the paper. Second, the selections of scalars in subdetectors are defined. Third, the spear phishing attack detection method of detector and the methods of traditional detection are compared and analyzed. And then, the obvious advantages of the detector are discussed. The prospection of the spear phishing attack detection development is also given at the end of the paper.

Keywords: Spear phishing · Phishing detecting · Credential · DAS

1 Introduction

Phishing is a frequently used method of attack. As of September 2017, Anti-Phishing Alliance of China (APAC) has identified and punished 402239 phishing websites [1]. In recent years, with the intensification of system defense and users' awareness of security, the phishing attack methods have been constantly emerging, among which spear phishing has become a new and highly targeted one. Statistics show that in a large number of data leakage incidents and social engineering events, the number of spear phishing attacks accounts for about 92% [2].

In August 2017, several security researchers in the University of California at Berkeley and Lawrence Berkeley National Laboratory (LBNL) proposed a method of detecting credential spear phishing attacks in the enterprise environment [3], which analyzed the characteristics and two key stages of the spear phishing attack process. They designed a new group of reputation features, and introduced a new Directed Anomaly Scoring (DAS), which ran in a nonparametric way without any marked training data, and used reputation to detect attacks. The researchers collaborated with LBNL's security team to evaluate e-mail data (about 370 million e-mails) and related HTTP logs for nearly 4 years (from 2013 to 2017) to verify this method to be effective in detecting credential spear phishing attacks.

This paper consists of five parts. In the first part, the background of the new detection method is introduced which was put forward by the researchers at the 26th USENIX security conference in 2017 [3]. In the second part of the process, the forms of spear phishing attack and the classification of spear phishing attackers are analyzed profoundly. The third part focuses on the structure, the detection index and the algorithm. And then the comparison between this method and the traditional spear phishing detection methods is discussed in the next part. Finally the reference value of this paper for the research and application of the spear phishing attack detector is probed in the last part.

2 Analysis of Spear Phishing Attacks

Phishing is a form of “social engineering attack”. The attacker creates visual confusion by copying the target website (the definition of phishing attacks on the document [4]), and then using e-mail as a carrier to send the deceptive electronic mail from a legitimate enterprise or institution to users. The users reply to their personal identity data or financial account vouchers on the phishing websites, or download malicious software through links. In this way, the users’ personal information is leaked or trafficked or even generate more serious economic loss, which is reckoned as crime frauds.

2.1 Attacks Process

The spear phishing attack process, as shown in Fig. 1, is divided into two stages: the first stage is the lure stage (step 1–4 of Fig. 1). Attackers build a phishing infrastructure (such as phishing website), which imitates a legitimate website or hosts a phishing website on a highly visited website to make it visually confusing. Then they edit the e-mail of a phishing site link or attachment in a trusted identity and send baits luring users to visit the phishing website. The second stage is the exploit stage (Fig. 1 step 5–6). After obtaining the trust of the target user, attackers use this trust to induce the receiver to perform dangerous operations, such as to input personal information or to download malware. Then attackers use the privacy of the user to gain illegal interests from the online bank or the enterprise where the user registered.

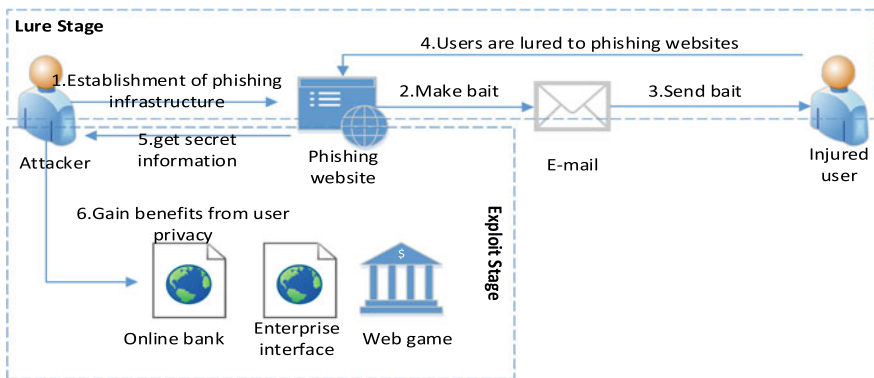


Fig. 1. Phishing attack process

2.2 Attacks Form

Casting net phishing attacks are not customized for attack target users but the spear phishing is different. Before the attack in the lure stage, attackers investigate and collect the target information then send deceiving e-mail to target users with customizing personalized bait, in order to let target users perform dangerous operations without doubt. The attacker intentionally disguises his attack mails as legal ones and the victims are more likely to be deceived, traditional reputation and spam filtering systems often fail to detect the malicious content contained in the mails, so it is difficult to prevent spear phishing [5].

The spear phishing attack has two forms. One is the malicious attachment attack which deceives users to download the malicious attachments in the e-mail, but the actual application of this form is rarely successful. The second is the voucher theft attack which induces the recipient to click on the link of the e-mail and enter personal information credentials on the generated web page. Among them this paper mainly focuses on the second form, because the voucher theft attack is likelier to perform than the vulnerabilities exploit attack. The detection method proposed by the UNSIX conference is also aimed at the second form of attack.

2.3 Lure Stage Attackers Classification

An e-mail mainly contains three parts: mail header, text, and attachment. Mail header consists of several predefined formatted fields, such as From, To, Subject and Message ID which includes the name and the address of the e-mail.

In the lure stage of the phishing attack, there are four types of attackers: the first type is the address spoofer who uses the e-mail address of the trusted individual as the "From" field to attack. The second is the name spoofer who impersonates the credible e-mail name instead of the e-mail address, and such attackers evade e-mail security mechanisms. The third type is previously unseen attacker who forges e-mail name and address looking like real name and real address at the same time. The fourth type is lateral attacker who sends phishing messages to other users from the trusted accounts which have been attacked by attackers. For the first type of attackers, e-mail security mechanisms such as Domain Keys Identified Mail (DKIM) and Domain-based Message Authentication Reporting and Conformance (DMARC) can handle it, so the latter three types of attackers are the focus of the research.

3 Analysis of the Detection Mechanism of the Detector

The detector is a new method to mitigate the risk of the spear phishing attack in the enterprise environment. Using network traffic log and machine learning, the alarm can be triggered in real time when the user clicks on the suspicious URL embedded in the e-mail [3].

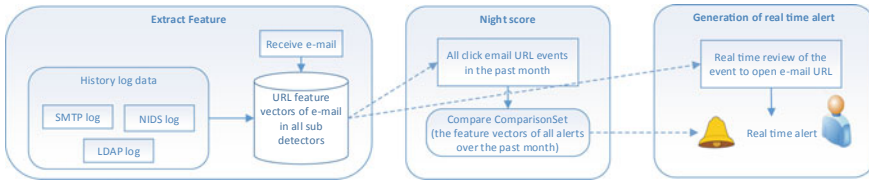


Fig. 2. General design of the detector

3.1 Overall Structure

The total design of the detector, as shown in Fig. 2, is divided into three parts: feature extraction, nightly scoring, and real-time alert generation.

The first part is feature extraction, including three sub-detectors. This part uses three kinds of logs from the company network to extract and save three feature vectors (FV) for each URL in e-mail, and each sub-detector has one FV. Network traffic logs include Simple Mail Transfer Protocol Logs (SMTP Logs), Network Intrusion Detection System Logs (NIDS Logs), and Lightweight Directory Access Protocol Logs (LDAP Logs). With network traffic logs, all clicks on URL in the e-mail messages are recorded, and the feature vectors extracted by the sub-detectors are used as the input of the nightly scoring and the real-time alert generation.

The second part is the nightly scoring. Every night, the events of clicking on URL in each sub-detector are collected and their FV through DAS are scored, and the most suspicious FV in this month is stored in the Comparison Set.

The third part is real-time alert generation. It can observe real-time network traffic, review events of the clicks on the e-mail URL, compare each sub detector's real-time click FV to the Comparison Set, and generate alert for the security team according to the need.

3.2 Reputation Characteristics

The attack targets of spear phishing are different from any regular phishing users but is aimed at the users who have some privileged access or capability. According to this selective target and motivation, the detector adopts a new classification method which characterizes the spear phishing attack in two dimensions which are also corresponding to the two key stages of the spear phishing attack.

The detector has set up two kinds of reputation characteristics for two key stages of phishing attack. The first kind is the domain reputation feature which describes the possibility of users accessing the URL according to the domain name. This feature can reflect the trust degree of the domain name and be used to capture the feature vector in exploit stage. The second kind is the sender reputation feature which describes whether the sender of the e-mail is belonging to the three types of the spear phishing attackers. This feature can reflect the trust degree of the sender used to capture the feature vectors in lure stage.

3.3 Selections of Scalar and DAS

As attackers use different attacks and the reputation features of the sender are different. The detector uses three sub-detectors based on three attack forms and each sub-detector extracts one feature vector containing four scalar values (two for the domain reputation, and two for the sender reputation) which are presented in Table 1.

With respect to domain reputation feature, if few users trust the linked domain name and access its URL in the experimental environment, the link is considered suspicious; if no employee has ever visited a domain name at present, the link will be considered dangerous.

Table 1. Sub detector extracts scalar

Reputation feature	Attack mode		
	Attacker of impersonate the name of e-mail	Unknown attacker	Lateral attacker
Domain reputation feature	C_{URL}	C_{URL}	C_{URL}
	T_d	T_d	T_d
Sender reputation feature	$D_{receiver}$	D_{name}	C_{user}
	C_{week}	$D_{address}$	C_{time}

Scalar C_{URL} comes from the statistics of the NIDS log statistic and calculates the number of URL which has the same fully qualified domain name (FQDN) as the clicking URL.

Scalar T_d , presented in Formula (1), calculates the difference between T_{email} which is the e-mail containing a domain name URL first arrived in the experimental environment and $T_{1stvisit}$ which is any users in the experimental environment for the first time to click on the domain name.

$$T_d = T_{email} - T_{1stvisit} \quad (1)$$

With respect to the sender reputation feature, address spoofed attacker impersonates an e-mail name of a trusted user but does not use the real e-mail address with this name, so the e-mail address of spear phishing e-mail sender does not match the address of any historical e-mail with this e-mail name. The name of the attacker is assumed to be associated with other trusted users, and the name used to send e-mail frequently to other users will be more reliable.

Scalar $D_{receive}$, shown in formula (2), calculates the difference between D_{past} which has the same name and address as the received e-mail and the current date $D_{current}$.

$$D_{receive} = D_{current} - D_{past} \quad (2)$$

Scalar C_{week} calculates the total number of weeks that meet the following conditions: using the name remained to be scored and sending e-mails to other users, sending at least one e-mail per working day in a week. C_{week} measures the credibility of the e-mail name field.

Previous unseen attacker chooses the name and address of an entity that resembles a known or authoritative entity, but the name and address does not match the actual real value of the entity, and the attacker will try to avoid the phishing attack detection. When an attacker uses deceptive identity, users who interact with him are likely to realize that the opponent uses forged names or e-mail addresses and reports. Therefore, the attacker will take a huge risk so they rarely use deceptive identity to avoid exposure.

Scalar D_{name} calculates the number of days between two e-mail messages which send using e-mail name remained to be scored.

Scalar $D_{address}$ calculates the number of days between the two e-mail messages which send using e-mail address remained to be scored.

Lateral attacker uses a controlled user's account to send a phishing mail to other users (not using any deception). When other users click on the link in the e-mail, if the e-mail is sent by the controlled account, the LDAP logs will be checked to confirm whether the account is using the IP address that has never been used before. If so, the sub-detector calculates the geographic location of the IP Address City, such as the city X, and then extracts two scalars of the city X.

Scalar C_{user} is the number of different users who log in from X city.

Scalar C_{time} is the previously logs number of the controlled user accounts with the same IP address in the city X.

The detector introduces the Directed Anomaly Scoring technology (DAS), which is used to automatically select the most suspicious events in the unmarked data set. The DAS scores the event according to the suspicion and then ranks the event according to the score. After all the events finished ranking, the DAS only chooses the most suspected (the highest ranking) N events, where N is the alert budget of the security team, that is, the number of acceptable alerts.

3.4 Application Situation

The spear phishing attack is rare. When American security researchers assessed the application of the detector, they used the collected 4 years enterprise data sets but only 10 cases of the known spear phishing. Extract feature vector by analyzing the network traffic logs of LBNL which includes the SMTP logs, NIDS logs, and LDAP logs. The SMTP logs record all e-mail messages sent by organization employees (including e-mails between two employees). The NIDS logs record information about HTTP GET and POST requests and include the complete URL of the access. The LDAP logs record the user's e-mail address, the login time and the IP address of the user's authentication. The researchers set the alarm budget as 10 which was evaluated from 370 million e-mail data sets of LBNL, and 17 of 19 phishing activities were identified successfully, accounting for 89% and the average false positive rate was 0.004% [3], which proved the effectiveness of the method.

4 Comparison and Analysis of Traditional Phishing Methods

Phishing detection technology extracts some characteristics of phishing attack behavior and uses these features to perform comparison and analysis. In this way it can identify phishing attack and achieve the prevention and defend of attack. From 2008 to 2016 there are few summary articles about the comparative analysis of the traditional detection methods for phishing attack. Although there have been a lot of researches and implementations about phishing detection technology, phishing websites have the characteristics of highly camouflage, strong timeliness, short viability, and wide attack targets [6]. What's more, the attackers constantly update the attack methods and strategies, and it is still unable to detect and defend all phishing attacks radically and effectively at the moment.

4.1 Analysis of Traditional Phishing Attack Detection Method

According to the content, phishing detection methods can be divided into three kinds: the analysis based on the communication approach, the analysis based on the website entrance and the analysis based on the content of the website. Communication channels include e-mail, telephone, social platform, SMS and so on, in which e-mail is the main channel of communication, referring to [7]. If the site entrance is URL, the attacker will make it similar to the common standard URL to aggravate the visual confusion. In the detection methods of phishing attack, the URL is analyzed frequently, but URL is not the only standard to judge a phishing website, so few people only analyze URL in the process of detection, and often use it in combination with web content analysis, referring to paper [8]; the content of the site includes the similarity of the web icon, the similarity of the Favicon, the similarity of the CSS architecture, the similarity of CSS architecture, the similarity of layout and homepage vision, and analysis of the underlying HTML of web pages, referring to paper [9].

According to the technical means, detections can be divided into blacklist-based phishing detection, heuristic phishing detection, and similarity-based phishing detection. The blacklist-based phishing detection constantly updates the blacklist generated by URL, IP address, domain name, certificate or keywords of known phishing, and then the phishing website is filtered according to the blacklist. For example, Liu et al. [10] probed the phishing website detection method based on the network spider and the LD algorithm. This method uses web spider which can grasp the link indefinitely and take the real URL link as the benchmark, then method uses the web spider to grab the similar URL as the blacklist, and hence analyzes the similarity of the content between these websites with the LD algorithm. Heuristic phishing detection uses machine learning. First, it extracts the features of a part of the phishing website as a training set, then uses classification or clustering machine learning algorithm to set threshold, then trains the machine to identify the features extracted by suspicious websites, in this way to judge whether it is phishing attack. Huang et al. [11] proposed technique of detecting unusual features of URL which extracts 12 features about results and vocabulary from the phishing website URL. These features are used as a training set. Then the machine is trained based on the characteristics of URL by SVM (Support Vector Machine) to determine whether the site is a phishing site. Phishing detection based on image

similarity is concerned with the visual similarity of pages or applications and is not concerned with the underlying code and features, such as the image similarity based phishing website detection method proposed by Lu Kang et al. [12]. It bases on the Selenium picture acquisition technology to obtain the protected pictures of official websites. It uses the perceptual hash algorithm to establish “Perceptual hash fingerprint library” in the website. When the new website appears, it can be judged by the comparison with the library to determine whether it is a phishing website.

4.2 Comparison of the Detector and Traditional Detection Method

When applying the traditional detection methods of phishing attack to detect the credential spear phishing attack, there are four limitations:

First, characters’ real distribution of the spear phishing attack is not determined, the selection of the parameters is blindness, and the manual setting of the features’ threshold will lead to attack missing. The phishing detection based on the blacklist is ephemeral. The phishing attacker will often change the IP address and URL. The blacklist needs to be updated in real time, but in actual application the time recorded in the blacklist time has been delayed. About 47–83% of the phishing can be added to blacklist after discovery of 12 h, while in fact 63% of phishing behavior will end in 2 h after the occurrence [13]. The delay greatly affects the accuracy of detection based on the blacklist, and it is unable to effectively defend the 0-h phishing attack and phishing detection based on the similarity.

Second, the spear phishing sample is relatively scarce and difficult to assemble into a large enough training set, and it is highly uneven in the classification of malicious samples and benign samples, which is not conducive to supervised learning. Heuristic phishing detection can effectively defend 0-h phishing attack and solve the limitation based on blacklist detection, but it needs a machine learning process which needs a large number of features as training set. There are few spear phish samples and unable to provide enough data. So using heuristic phishing detection will increase misjudging rate.

Third, in many security settings, scalar features usually possess numeric directionality. For example, the less numbers of access in a domain, the more suspicious it is, and the traditional detection technology unable to incorporate the concept of asymmetry or directionality into its calculation. In the application of heuristic phishing detection method, the selection of characteristics and the setting of threshold are mainly according to the statistical characteristics and artificial summaries of phishing, which is not objective enough. And some characteristics of the legitimate content conform to the set thresholds and result in misjudgment.

Fourth, even if only one or a few of the events’ characteristics are statistically abnormal, the traditional detection technology will regard the event as an attack, causing a waste of the alarm budget. Phishing detection based on image similarity sets the threshold of similarity as criterion only. This method assumes that the phishing website is similar to the legal website, which shows certain blindness. For the phishing websites which only partial copy the legitimate websites (less than 50%), the methods based on image similarity cannot detect successfully [14]. There may also be a strong

similarity between legitimate websites and other legitimate websites which can result in misjudgment.

The spear phishing detector presents obvious advantages in view of the above limitations. In real time, the real-time alert system can play the role of real-time monitoring. It does not need to form a training set of phishing websites in data, only the network traffic logs are needed. As for the extracting and analyzing on the features, three sub-detectors extract four scalars about the reputations of the receiver and the sender. The anomaly scoring standard is used at night which is objective and directional, and the alert budget can be formulated in order to reduce waste.

5 Conclusions

In this paper, based on the overview of the spear phishing attack, the overall structure, and key technology of the spear phishing detector are analyzed. Compared with the traditional detection methods of phishing attacks, the spear phishing detector shows obvious advantages, which puts forward the new reputations and scalars and introduces a new anomaly detection technique. The detector provides a practical approach to detect credential spear phishing attack, but the limitation in visibility, alert budget, ineffective coping strategies, and historical data is still worth exploring. These limitations may be broken through with future research and a more effective warning mechanism will be designed as a part of preventive defenses for spear phishing attacks.

Acknowledgements. This research was financially supported by the National Key Research and Development Plan (2018YFB1004101).

References

1. Anti-Phishing Alliance of China (APAC): [R/OL]. [2017-10]. <http://apac.cn/gzdt/201710/P020171012610151655807.pdf>
2. Huan, W.: Research on defense based on harpoon phishing attack. *Comput. Knowl. Technol.* **12**(8), 51–54 (2016)
3. Ho, G.: Detecting Credential Spear phishing Attacks in Enterprise Settings [EB/OL]. [2017-8-16]. <https://www.usenix.org/system/files/conference/usenixsecurity17/sec17-ho.pdf>
4. Sha, H., Liu, Q., Liu, T., et al.: Survey on malicious webpage detection research. *J. Comput.* **39**(03), 529–542 (2016)
5. Akinyelu, A.A., Adewumi, A.O.: Classification of phishing email using random forest machine learning technique. *J. Appl. Math.* (2014-4-3), **2014**(1), 89–108 (2014)
6. Anti-Phishing Working Group (APWG): Global phishing survey: trends and domain name use in 2H2014 [EB/OL]. [2014-5-27]. http://apwg.org/download/document/245/APWG_Global_Phishing_Report_2H_2014.pdf
7. Hongfu, H.U., Peng, G.: Mechanism of phishing email detection based on user interaction and its realization. *Comput. Eng. Appl.* **53**(15), 7–13 (2017)
8. Ramesh, G., Krishnamurthi, I., Kumar, K.S.S.: An efficacious method for detecting phishing webpages through target domain identification. *Decis. Support Syst.* **61**(5), 12–22 (2014)

9. Yan, Z., Liu, S., Wang, T., et al.: A genetic algorithm based model for chinese phishing e-commerce websites detection. In: *The International Conference on HCI in Business, Government, and Organizations: Commerce and Innovation*, pp. 270–279. Springer International Publishing (2016)
10. Liu, S., Peng, H., Lin, H., et al.: Detection technology of phishing web based on network spider and LD algorithm. *Autom. Instrum.* **188**(6), 165–166 (2015)
11. Huang, H., Qian, L., Wang, Y.: Detection of phishing URL based on abnormal feature. *Inf. Netw. Secur.* **2012**(1), 23–25 (2012)
12. Lu, K., Zhou, A.: Phishing-website detection based on image similarity. *Inf. Secur. Commun. Priv.* **54**(3), 115–117 (2016)
13. Xu, H., Xu, H., Lei, L.: Phishing recognition technology based on fusion of multiple features classification and recognition algorithm. *Appl. Res. Comput.* **34**(04), 1129–1132 (2017)
14. Guan, Y., Zou, F., Yi, P.: Web phishing detection algorithm based on graph mining. *Microcomput. Appl.* **32**(7), 1–5 (2016)



Memory Confidentiality and Integrity Protection Technology

Hongjin Wang¹, Huimin Meng^{1(✉)}, Nianmin Yao²,
and Yishun Cheng²

¹ School of Information Science and Engineering, Dalian Polytechnic University, Dalian 116034, China

Wanghongjin201@163.com, menghm@dlnpu.edu.cn

² School of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China

lucos@dlut.edu.cn

Abstract. In most existing computer systems, data transmission and preservation in the form of plaintext are vulnerable to various attacks. In this paper, we use Parallelized memory Confidentiality and Integrity Protection technology (PCIP) algorithm to ensure the confidentiality and integrity of memory data. On the basis of PCIP, we use PCIP Bonsai Merkle Tree (PCIP+BMT) to protect the counter values of off-chip to reduce system delay and overhead. PCIP is that uses counter mode encryption to encrypt data while adding redundant data for integrity checking. Finally, we use the SimpleScalar Tool to simulate the PCIP and PE-ICE algorithms. The results show that PCIP is encrypted more effectively than the PC-ICE. Compared with the Hash algorithm, it can reduce the system delay and reduce the internal memory overhead. The tree mechanism adopted in this paper reduces the impact on system performance.

Keywords: PCIP · Confidentiality · Integrity · Counter mode encryption · System delay · Overhead

1 Introduction

With the rapid growth of digital information stored in modern computer systems, information security has been paid more and more attention. The phenomenon of commercial software being cracked and confidential information being leaked are increasingly emerging. Protecting the confidentiality and integrity of data and preventing attackers from stealing and modifying system critical data has become an important direction in the current security field [1].

For memory security problems, it is mainly to prevent spoofing attacks, splicing attacks, and replay attacks [2]. Best et al. proposed bus encryption technique in 1979 to protect the confidentiality of data on the system bus, which was mainly to solve the problem of digital copyright protection. In 1984, Tim Maude et al. proposed hardware protection mechanisms to solve software copyright issues, but the method was costly [2, 3]. The protection of memory confidentiality was mainly achieved by encrypting the contents of memory. Earlier research methods were carried out by direct memory

encryption, such as Direct Block Encryption adopted by eExecute Only Memory (XOM) and AEGIS [4, 5]. The Direct Block Encryption was easy to operate and implement. However, during a read operation, if the cache missed, it would increase the system delay greatly. Jun Yang et al. proposed One-Time Password (OTP) Encryption in 2003, this method could reduce the system delay, but it must maintain the non-tampering counter that corresponded to the data for encryption algorithm, so which increased the maintenance overhead [6, 7]. In terms of memory integrity protection, both of XOM and AEGIS architecture were using Cryptographic Hash function, Message Authentication Code (MAC), and Added Redundancy Explicit Authentication technology [8–10]. To prevent replay attacks, Brian Rogers et al. introduced Hash Tree into AEGIS architecture, but it needs to maintain a tree that occupied large space, moreover, it is necessary to calculate the hash value of each layer of the tree when performing integrity check, therefore, the resulting system delay is greater [11, 12].

In this paper, the PCIP is to use counter encryption to reduce system delay and improve system operation efficiency. At the same time, for the protection of the data and counters, we adopt PCIP+BMT which integrate integrity check into encryption operation to improve the efficiency of integrity check. Finally, we use the SimpleScalar Tool to simulate this method. The results show that PCIP+BMT protection mechanisms can reduce system delay and internal storage overhead, and protect the confidentiality and integrity of memory.

2 Parallelized Encryption and Integrity Checking Engine

PE-ICE is an encryption and integrity check engine proposed in [13]. This method is using AREA technology and the diffusion characteristic of block encryption algorithm to integrate the integrity authentication into the cryptographic operations.

PE-ICE adds the verification flag T generated by the counter in the CPU to the plaintext to perform encryption after the diffusion and then writes the ciphertext into RAM. When CPU performed the read operation, the ciphertext is decrypted and extract the corresponding T' from the decrypted data, then compared T' with T . if they are the same, the authentication is successful. Otherwise, the data is tampered with. PE-ICE encryption is a Direct Block Encryption method that data encryption and integrity check can be performed in parallel. But it increases system delay greatly compared to GCM mode counter encryption.

3 PCIP Technology

Based on PE-ICE, this paper proposes a counter encryption PCIP method and adds redundancy data for integrity verification after decryption. PCIP will diffuse the data block before data encryption to ensure that all attacks could be detected. Each state bit in the data block after diffusion will depend on all the state bits before the diffusion operation, and changing one bit of data may affect half of the state bits after diffusion [14].

3.1 Generate Verification Data

Because the read-only data is only affected by Spoofing attacks and Splicing attacks, PCIP can use the corresponding address as the Tag to form the plaintext. For the read-write data, it is not only affected by Spoofing attacks and Splicing attacks, but also sensitive to Replay attacks. PCIP use two methods to generate verification data. One of the ways is to generate Tag by counter plus 1. But the Tag can only be used once, and the CPU performs a write operation one time, the counter plus 1. When a counter overflows, all related memory areas must be re-encrypted. The other way is to add a random number RV after the corresponding address of the data to form a Tag, which can prevent Spoofing attacks, Splicing attacks and Replay attacks.

3.2 Data Diffusion and Encryption Operation

In order to ensure all attacks can be detected during the integrity check, the ciphertext needs to achieve full diffusion. The data blocks need to be diffused before data encryption. Regard each block as a two-dimensional array of 4×4 , and every Byte is considered as an element in $GF(2^8)$. Then the full data diffusion is achieved through state column mixing and Byte transposition [15].

First, regard every column in 4×4 matrix as a cubic polynomial of $GF(2^8)$, then define two random polynomials $a(x)$ and $b(x)$ of variable coefficients, and a polynomial $c(x)$ of constant coefficients.

$$\begin{cases} a(x) = a_0 + a_1x + a_2x^2 + a_3x^3 \\ b(x) = b_0 + b_1x + b_2x^2 + b_3x^3 \\ c(x) = c_0 + c_1x + c_2x^2 + c_3x^3 \end{cases} \quad (1)$$

Then, multiply $a(x)$ with $c(x)$ in $\text{mod}(x^4 + 1)$ to get $b(x)$.

$$\begin{aligned} b(x) &= a(x) \cdot c(x) \\ &= (a_0 + a_1x + a_2x^2 + a_3x^3) * (c_0 + c_1x + c_2x^2 + c_3x^3) \\ &\equiv (b_0 + b_1x + b_2x^2 + b_3x^3) \text{mod}(x^4 + 1) \end{aligned} \quad (2)$$

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} c_0 & c_3 & c_2 & c_1 \\ c_1 & c_0 & c_3 & c_2 \\ c_2 & c_1 & c_0 & c_3 \\ c_3 & c_2 & c_1 & c_0 \end{bmatrix} \times \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} \quad (3)$$

According to the Rijndael algorithm [16]. Define $c(x)$ as Eq. (4). Then each column in the matrix is multiplied by a fixed polynomial $d(x)$. Each column element in matrix b is related to each column element in the original matrix.

$$c(x) = 02 + 01 \cdot x + 01 \cdot x^2 + 03 \cdot x^3 \quad (4)$$

$$(02 + 01 \cdot x + 01 \cdot x^2 + 03 \cdot x^3) \cdot d(x) \equiv 1 \pmod{(x^4 + 1)} \quad (5)$$

$$d(x) = 0E + 09 \cdot x + 0D \cdot x^2 + 0B \cdot x^3 \quad (6)$$

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 0E & 0B & 0D & 09 \\ 09 & 0E & 0B & 0D \\ 0D & 09 & 0E & 0B \\ 0B & 0D & 09 & 0E \end{bmatrix} \times \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} \quad (7)$$

Finally, cyclically shift the rows in the matrix by using different offsets to achieve the effect of the full diffusion of the data, so that all the attacks can be detected when the data is checked for integrity.

After the data block reaches full diffusion effect, the intermediate data needs to be encrypted and stored in the RAM. PCIP uses a counter plus 1 to generate the corresponding seed for each data to form the corresponding seed, and then encrypts the key using the AES algorithm to generate the key, and at the same time, the combined data and the corresponding check value generate the plaintext data blocks, and get the intermediate data after diffusion. Finally, the key and the intermediate data are XORed to generate a ciphertext.

3.3 Read Operation and Write Operation

When the CPU performs writing operation. Firstly, accord the data block's CTR and address to generate Tag. Then combine the Payload data and Tag to form the Plaintext Block, and diffuse the Plaintext Block. Finally, the diffused plaintext and the encrypted CTR are XORed to generate a Ciphertext Block, and write it on-chip.

When the CPU performs reading operation. First, read the Ciphertext Block and the corresponding CTR from external memory. Then decrypted CTR and Ciphertext are XORed to get intermediate data, and inverse operations on the diffusion of intermediate data to restore the Plaintext Block. Finally, extract Tag from the Plaintext Block for integrity check. If it is equal to the Tag before encryption, the corresponding valid data is read into the CPU. Otherwise, the integrity check fails or the data may be tampered.

3.4 The Structure of PCIP+BMT

In the memory protection process, each data block has a corresponding CTR, so a large number of CTRs need to be preserved. Due to limited internal storage space, these CTRs need to be stored off-chip. This paper proposes PCIP+BMT to protect these CTRs. PCIP+BMT use the MAC function to compute corresponding MAC for these CTRs, apply Hash Tree to prevent Replay attacks. Since the ciphertext block and the corresponding CTR are secure in RAM, the decrypted CTR is also trusted, which only needs to consider the Replay attacks. Therefore, we can build a Hash Tree in CTRs to prevent such attacks. First, use the Hash Algorithm to calculate the MAC for blocks composed of multiple CTRs, and multiple MACs to form a node in the tree. Then continue to calculate the MAC of these MAC nodes until generated a root node, and stored the root node in internal memory. The tree structure is shown in Fig. 1.

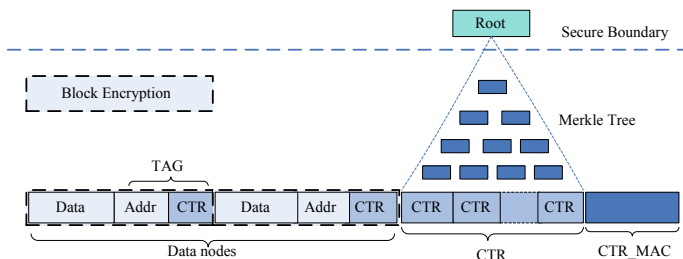


Fig. 1. The tree of protect CTRs

On the one hand, PCIP+BMT build the Merkle Tree only for CTR, compared to the traditional method of constructing a Merkle Tree for the entire memory area, the coverage is reduced greatly, and the verification path from the leaf node to the root node is shortened significantly. On the other hand, because Incremental Hash Function has incremental updates, it can be used to calculate the MAC of these tree nodes, enlarge the size of the block, reducing the tree nodes and the height of the tree, thus further improving the performance.

4 Simulation Experiment and Result Analysis

In this paper, we use SimpleScalar to simulate the proposal. In the experiment, 10 SPEC2000 benchmark programs (ammp, art, bzip2, equake, gzip, mcf, mesa, parser, vortex, vpr) are used to assess the performance of the proposed program [17–19]. The size of Cache data block is 64 Bytes and 128 Bytes, respectively, and the capacity of L2Cache capacity is 256 KB and 1 MB, respectively, and statistics the Instruction-Per-Cycle (IPC) of each test program [20, 21].

The experiment assumes that the Tag has been stored on-chip. When the data is read into the processor for decryption, it can be checked for integrity immediately. The PCIP proposed in this paper is compared with the non-protection method, the method of OTP encryption, and the method of PE-ICE encryption. The results are shown in Fig. 2.

Figure 2 shows the effects of various algorithms on program performance under different L2 cache capacities. And different Cache data block sizes have different effects on different test programs. When the capacity of L2 cache is small, PCIP has greater advantages than PE-ICE on program performance. The larger the cache capacity, the higher the Cache hit rate and the higher the corresponding IPC, which makes the impact of the PCIP encryption method on the performance of the program gradually reduced according to $slowdown = 1 - x/baseline$.

PCIP+BMT is simulated under different L2 cache capacities and compared with PCIP without tree protection mechanism. The results are shown in Fig. 3.

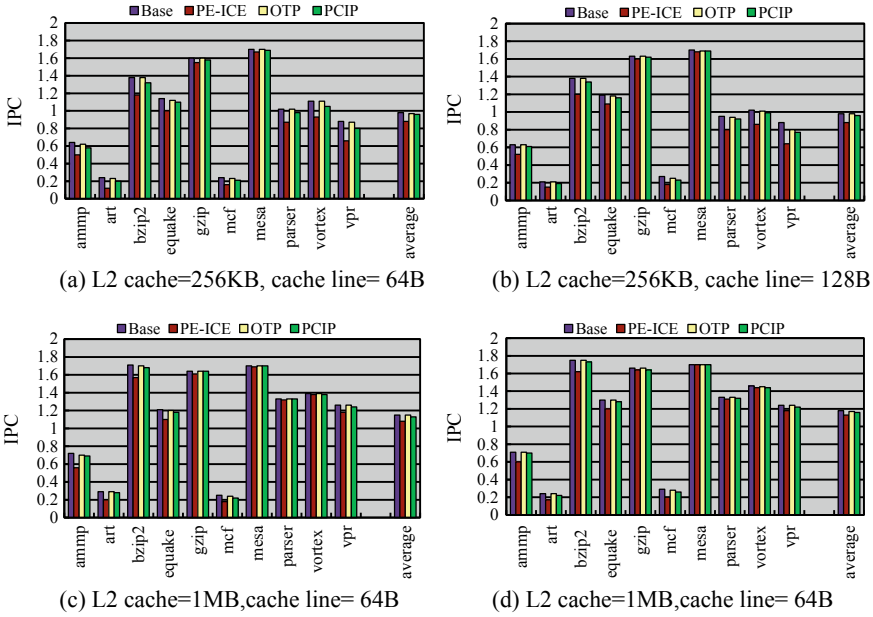


Fig. 2. The IPC of different encryption schemes

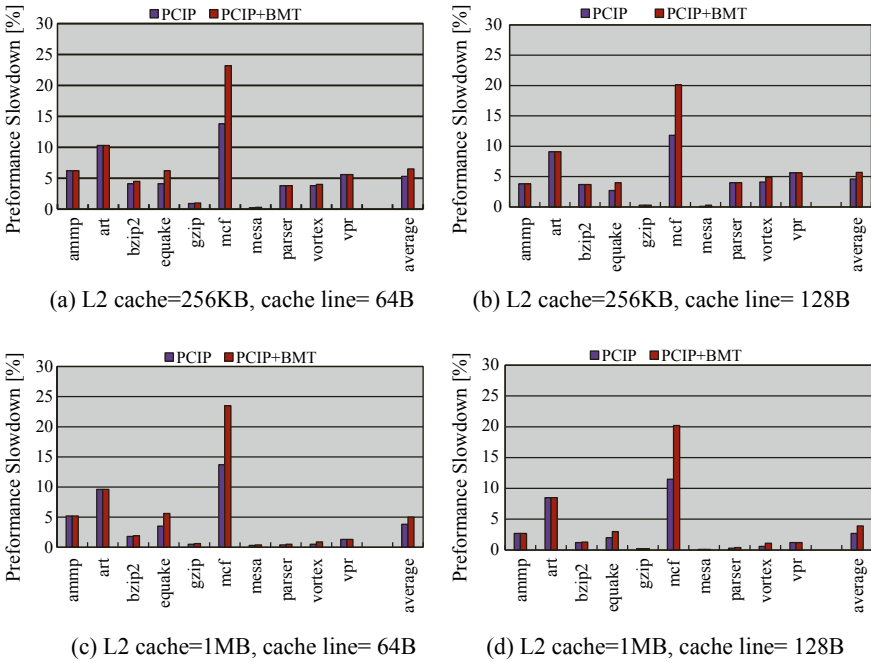


Fig. 3. The effect of integrity protection scheme on program performance

Due to the high hit rate of CTR, PCIP+BMT generate less delay when the CTR cache misses. Some programs performance degradation after adopting the tree protection mechanism, such as mcf and equake, it is due to the large cache miss rate, so the overhead of PCIP+BMT performing integrity checks increases rapidly.

5 Conclusion

For the memory security problem, this article introduces the related technologies of confidentiality protection and integrity protection of memory. In view of their shortcomings, we propose PCIP. Diffusing the data and its Tag, and encrypt it with counter mode. Compared with the PE-ICE, the encryption performance has been improved greatly. For example, the L2cache capacity is 256 KB, and the cache line capacity is 128 B. The average influence of PCIP encryption on program performance has decreased from 15.56% of PE-ICE to 4.54%. Then to prevent Replay attacks, PCIP+BMT are used to protect the CTR that stored off-chip, which also reduces the system delay and storage overhead.

References

1. Chen, X., Shang, T., Kim, I., et al.: A remote data integrity checking scheme for big data storage. In: IEEE Second International Conference on Data Science in Cyberspace, pp. 53–59. IEEE (2017)
2. Patrick, C., Jiawen, Z., James, G., et al.: Protecting data on smartphones and tablets from memory attacks. *Twent. Int. Conf. Archit. Supp. Progr. Lang. Oper. Syst.* **50**(4), 177–189 (2015)
3. Rejani, R., Murugan, D., Krishnan, V.: Digital data protection using steganography. *ICTACT J. Commun. Technol.* **7**(1), 1245–1254 (2016)
4. Abdellatif, K.M., Chotin, A.R., Mehrez, H.: AES-GCM and AEGIS: efficient and high speed hardware implementations. *J. Sig. Process. Syst.* **88**(1), 1–12 (2017)
5. Ueno, R., Morioka, S., Homma, N., et al.: A high throughput/gate AES hardware architecture by compressing encryption and decryption datapaths. In: International Conference on Cryptographic Hardware and Embedded Systems, pp. 538–558. Springer, Berlin, Heidelberg (2016)
6. Wen, W., Jun, Y., Youtao, Z.: Optimizing power efficiency for 3D stacked GPU-in-memory architecture. *Microprocess. Microsyst.* **49**, 44–53 (2017)
7. Mohan, P., Chelliah, S., Li, Y.: An authentication technique for accessing de-duplicated data from private cloud using one time password. *Int. J. Inf. Secur. Priv.* **11**(2), 1–10 (2017)
8. Rasjid, Z.E., Soewito, B., Witjaksono, G., et al.: A review of collisions in cryptographic hash function used in digital forensic tools. *Procedia Comput. Sci.* **116**, 381–392 (2017)
9. Elbaz, R., Torres, L., Sassatelli, G., et al.: Block-level added redundancy explicit authentication for parallelized encryption and integrity checking of processor-memory transactions. *Trans. Comp. Sci. X* **10**, 231–260 (2011)
10. Chen, C.: QR code authentication with embedded message authentication code. *Mob. Netw. Appl.* **22**(3), 1–12 (2016)
11. Suh, G.E., Fletcher, C., Clarke, D., et al.: Author retrospective AEGIS. In: Anniversary International Conference, pp. 8–70 (2014)

12. Garg, N., Bawa, S.: RITS-MHT: relative indexed and time stamped Merkle hash tree based data auditing protocol for cloud computing. *J. Netw. Comput. Appl.* **84**, 1–13 (2017)
13. Elbaz, R., Torres, L., Sassatelli, G., et al.: A parallelized way to provide data encryption and integrity checking on a processor-memory bus. In: *Design Automation Conference*, pp. 506–509. ACM (2006)
14. Mohanraj, V., Srinivasan, S.: Development of advanced encryption standard architecture with sbox parity. *Walailak J. Sci. Technol.* **14**(4), 315–325 (2017)
15. Kermani, M.M., Jalali, A., Azarderakhsh, R., et al.: Reliable inversion in $GF(2^8)$ with redundant arithmetic for secure error detection of cryptographic architectures. *IEEE Trans. Comput.-Aided Des. Integr. Circ. Syst.* **PP**(99), 1–1 (2017)
16. Sharma, N.A., Farik, M.: A performance test on symmetric encryption algorithms—RC2 Vs Rijndael. *Int. J. Sci. Technol. Res.* **6**(7), 292–294 (2017)
17. SimpleScalar. <http://www.simplescalar.com/>
18. Arora, H., Bhatia, U., Pandita, S.: Comparative study of different cache models on SimpleScalar architecture. In: *Communication and Computing*, pp. 302–304. IET (2015)
19. SPEC 2000. <http://www.spec.org/cpu2000/>
20. Uddin, I.: One-IPC high-level simulation of microthreaded many-core architectures. *J. Syst. Architect.* **60**(7), 529–552 (2014)
21. Zhang, W., Ji, X., Lu, Y., et al.: Prophet: a parallel instruction-oriented many-core simulator. *IEEE Trans. Parall. Distrib. Syst.* **PP**(99), 1–1 (2017)



Privacy Preservation of Semi-structured Data Based on XML

Cheng Shi^{1(✉)}, Mingda Yang^{2(✉)}, and Bo Ning^{1(✉)}

¹ Dalian Maritime University, Dalian 116026, China
shicheng2015@126.com, ningbo@dlmu.edu.cn

² Cornell University, Ithaca, USA
my432@cornell.edu

Abstract. In the information age, people's various behavioral data are collected in large quantities. The sharing of information makes it convenient for some scientific investigations, but there is a leakage of personal privacy at the same time. The current research on privacy preservation is mostly based on relational tables or social network graphs. This paper focuses on semi-structured data, which is often ignored in privacy preservation. We propose a new privacy guarantee called $X-k^m$ -anonymity and propose a bottom-up heuristic algorithm that provides protection by satisfying $X-k^m$ -anonymity. We verified the feasibility of the algorithm through a reliable utility analysis method on the simulation data.

Keywords: Privacy preservation · Semi-structured data · $X-k^m$ -anonymity · XML

1 Introduction

With the advent of the era of big data informatization, personal information is collected in large numbers by various organizations and organizations, and information is published and shared. The sharing of information makes it convenient for some scientific research investigations, but at the same time, there is a disclosure of personal privacy. Although information such as name and the ID number is concealed when the information is released, this does not completely protect personal privacy information. Attackers still can reidentify personal identification by linking multiple sets of public information [1]. Therefore, the issue of privacy protection has attracted the attention of more and more researchers.

At present, the research on privacy protection is mostly based on two-dimensional relational tables or social network diagrams, and the intermediate situation is often ignored. The existing research on privacy protection based on semi-structured data is very scarce, and existing methods are not well-suited to XML document privacy protection. Paper [2] uses generalization and structure disassociation methods to protect the semi-structured data, but structural disassociation, if applied to XML, will destroy

the DTD structure of the document itself. Paper [3] reconstructs the DTD for the privacy document and reassembles the quasi-identifier information and sensitive information in groups. This method is less versatile and cannot be applied to all forms of XML, and the user's personal information will no longer be accurate in private document, so it has a big impact on the availability of data.

Our main contributions include the following:

- (1) We define the problem of anonymizing semi-structured data and we define the data structure for anonymous documents.
- (2) We define the X - k^m -anonymity privacy guarantee for XML data and explain how it is efficient in concrete attack scenarios.
- (3) We propose a novel anonymization algorithm and an information loss metric that takes document similarity into account.
- (4) We evaluated the proposed anonymization method through experiments and proved that it can provide anonymous datasets with limited information loss.

2 Problem Definition

2.1 Attack Model Against XML Data

We consider an XML document D that is composed of the root and several elements, each element E contains text value V and attribute S , denoted as $D = (root, E)$, where $E = (V, S)$. The text values take values from a domain F , the same tags represent a class G , each class G has a domain F_G , the domain of each class does not intersect, and the union of all class domains is F .

We consider attackers who have the information of m elements, denoted as set E_A , and they tried to use this information to reidentify the entire record. We assume that an attacker has only positive knowledge and they do not know the structural relations between the elements in E_A .

2.2 Privacy Guarantee

We know the k^m -anonymity [4] is a kind of data protection technology extended on the basis of k -anonymity [1], it specifies the maximum information leakage risk that the user can afford by specifying the parameter k and m . In XML, to satisfy k^m -anonymity, it requires that there are at least k combinations of any m nodes in the anonymous document, and depending on the number of nodes and the distribution of nodes, these combinations may be a single node or they may be path or twig, so we call this k^m -anonymity on XML X - k^m -anonymity. The X - k^m -anonymity protects individual privacy to a certain degree, but at the same time, it reduces the availability and accuracy of data. Therefore, our work mainly focuses on protecting private information while improving the availability and accuracy of data.

3 Anonymization Operations

A XML document D can be transformed to a private XML document D' which complies to X - k^m -anonymity by a series of transformations. There are two steps, first we should find out which nodes or twigs are easily identified, and second, we replace rare values with a more common one which can meet the privacy guarantee.

3.1 Sensitive Query

Definition: Execute the sensitive query to the original XML document D , the corresponding result obtained by each query pattern is called an instance of the pattern. If the number of instances of a query pattern is less than k , such an instance is called sensitive fragment, denoted as l .

In the process of querying sensitive fragments, we mainly use the classic StackTree [5], PathStack and TwigStack [6] algorithms to get fragments that do not meet X - k^m -anonymity. After the sensitive query, all the sensitive fragments are stored in the set L , where $L = \{ l_1, l_2, \dots, l_n \}$ ($n \geq 0$). The values of all sensitive fragments constitute the domain Γ of L .

3.2 Generalization

We use generalization [7, 8] method to protect nodes or twigs from privacy breaches. First, we build the data generalization hierarchy (DGH) for each class of tag, the concept of elements gradually turns general from leaf to root. Let us see an example shown in Fig. 1, here are the DGH of element *gender* and *disease*.

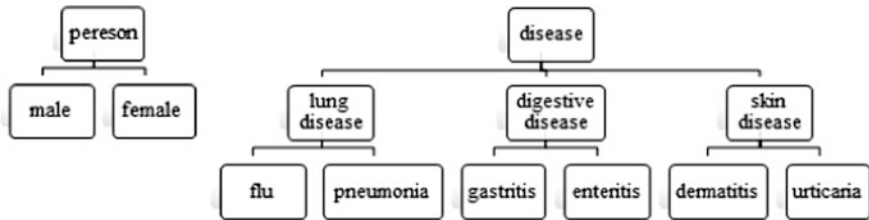


Fig. 1. The DGH for element *gender* and *disease*

In generalization, when a value is generalized, then all its appearances in the document are also generalized to the same value. After the document has performed all the generalization operations, check whether there are sibling nodes with the same value in the document. If there are such sibling nodes, check whether their subtrees are the same. If they are the same, merge the sibling nodes and their subtrees, if they are different, remain as it is.

4 Utility Analysis

Data anonymization must be accompanied by information loss, which has an important impact on the content and query accuracy of the document. This paper will measure the loss caused by anonymization from the similarity of the document [9]. In the latter algorithm, we will use this method as criteria to get the optimal solution, so that we can further optimize our algorithm by minimizing information loss and increasing accuracy.

We use cosine similarity calculation method to measure the similarity of text content. We represent the text vector of the original XML document D as $D(d_1, d_2, \dots, d_i, \dots, d_n)$ and the anonymous XML document D' using the vectors $D'(d_1', d_2', \dots, d_i', \dots, d_n')$, where d_i or d_i'' represent the weights of text value of the corresponding element in D or D' . The weight of text value is calculated as: $\text{weight}(v) = \frac{\text{depth}(v)}{\text{depth}(DGH)}$, where $\text{depth}(v)$ is the depth of v in the DGH and $\text{depth}(DGH)$ is the depth of DGH. According to the cosine similarity calculation, the text similarity calculation formula between D and D' can be shown as follows:

$$\text{contSim}(D, D') = \frac{\sum_{i=1}^n d_i \times d_i'}{\sqrt{\sum_{i=1}^n d_i^2} \times \sqrt{\sum_{i=1}^n d_i'^2}} \quad (1)$$

when the document length is long enough, the text vector of the document is too large and it is not convenient to calculate. Thus, we can split the document according to certain rules, and then construct the text vector separately. Therefore, we improve the above formula as follows:

$$\text{AVGContSim}(D, D') = \frac{\sum_{s=1}^a \text{contSim}(D_s, D'_s)}{a} \quad (2)$$

where a is the number of parts the document is split. The formula adds the cosine similarity of each part and divides it by the total part number to obtain the average cosine similarity of the overall document.

5 Anonymization Algorithm

We use a heuristic algorithm to find the optimal generalization solution that enables the document to satisfy $X-k^m$ -anonymity. The algorithm all plan search (APS) describes the entire process for finding the optimal solution. To facilitate the understanding of the algorithm, we give an example to explain the basic idea and the process of the algorithm. The algorithm first executes a sensitive query on the original XML document. After the sensitive query, if the text values contained in the query result that does not satisfy $X-k^m$ -anonymity include *male*, *female*, *pneumonia*, *flu*, *gastritis*, *enteritis*, *dermatitis*, and *urticaria*, then we need to combine the DGH of *gender* and *disease* to form an overall DGH with the root node being “*” and representing “any”.

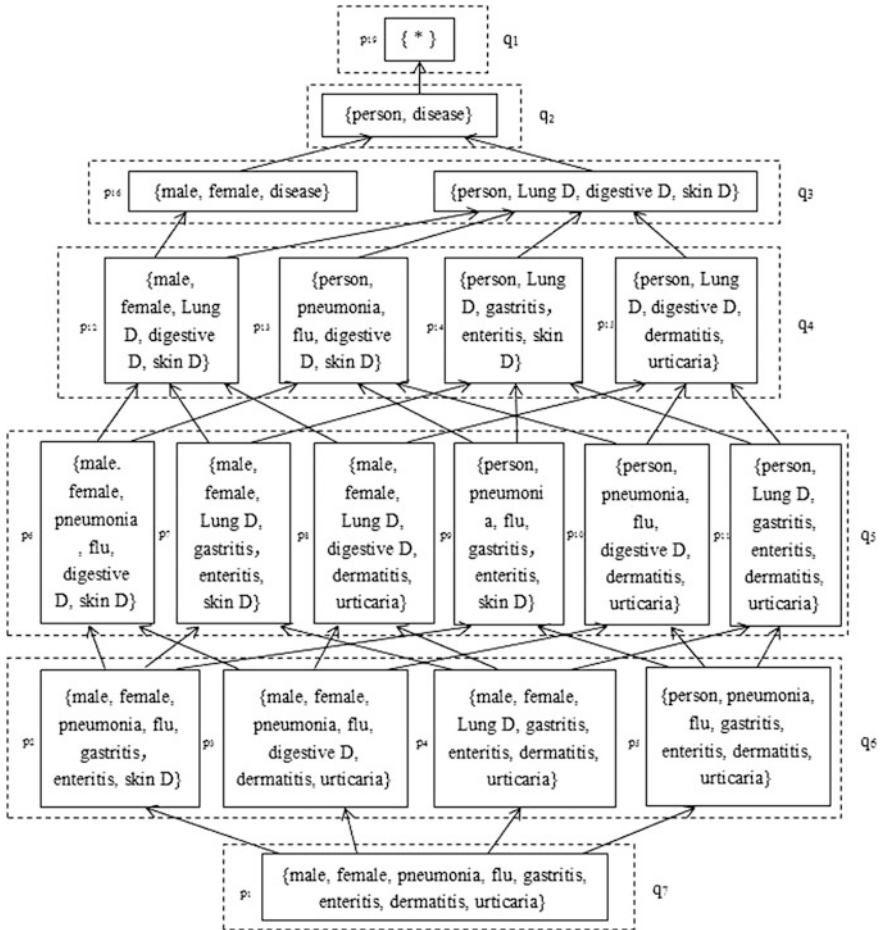


Fig. 2. The Plan-enumeration Tree for the DGH of Fig. 1

According to the DGH in Fig. 1, we can construct the plan-enumeration tree (PET) as shown in Fig. 2. The figure lists all possible generalization plans. Each node represents a generalization plan. The directed edges between nodes show how one generalization plan generates another generalization plan by generalizing one of the elements. The algorithm APS first starts with the most special generalization plan p_1 , i.e., the initial text values $\{male, female, pneumonia, flu, gastritis, enteritis, dermatitis, urticaria\}$. The algorithm only traverses all the plans at most once. The pseudocode of the APS algorithm is displayed by Algorithm 1.

Algorithm 1. AllPlanSearch(APS) Algorithm

Input: D , DGH, k , m **Output:** $BestPlan$, D'

```

1:  $L = \text{SensiQuery}(k, m)$ ;
2: create the generalization hierarchy of  $L$ ;
3: create the PET;
4: for  $q = i-1, i-2, \dots, 1$  do
5:   for all plans of  $PET_q$  do
6:     if the plan need to check then
7:       if the plan satisfy  $X$ - $k^m$ -anonymity then
8:         add this plan into  $P$ ;
9:   if  $P \neq \text{NULL}$  then
10:    break;
11: calculate the costs of all plans in  $P$ ;
12:  $BestPlan =$  the plan with the lowest cost;
13: update the raw XML with  $BestPlan$ ;
14: return ( $BestPlan$ ,  $D'$ );

```

The algorithm first performs a sensitive query on the original document, puts the query result that does not satisfy X - k^m -anonymity into the set L , then it creates DGH for all the elements contained in the set L , and generates the PET according to the DGH. Consider the PET of Fig. 2, APS would first start all values to $\{male, female, pneumonia, flu, gastritis, enteritis, dermatitis, urticaria\}$, as shown in plan p_1 of Fig. 2. Then APS proceeds to the last layer, there are four plans that can be generalized: p_2, p_3, p_4 , and p_5 . If there is only one record has *urticarial*, then p_3, p_4, p_5 have no need to be checked. After checking for X - k^m -anonymity, the plans which satisfy X - k^m -anonymity are put into set P . If there's no plan that satisfy X - k^m -anonymity, then APS proceeds to the last layer, and repeat the above process until all values are generalized to $p_{19} \{*\}$. Then calculate the costs of all the plans in P , and pick the plan with the lowest cost as described in the previous chapter. Finally, use $BestPlan$ to update the original document, then output $BestPlan$ and new documents, and APS algorithm terminates.

When the data is more intensive, that is, fewer layers of data are generalized, so that the bottom-up generalization method is more efficient than the top-down generalization method. Since we want to keep the usability of data while protecting data, our method can meet this claim greater.

6 Experiment

In the section, we present the experimental evaluation of our algorithms. All implementations were performed on an Intel Core i5 CPU, with 4 GB RAM, running Windows 10.

We implement the algorithm code on the simulation data, and the XML document is composed of medical records of different individuals, the elements are *age*, *gender*, *address*, *department_id*, *date*, *disease*, and *medicine*. Experiments are performed on datasets size of 50 and 500, respectively. Because of the small quantity of research on privacy protection in XML, there is no similar algorithm to this article for comparison.

From Fig. 3, we can see that as m increases, AVGContSim gradually decreases, and when the value of m starts to increase from 2, the similarity between D and D' is almost the same no matter what k is. Therefore, when m takes a value of 2 or larger, we can set k as large as possible so that the probability of a single record being reidentified is greatly reduced.

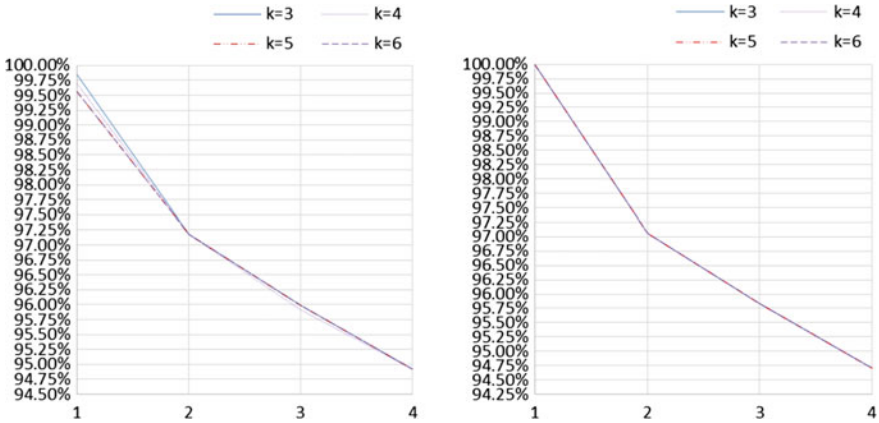


Fig. 3. AVGContSim with $|D| = 50$ and with $|D| = 500$: effect of m

7 Conclusion

In this article, we are working on the privacy protection of semi-structured data. We define X - k^m -anonymous privacy assurance which involves values and XML structures. We propose an all plan search algorithm to create X - k^m -anonymous datasets, by using value generalization and data transformation. The privacy protection methods presented in this paper can be applied to any schema of XML datasets and larger datasets.

Acknowledgements. This work is supported by the National Natural Science Foundation of China (U1401256), the National Natural Science Foundation of Liaoning province (201602094) and the National Natural Science Foundation of China under Grant Nos. 61602076.

References

1. Sweeney, L.: K-anonymity: a model for protecting privacy. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **10**(5), 557–570 (2002)
2. Gkountouna, O., Terrovitis, M.: Anonymizing collections of tree-structured data. *IEEE Trans. Knowl. Data Eng.* **27**(8), 2034–2048 (2015)
3. Landberg, A.H., Nguyen, K., Pardede, E., Rahayu, J.W.: δ -dependency for privacy-preserving XML data publishing. *J. Biomed. Inf.* **50**, 77–94 (2014)
4. Terrovitis, M., Mamoulis, N., Kalnis, P.: Privacy-preserving Anonymization of set-valued data. *PVLDB* **1**(1), 115–125 (2008)
5. Al-Khalifa, S., Jagadish, H.V., Koudas, N., Patel, J., Srivastava, D., Wu, Y.: Structural joins: a primitive for efficient XML query pattern matching. In: *Proceedings of the IEEE International on Data Engineering*, 141–152 (2002)
6. Bruno, N., Koudas, N., Srivastava, D.: Holistic twig joins: optimal XML pattern matching. *ACM SIGMOD* 310–321 (2002)
7. Samarati, P., Sweeney, L.: Protecting privacy when disclosing information: K-anonymity and its enforcement through generalization and suppression. Technical report, CMU, SRI (1998)
8. Sweeney, L.: Achieving K-anonymity privacy protection using generalization and suppression. *Int. J. Uncertain. Fuzziness Knowl.-based Syst.* **10**(5), 571–588 (2002)
9. Zhao, N.N., Liang, Y.W.: Combining structure and content similarities measure for XML document. *Microelectron. Comput.* **33**(4), 69–76 (2016)



A Research on Detection Algorithm of Vehicle Illegal U-Turn

Wenrui Wang^(✉), Yiran Zhou, Qiurun Cai, and Yiyang Zhou

College of Marine Electrical Engineering, Dalian Maritime University,
Dalian, China

wrr362901012@dlmu.edu.cn

Abstract. This paper focuses on the automatic detection of illegal U-turn and recording the number of license plates. The algorithm mainly includes the position detection of vehicles' heads and tails, the marking of vehicles, the license plate recognition, and the judgment of illegal U-turn. The positions of heads and tails are detected by monitoring the headlights and the taillights in RGB color gamut. Most of the headlights are white and the taillights are largely red. In order to avoid the impact of white vehicles on results, the HOUGH Transform and face detection are utilized to determine the position of the steering wheel and the driver, which complement and verify the detection of head position. The edge of the input image needs to be enhanced to improve image quality. It will be judged whether the moving vehicle has had an illegal U-turn. The effectiveness of the algorithm is verified by simulation with MATLAB. Through the analysis of experimental results, it is demonstrated that the algorithm designed in this paper is feasible and achieves the desired design requirements.

Keywords: Vehicle detection · Color detection · Violation detection · MATLAB · HOUGH transform · Face detection

1 Introduction

Traditional detection methods such as induction coils and ultrasonic detectors are difficult to meet the demands of low-cost and high reliability. The traffic system has transferred from traditional manual detection of vehicle violation to automatic detection by computer.

The detection of the head and tail is the first step. From the view of the head and tail of a car, there are several characteristics: 1. the driver is at the head of a car, so the head position can be detected by recognizing the driver's face. 2. There is a steering wheel in the front, which can be detected by its contour. 3. Each car has two lights on the front and rear, which can be detected by color position detection. There are two methods to locate faces. One is face recognition based on the skin color model [1]. It underlies the color of skin while the model is mainly based on color. The other one is the application of the Viola-Jones algorithm [2]. In the course of driving the vehicle, the position of the steering wheel must be in the same direction as the head of the car. And from the perspective of traffic surveillance cameras, most outlines of steering wheels can be

photographed by cameras, which is conducive to the further analysis and processing. Therefore, the position of vehicles' heads can be confirmed by detecting the position of steering wheels. Because the shape of a steering wheel is round, the HOUGH Transform can be utilized to detect the circle [3]. Specific color areas can be utilized to locate the head and tail. In some algorithms, for example, deep learning algorithms [4], the extraction of a high complexity feature may solve problems, but this will take more processing effects at the cost of processing more data. The color features do not need a large amount of calculation. Only the pixel value should be converted to the numerical value. So the color feature is a good feature due to its low complexity.

After detecting the tail and head of a vehicle, its information should be checked. The detection and recognition of license plates are of importance. License plate recognition consists of three parts: license plate location, character segmentation, and character recognition [5].

Three kinds of head and tail detection methods are integrated. The head and tail position is used to identify whether the vehicle has a U-turn. The video image is divided into regions, then the moving vehicles are marked, and the final decision of illegal U-turn will be made.

The rest of the paper is arranged as follows: Sect. 2 introduces the head and tail detection methods and license plate recognition. Section 3 demonstrates the concrete steps of the detection algorithm for illegal U-turn design. The relevant simulation results are shown in Sect. 4. Section 5 concludes the paper.

2 Design of Vehicle Tail and Head Detection Algorithm

2.1 Analysis of License Plate Recognition Algorithm

The steps used in this paper are as follows:

- (1) The image is transformed to grayscale image and filtered to enhance its edge. It is shown that this method not only preserves the information of the license plate area but also reduces the noise.
- (2) The edges of plates can be obtained, and the method is based on the statistical characteristics of blue pixel points. Then the projection area of the edge image is calculated to determine the location of the license plate. The ratio of width to height of the connected domains is calculated to eliminate useless connected domains.
- (3) The license plate area is got from the original image. The character location and segmentation method of projection detection are applied to get a single character.
- (4) The character recognition is based on the neural network algorithm [6], which is trained by the constructed character library [7]. The experimental results are shown in the Figs. 1, 2, and 3.



Fig. 1. Original image



Fig. 2. Position of plate



Fig. 3. License plate

2.2 Algorithm Analysis of Face Location Method

The Viola-Jones algorithm is based on the Haar eigenvalue of the face. The algorithm is divided into three parts.

1. Extract the Haar feature of the target.
2. Through the process of training weak classifiers into the optimal weak classifier, multiple optimal combinations of weak classifiers form a strong classifier [8].
3. Classifiers are cascaded to improve the speed and accuracy of detection [8].

Each Haar returns a single value given by the difference of sum of the pixels in the white region from the sum of pixels in the black region. For the speedy detection of the face features, we consider Haar features as the rectangular region. Some of the Haar filters are given in fig below [9] (Fig. 4).

The other method is based on the skin color model. The advantage of this method is of low complexity. But it is sensitive to the slow change of the background. The flow chart is shown in Fig. 5.

The detection results of the above two methods are shown in Figs. 6, 7 and 8.

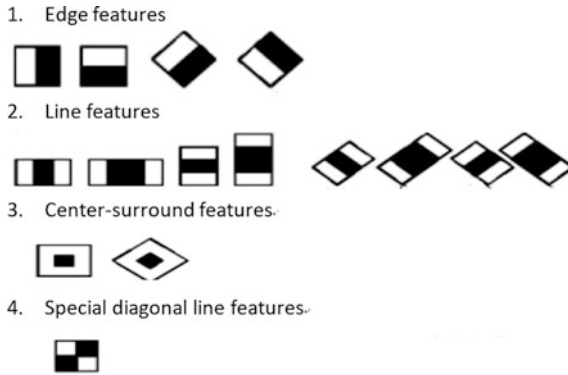


Fig. 4. Haar feature

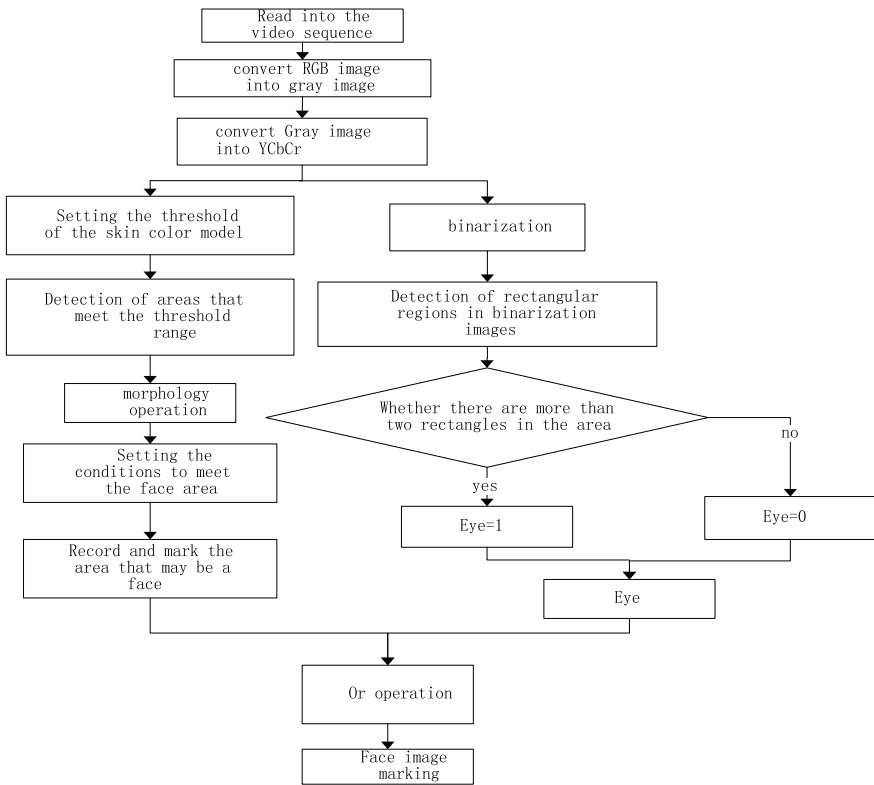


Fig. 5. The flow chart of the face detection

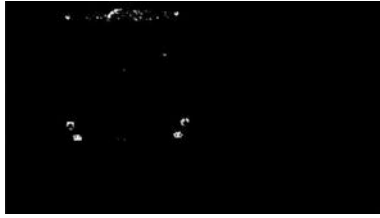


Fig. 6. Region of skin



Fig. 7. Result of skin color model (miss)



Fig. 8. Result of Viola-Jones algorithm

2.3 Analysis of Steering Wheel Positioning Method

As steering wheels are rarely obscured and their colors are deep, they can be easily extracted. The process consists of binarization, morphological processing, and HOUGH transform.

(1) Binary Processing

After repeated selections of the threshold, 0.15 is chosen as the final appropriate one.

(2) Morphological Processing

Mathematical processing has four basic operations: expansion, corrosion, opening, and closing [10]. The morphological processing results are shown in Fig. 9. Through morphological processing, the noise in the binary image is removed.



Fig. 9. Images before and after morphological processing

(3) HOUGH Transform

For any one circle, assuming that center pixel (x_0, y_0) is known, the radius is R , if circles with a radius of R are made at every point in the circle, the value of (x_0, y_0) will be the biggest [3]. This is the mathematical principle of HOUGH transform to find circles. The flow chart is shown in Fig. 10.

The result of steering wheel detection is shown in Fig. 11.

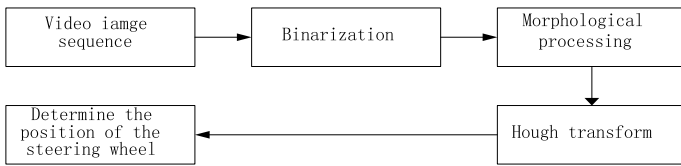


Fig. 10. Flow chart of the steering wheel

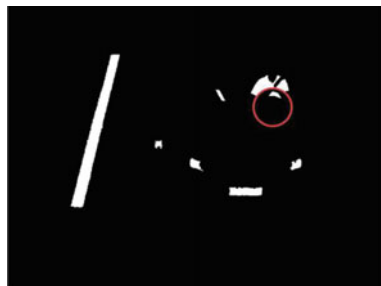


Fig. 11. Result of steering wheel detection

2.4 Analysis of Color Positioning Method

(1) Binarization

This binarization differs from the former one. The transformation assigns 1 to pixels which meet the requirements and assigns 0 to other pixels [11]. This method can locate the head or the tail of a car easily and effectively.

(2) Morphology Operation

Sometimes binary images cannot satisfy demands. Images should be operated morphologically [10]. The test results are shown below. The white area in the figures represents the position of the car lights. (License plate number is “苏D v992” This picture is from the internet.) (Figs. 12 and 13).

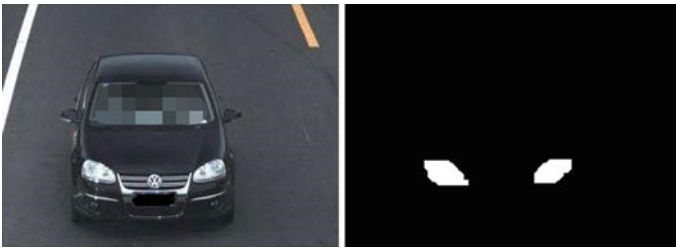


Fig. 12. Result of steering wheel detection

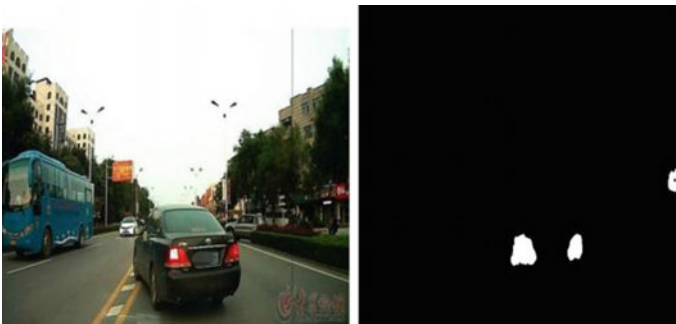


Fig. 13. Result of steering wheel detection

3 Design of Detection Algorithm for Illegal U-Turn

The specific detection methods are as follows:

- (1) The violation detection area is defined as the region of interest.
- (2) The vehicles in the interested region are marked.
- (3) The locations of the vehicles are used to determine whether the vehicle has had illegal U-turn. The flow chart is shown in Fig. 14.

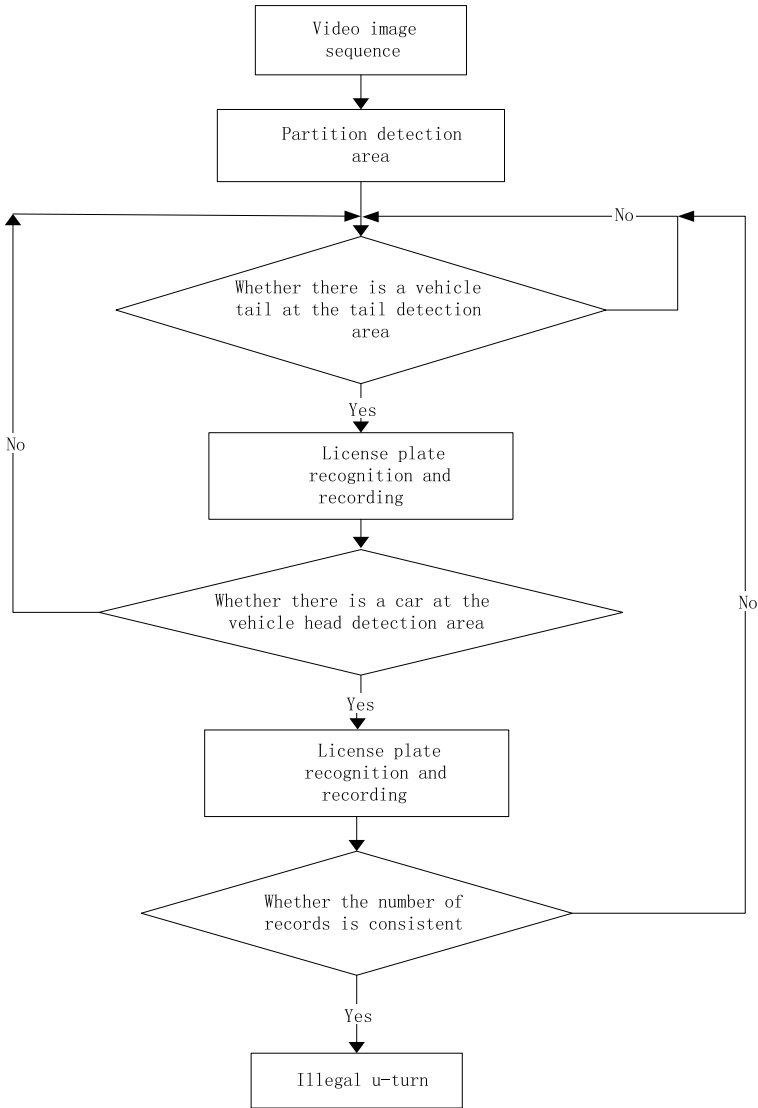


Fig. 14. Flow chart of the detection

3.1 Decision Algorithm for Illegal U-Turn

The region division image is shown in Fig. 15. The brown rectangle area represents the detection area of the tail of the car and the pink rectangle area represents the detection area of the head of the car.

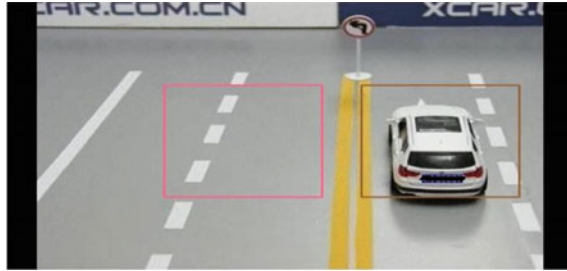


Fig. 15. The region division image

Whether the license plate number of the vehicle appearing in the head detection area is consistent with the one in the tail detection area that can be utilized to determine the illegal U-turn. This is done to identify whether the head part and the tail part belongs to the same car. If the car's head and tail turned up in a short interval, there could be an illegal U-turn. The vehicle's plate number will be recorded if there is an illegal U-turn.

4 Experimental Results and Analysis

The interface is shown in Fig. 16. S1, S2, and S3 are the license plate numbers detected, respectively, in the tail detection area, the head detection area, and recorded car. The head and tail of the vehicle are marked in different colors. The marked image is shown in Fig. 17.

The result shows that the vehicle whose license plate number is “粤N*****” has had an illegal U-turn. As shown in Fig. 17, the vehicle did have a violation of the illegal U-turn. Therefore, it can be judged as a violation of illegal U-turn.


```
Command Window
>> s1

s1 =

'粤N00[REDACTED]'

>> s2

s2 =

'粤N00[REDACTED]'

>> s3

s3 =

'粤N00[REDACTED]'
```

Fig. 16. Record of result

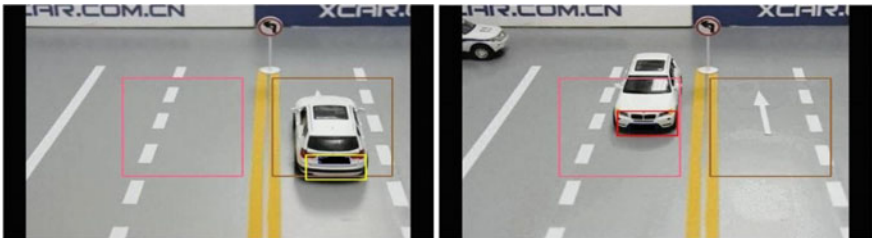


Fig. 17. The marking of car

5 Conclusion

This paper studies the detection algorithm of vehicles' illegal U-turn. It primarily studies the illegal U-turn of vehicles on the straight lane. The main contents include the detection of vehicles' head and tail and the recognition of license plate number. The feasibility of the violation detection algorithm is verified by experiments.

This paper researches the vehicle violation detection algorithm in the environment with little change in the daytime or bright conditions, but the algorithm has some limitations. For the environment of the night and the environment with large brightness change, the detection algorithm has not been studied, so the follow-up can be studied in this field in the future.

References

1. Yan, C., Wang, Z., Zhou, X.: Design of a detection system of faces intercepted by video based on the skin color model. In: 2017 4th International Conference Systems and Informatics (ICSAI), 11–13 Nov. 2017
2. Fernandez, M.C.D., Gob, K.J. E., Leonidas, A.R.M.: Simultaneous face detection and recognition using Viola-Jones algorithm and artificial neural networks for identity verification. In: Proceedings of the 2014 IEEE Region 10 Symposium, 14–16 Apr 2014
3. Ye, H., Shang, G., Wang, L.: A new method based on Hough transform for quick line and circle detection. In: 2015 8th International Conference on Biomedical Engineering and Informatics (BMEI), 11 Feb 2016
4. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521**, 436–444 (2015)
5. Xu, L.: A new method for license plate detection based on color and edge information of lab space. In: Proceedings of the 2011 International Conference on Multimedia and Signal Processing (CMSP), 18 July 2011
6. Madani, K.: Artificial neural networks based image processing & pattern recognition: from concepts to real-world applications. In: First Workshops on Image Processing Theory, Tools and Applications, 2008. IPTA 2008, 09 Jan 2009
7. Noor, K., Siddiquee, E.A., Sarma, D.: Performance analysis of a surveillance system to detect and track vehicles using Haar cascaded classifiers and optical flow method. In: 2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA), 18–20 June 2017
8. Wang, M., Huang, X, Wu, W., Li, Y.: A method of characters in vehicle number-plate using pattern match and neural networks. *J. Huazhong Univ. Sci. Technol.* (2001)
9. Nehru, M., Padmavathi, S.: Illumination invariant face detection using viola jones algorithm. In: 2017 4th International Conference on Systems and Informatics (ICSAI), 11–13 Nov. 2017
10. Yang, H.: Improved image denoising based on mathematical morphology algorithm. *Bull. Sci. Technol.* (2012)
11. Luo, X., Zhao, H., Ogai, H.: A novel video detection design based on modified adaboost algorithm and HSV model. In: Proceedings of the 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 02 Oct 2017



A Dredge Traffic Algorithm for Maintaining Network Stability

Yanan Zhao^{1,2}, Fusheng Dai^{1,2(✉)}, and Jun Shi^{1,2}

¹ Harbin Institute of Technology at Weihai, Weihai 264209, China
dfs1963@163.com

² Science and Technology on Communication Networks Laboratory,
Shijiazhuang 050081, China

Abstract. In order to solve the stability problem of communication network, network traffic instability is prevented, which is caused by local node or link failure. First, based on Lyapunov theory, a random queue model is established for communication network nodes to clear the traffic. Then, the factors are analyzed, which affects the network traffic stability. Finally, it proposes a dynamic scheduling multipath routing transmission algorithm to clear the network traffic steadily. Simulation and experimental results under the most unfavorable conditions of network stability, the performance of dynamic scheduling multipath route is verified to clear the traffic algorithm, which indicates that it is superior to traditional shortest path route dredging algorithm in ensuring network stability. It is proved that the dynamic scheduling multipath routing transmission algorithm can well suppress network turbulence by Lyapunov theory in the case of local network element failure, which ensures communication network stability.

Keywords: Network stability · Lyapunov · Routing algorithm · Network traffic · Multipath transmission

1 Instruction

With the development of the information age, people are increasingly inseparable from the network. They place stringent demands on the Quality of Service (QoS) of network reliability, stability, and quality of service. Network turbulence caused by local device failures, and trigger a large area of the network paralysis must be prevented. Some scholars have tried to use complex network methods to study communication network problems, and tried to solve the problem of linkage-induced failures under the framework of complex network theory [1, 2]. Although these research results are quite theoretically valuable, no consideration has been given to communication network entities, protocol formation, and QoS requirements. Some scholars have also tried to study network stability issues from the aspects of network topology and protocol composition [3, 4]. Although these research results have improved the stability of the network to some extent, there is no clear theoretical support. Retrieving domestic and foreign documents, there are few data on the stability of communications networks from the perspective of modern control theory. If we can study the stability of networks

from the perspective of control theory and find a control method that is easy to implement, we can guarantee the long-term stable operation of communication networks, especially in the case of failure of local network elements. So, we studied the method to solve the network stability problem based on Lyapunov theory and proposed a dynamic scheduling [5] multipath routing transmission algorithm that can steadily clear the network traffic.

2 Network Dredging Traffic Model Based on Lyapunov Theory

2.1 Network Node to Clear Traffic Queue Model

Assume that there are N nodes in the network, constituting a set of nodes V_s . The data flow of access and clearing of edge nodes is f , and the data flow f accessed by all edge nodes constitutes flowing set F . Taking node n as the research object, its adjacent node m forms a set of adjacent nodes $N(t)$ at time t , and there is a queue update equation [6].

$$\begin{aligned}
Q_n(t+1) &= \max[Q_n(t) - b_n(t), 0] + a_n(t) \\
&= \max[Q_n(t) - \sum_{m \in N(t)} b_{nm}^i(t), 0] + \sum_{m \in N(t)} b_{nm}^i(t) + \sum_{f \in F} a_f(t) \delta_{n=n_f^s} \\
&= \max[Q_n(t) - \sum_{m \in N(t)} b_{nm}^i(\alpha_{nm}^i(t, \delta_m^d), S_{nm}^i(t)), 0] \\
&\quad + \sum_{m \in N(t)} b_{nm}^i(\alpha_{nm}^i(t, \delta_m^d), S_{nm}^i(t)) + \sum_{f \in F} a_f(t) \delta_{n=n_f^s}
\end{aligned}$$

Among them, $Q_n(t)$ represents the queue backlog of node n in slot t , and the unit is the number of data packets; $b_n(t)$ denotes the total data packet transmitted by t -slot node n to its neighboring nodes, that is, the data packet that egressed from node n . $b_{nm}^i(t)$ represents the data packet that is transmitted by t -slot node n to its neighboring node

$$m. \quad b_{nm}^i(t) = b_{nm}^i(\alpha_{nm}^i(t, \delta_m^d), S_{nm}^i(t)) = \begin{cases} 0, & \alpha_{nm}^i(t, \delta_m^d) = 0 \\ S_{nm}^i(t), & \alpha_{nm}^i(t, \delta_m^d) = 1 \end{cases} \text{ is the function of }$$

$$S_{nm}^i(t) \text{ and } \alpha_{nm}^i(t, \delta_m^d). \quad \alpha_{nm}^i(t, \delta_m^d) = \begin{cases} 1, & \delta_m^d = \min\{\delta_m^d | m \in N(t)\} \\ 0, & \text{else} \end{cases} \text{ is node } n\text{'s routing}$$

decision for neighboring node m . δ_m^d indicates the relative route hops from the neighbor node m to the destination node d . If δ_m^d of the neighbor node m is the smallest of all the downlink neighbor node sets $N(t)$, node n selects the neighbor node m as the next hop, and pass data packets through the link i between nodes n , m , and this time $\alpha_{nm}^i(t, \delta_m^d) = 1$. $S_{nm}^i(t)$ represents the state of the link i between the nodes n and m , that is, the data packet actually transmitted through the link i ; Obviously, we have

$$b_{nm}^i(t) \leq b_{nm}^{i*}(\alpha_{nm}^i(t, \delta_m^d), S_{nm}^i(t)) = \begin{cases} 0, & \alpha_{nm}^i(t, \delta_m^d) = 0 \\ S_{nm}^i(t) = e_{nm}, & \alpha_{nm}^i(t, \delta_m^d) = 1 \end{cases}, \quad e_{nm}, \quad e_{mn} \text{ represents the link capacity between nodes } n \text{ and } m, \text{ and its physical meaning is the data packet } D_n(out) \text{ and } D_m(out) \text{ that can be transmitted between node } n \text{ and node } m \text{ at}$$

most. $a_n(t)$ represents the data packet that arrives at node n in the t -slot, consisting of two parts: one is $\sum_{m \in N(t)} b_{mn}^i(t)$, which means that the adjacent node m reaches the total packet of node n . Another is $\sum_{f \in F} a_f(t) \delta_{n=n_f}$, indicating the total packet that is accessed or cleared by the edge node at time slot t .

2.2 Constructing a Lyapunov Function for a Node Queue and Calculate Its Lyapunov Offset

Based on the network stability theorem given in [6], the Lyapunov function is constructed with the backlog of node queue, that is $L(\vec{Q}(t)) = \frac{1}{2} \sum_{n \in V_s} Q_n^2(t)$. If there is a

constant M such that there is $L(\vec{Q}(t)) \leq M$ for all t , then the backlog of each node queue is limited to $\sqrt[3]{2M}$, and all queues are stable. The definition of the Lyapunov offset given in [7] is $\Delta(\vec{Q}(t)) = E\{L(\vec{Q}(t+1)) - L(\vec{Q}(t)) | \vec{Q}(t)\}$, indicating the mean value of the mean square value of the queue backlog over time fluctuations. Calculate the change limit of the Lyapunov function from t -slot to $t + 1$ slot as follows:

$$L(\vec{Q}(t+1)) - L(\vec{Q}(t)) = \frac{1}{2} \sum_{n \in V_s} [(\max[Q_n(t) - b_n(t), 0] + a_n(t))^2 - Q_n^2(t)] \quad (1)$$

Inequality scaling for Eq. (1) we have:

$$L(\vec{Q}(t+1)) - L(\vec{Q}(t)) \leq \frac{1}{2} \sum_{n \in V_s} a_n^2(t) + b_n^2(t) + \sum_{n \in V_s} Q_n(t)(a_n(t) - b_n(t)) \quad (2)$$

Substituting Eq. (2) into $\Delta(\vec{Q}(t))$ yields:

$$\begin{aligned} \Delta(\vec{Q}(t)) &\leq \frac{1}{2} E\left\{ \sum_{n \in V_s} \left[\left(\sum_{m \in N(t)} b_{mn}^i(\alpha_{mn}^i(t, \delta_n^d), S_{mn}^i(t)) + \sum_{f \in F} a_f(t) \delta_{n=n_f} \right)^2 \right. \right. \\ &\quad \left. \left. + \left(\sum_{m \in N(t)} b_{nm}^i(\alpha_{nm}^i(t, \delta_m^d), S_{nm}^i(t)) \right)^2 \right] | \vec{Q}(t) \right\} \\ &\quad - E\left\{ \sum_{n \in V_s} Q_n(t) \left(\sum_{m \in N(t)} b_{nm}^i(\alpha_{nm}^i(t, \delta_m^d), S_{nm}^i(t)) - \sum_{m \in N(t)} b_{mn}^i(\alpha_{mn}^i(t, \delta_n^d), S_{mn}^i(t)) \right) | \vec{Q}(t) \right\} \end{aligned} \quad (3)$$

For inequality (3), there are upper bounds for both $b_{mn}^i(\alpha_{mn}^i(t, \delta_n^d), S_{mn}^i(t))$ and $b_{nm}^i(\alpha_{nm}^i(t, \delta_m^d), S_{nm}^i(t))$, that is $D_n(in)$ and $D_n(out)$, respectively. According to the edge nodes adopt the access data flow control technology, we have $a_f(t) \leq \lambda_{\max}$. λ_{\max} represents the maximum data flow. As a result, there is $\sum_{f \in F} a_f(t) \delta_{n=n_f} \leq |F| \lambda_{\max}$. The formula (3) can be written as

$$\begin{aligned}
\Delta(\overrightarrow{Q}(t)) \leq & \frac{1}{2} |V_s| |V_s - 1| ((D_n(in) + |F| \chi_{\max})^2 + D_n^2(out)) \\
& - E \left\{ \sum_{m,n \in V_s} Q_n(t) \left(\sum_{m \in N(t)} b_{nm}^i(\alpha_{nm}^i(t, \delta_m^d), S_{nm}^i(t)) - \sum_{m \in N(t)} b_{mn}^i(\alpha_{mn}^i(t, \delta_n^d), S_{mn}^i(t)) \right) | \overrightarrow{Q}(t) \right\}
\end{aligned} \tag{4}$$

2.3 Routing Strategy Based on Lyapunov Method

For formula (4), [8] gives a theorem that minimizing the upper bound of $\Delta(\overrightarrow{Q}(t))$ can make the node queue tend to be stable, and stable the entire network. Therefore, the upper bound of $\Delta(\overrightarrow{Q}(t))$ is controlled in combination with a specific routing policy to minimize it. Design a routing strategy so that the second item in Eq. (4) as large as possible [9, 10], we have:

$$\begin{aligned}
& \max E \left\{ \sum_{m,n \in V_s} Q_n(t) \sum_{m \in N(t)} (b_{nm}^i(\alpha_{nm}^i(t), S_{nm}^i(t)) - b_{mn}^i(\alpha_{mn}^i(t), S_{mn}^i(t))) | \overrightarrow{Q}(t) \right\} \\
& \leq E \left\{ \sum_{n \in V_s} \sum_{m \in V_s} Q_n(t) - Q_m(t) \sum_{m \in N(t)} [b_{nm}^i(\alpha_{nm}^i(t), S_{nm}^i(t)) - b_{mn}^i(\alpha_{mn}^i(t), S_{mn}^i(t))] | \overrightarrow{Q}(t) \right\} \\
& s.t. \begin{cases} b_{mn}^i(\alpha_{mn}^i(t, \delta_n^d), S_{mn}^i(t)) \leq e_{mn}, b_{nm}^i(\alpha_{nm}^i(t, \delta_m^d), S_{nm}^i(t)) \leq e_{mn} \\ \alpha_{nm}^i(t, \delta_m^d), \alpha_{mn}^i(t, \delta_n^d) \in \{0, 1\} \end{cases}
\end{aligned}$$

In maximize the above problem, it is necessary to jointly consider $Q_m(t)$ and $b_{nm}^i(\alpha_{nm}^i(t, \delta_m^d), S_{nm}^i(t))$, which means when selecting the next hop node, the queue backlog of the node and the relative hops from the node to the destination node should be comprehensively considered. Calculate the routing factor that is $\gamma_m = \frac{1}{Q_m(t) \cdot \delta_m^d}$. Obviously, if the queue's backlog is smaller and the relative hop count to the destination node is smaller, the larger the routing factor is, and the larger the target value will be. Normalize the routing factor γ_m to derive the routing probabilities $P_m = \frac{\gamma_m}{\sum_{m \in N(t)} \gamma_m}$. When node m' performance is significantly lower than that of other adjacent nodes m' , its probability of selection will be small, so that better routes will bear more traffic. If the performance of several adjacent nodes is similar, the probability of selection will be similar, and the network load will be distributed among these routes, effectively avoiding the network turbulence caused by excessive network congestion, and thus maintaining the stability of the network in real time.

3 Experimental Simulation and Simulation Results Analysis

3.1 Simulation Environment Settings

The simulation is based on the OPNET [11]. The simulation scenario is as follows: 25 nodes are distributed as shown in Fig. 1, in which node 0 is the source node and node 24 is the destination node.

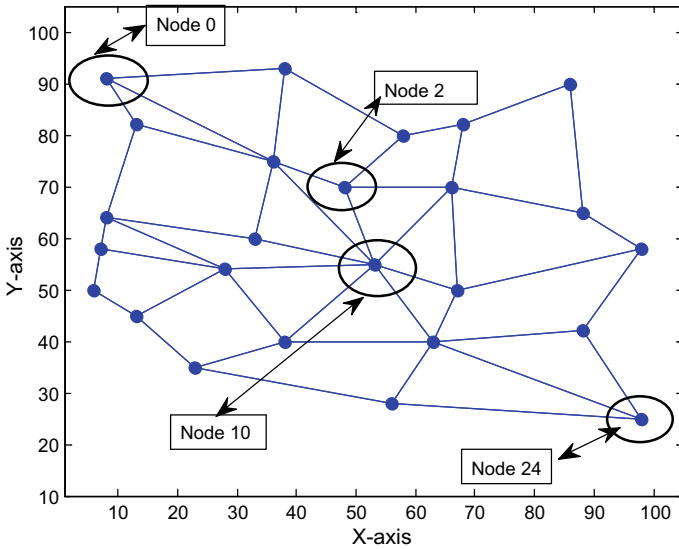


Fig. 1. Network topology

Set the background traffic as follows: during the simulation, the size of the data packet is exponentially distributed with an average value of 1024 bit. Without loss of generality, based on the above background traffic, the traffic flow with extremely large load is set as follows: the number of service data flows is 3 and the start times are 0 s, 2 s, and 4 s, respectively. The interval for generating packets follows the exponential distribution with a mean value of 0.02 s. The packet size follows an exponential distribution with a mean value of 1024 bit. The data rate of the physical layer is 1 Mbps. The simulation time is 400 time slots. Then, set the backbone node 10 to fail, and the traditional routing strategy based on the shortest path algorithm and the dynamic scheduling multipath routing strategy based on the Lyapunov method is simulated. Statistic the average of the backlog of cache queues of all nodes in the two routing strategies and the cache backlog of the MAC layer of the most affected node 2 around the failed node. In eliminating the influence of random factors, the statistics mode of the bucket is used to collect the queue lengths of all the nodes in the MAC layer. The statistical average value of the sampling values of every four time slots is used to obtain the backlog of MAC layer buffer queues of 100 time slots.

3.2 Simulation Results Processing and Analysis

The Lyapunov offset is calculated and curve fitted [7] from all the nodes' of the MAC layer buffer queues for 100 time slots. Figure 2 shows the comparison results of the Lyapunov offset curve fitting for the two routing strategies in the above simulation scenario. It can be seen that the Lyapunov offset of $Q(t)$ based on the shortest path algorithm traditional routing strategy varies drastically, and bounds of the steady are larger and the time is longer. Figure 3 shows that after the failure of the backbone node 10, for the traditional routing strategy, the traffic that the node 10 should have dredged is all pressed to its surrounding node 2, and this node bears almost 70% of the network's load. On the contrary, for the dynamically scheduled multipath routing strategy, traffic that should be dredged by the node 10 is distributed to its neighboring nodes, and the surrounding node 2 only bears 20% of the network's load, reflecting the load balancing. The process of Lyapunov's offset reaching a stable state tends to be gentle, and the time to reach a stable state is significantly shortened. Figure 4 shows the traditional routing strategy based on the shortest path algorithm has a serious deviation from the error curve of the stable center. The stability value of the Lyapunov deviation is above 300, and it does not become stable until the 80th time slot. In contrast, the dynamic scheduling multipath routing strategy controls the stability value of Lyapunov deviation below 150, and the steady-state time does not exceed 30% of the simulation time.

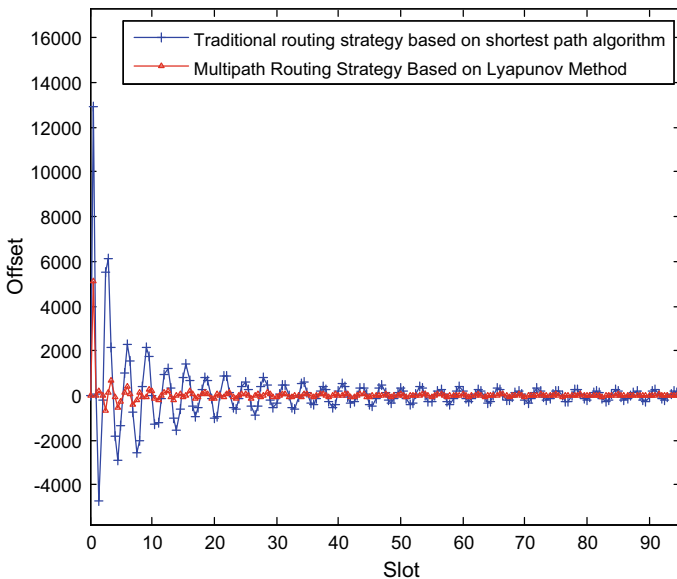


Fig. 2. Comparison of two routing strategies offset

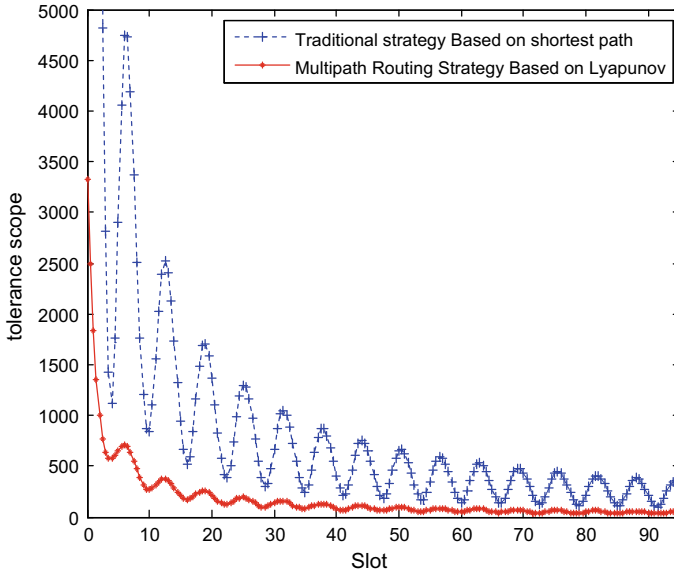


Fig. 3. Node 2 offset error curve local amplification

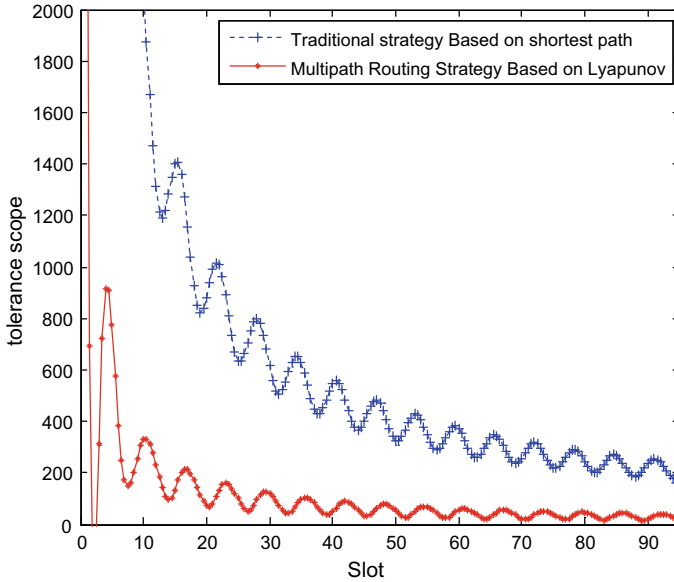


Fig. 4. Offset error curve partial enlargement

Through the simulation experiments, we found that the dynamic scheduling multipath routing strategy based on Lyapunov method is superior to the conventional routing strategy based on the shortest path algorithm in maintaining network stability. First, the multipath routing strategy can quickly clear the traffic transferred by the failed network elements, and the network congestion processing capability is improved by 50%. Compared with the obvious network turbulence presented by traditional routing strategy, it shows better control network stability. Second, the intensity of control over the bounds of Lyapunov's deviation has increased by 50%, and the control time for reaching a steady state has also been reduced by 30%.

4 Conclusion

Using the classical Lyapunov theory, a dynamic scheduling multipath routing strategy that can effectively maintain the stability of the network is proposed. Compared with the traditional routing algorithm, it considers the number of route hops and the backlog of cache queues at the MAC layer of the node during the routing process, and monitors the queue backlog of the entire network node in real time, thus ensuring the stability of the network. By comparing two routing strategies by setting the most unfavorable to the network stability. It can be seen that the dynamic scheduling multipath routing strategy based on Lyapunov method has better performance in maintaining network stability, and can be more suitable for the increase of network traffic load and network environment deterioration, and is more in line with the requirements of the stability of modern communication networks.

References

1. Yan, J., et al.: Integrated security analysis on cascading failure in complex networks. IEEE Press (2014)
2. Mei, D., et al.: Models of cascading failures based on complex networks theory. *Comput. Dig. Eng.* (2014)
3. Mahboub, A., et al.: Multi-zonal approach clustering based on stable election protocol in heterogeneous wireless sensor networks. In: *IEEE International Colloquium on Information Science and Technology* (2017)
4. Premkumar, H.C., et al.: Cognitive agent based stable routing protocol for vehicle-to-vehicle communication. In: *IEEE India Conference*, pp. 1–5 (2016)
5. Neely, M.J.: Universal scheduling for networks with arbitrary traffic, channels, and mobility. In: *IEEE decision and control*, pp. 1822–1829 (2012)
6. Neely, M.: Stochastic network optimization with application to communication and queueing systems. *Synth. Lect. Commun. Netw.* **3**(1), 211 (2010)
7. Tian-qi, G., et al.: Curve fitting method for closed discrete points. *J. Jilin Univ.* **45**(2), 437–441 (2015)
8. Su, P., et al.: Energy-constrained dynamic scheduling and dynamic pricing algorithm in wireless cloud computing. *Dianzi Yu Xinxu Xuebao/J. Electr. Inf. Technol.* **39**(5), 1150–1156 (2017)

9. Neely, M.J.: Stability and probability 1 convergence for queueing networks via Lyapunov optimization. *J. Appl. Math.* **2012**(1110-757X), 203–222 (2012)
10. Retchkiman Königsberg, Z.: Modeling stability analysis and scheduling design for a class of queueing systems by means of timed petri nets Lyapunov methods and max-plus algebra. *Int. J. Pure Appl. Math.* **77**(3), 1457–1476 (2011)
11. Min, C.: *OPNET Network Simulation*. Tsinghua University Press (2004)



Attacks Detection Method Based on Free Space Quantum Secure Direct Communication

Jinlong Liu, Zhutian Yang, and Zhilu Wu^(✉)

School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150006, China
{yq20,yanzhutian,wuzhilu}@hit.edu.cn

Abstract. Quantum secure direct communication is a new communication mode in the free space. Quantum secure direct communication utilizes single photon with the information to transmit ciphertext in the quantum channel directly, and it distinguishes from Quantum key distribution. Its outstanding characteristics are as follows: super velocity of light, high power capacity, high security, and anti-interference, so quantum secure direct communication can break the traditional secret communication frame to ensure the absolute secure communication. However, the imperfection of optical devices can leave loopholes for quantum attacks. Based on free space, this paper puts forward one kind of attack detection method in the field of quantum secure direct communication. This paper introduces quantum bit error rate analysis and decoy-state photon transmission rate analysis for designing security detection process to effectively detect the eavesdropping and improve security.

Keywords: Free space · Quantum secure direct communication · Quantum bit error rate · Decoy state

1 Introduction

Security is the root of quantum communication, but more and more attacks come into quantum secure direct communication, and some attacks greatly threaten the security [1]. Three kinds of attacks exist in the transmission process, such as perfect light source attack, practical light source attack, and optical device attack [2]. Perfect light-source attack mainly includes incoherent attack, Trojan horse attack, photon number separate attack (PNS), and these attacks utilize intercepting quantum, camouflaging quantum, analyzing code rule, and separating photon number to break security [3–5]. Some detection methods have been studied for the perfect light source attack. Comparing with the perfect light source attack, practical light source attack, and optical device attack possess more serious threat for the security, because of the practical optical channel which the transmission process executes in, such as intensity fluctuation attack [6], untrusted source attack [7], phase-remapping attack [8], multi sources attack [9], and passive Faraday-mirror attack [10]. Recently, many and varied attacks are constantly updated for quantum secure direct communication, but the coping strategy is studied lately, so this paper proposes multiple attacks detection method based on free-space quantum secure direct communication.

2 Security Detection Process

Based on the imperfect optical element, the encrypted data are intercepted or eavesdropped during quantum secure direct communication, so the security detection needs to introduce quantum bit error rate analysis and decoy-state photon transmission rate analysis. On the basis of QSDC transporting mechanism, the security detection process is designed as follows:

- Step 1: Alice and Bob execute channel detection before quantum secret communication. The transmitter Alice utilizes the damped low power laser to take the place of the single-photon source and manufactures one single-photon sequence S_b .
- Step 2: Decoy-state photons are added into S_b before they are transmitted.
- Step 3: Alice transmits the single-photon sequence S_b to Bob, and Bob randomly chooses the conjugate bases to measure S_b , then Bob informs which the conjugate base is chosen to Alice. Alice replies the correct position of the conjugate base and the position of decoy-state photon to Bob. Alice and Bob compare the decoy-state photon transmission rate in the transmission to detect whether the eavesdrop exists or not.
- Step 4: Alice and Bob execute the physical detection to eliminate eavesdrop and attack when there are the eavesdrop and the attack in the quantum channel.
- Step 5: Alice and Bob wipe off the detected decoy-state photons after eliminating eavesdrop and attack, and choose a part of random photons in S_b . Alice and Bob execute the quantum query to publicly compare the quantum bit error rate.
- Step 6: When the quantum bit error rate is less than the threshold value which the two parties set, Alice and Bob wipe off the compared single photons, and execute quantum information processing for k single photon to generate the control code N_k .

Alice and Bob execute the channel detection before the quantum secret communication, Alice generates one single-photon sequence S_b which is added with the decoy-state photons. Alice transmits the single-photon sequence S_b to Bob, Fig. 1 shows the transmission of S_b , and Bob randomly chooses the conjugate bases to measure S_b , then both sides compare quantum bit error rate and the decoy-state photon transmission rate Q_μ . When the quantum bit error rate is less than the threshold value and the decoy-state photon transmission rate $Q_\mu \gg Q'_\mu$, $Q'_\mu = Q''_\mu$, the quantum channel is determined in the security. The control code N_k is generated after eliminating the quantum photon.

At the discovery stage, each node needs to know which node is its neighbor node and how far it is from the root node. A wireless node generates its own routing table through a dual program: (I) a typical tree based addressing scheme can reduce the size of the address field in the message header, (II) a link between neighbors in the neighborhood to generate the final grid tree topology. After the root node receives the hierarchical discovery of the child node broadcast, the child node is added to its child node table. The root node is responsible for the continuous 16 bits' logical block address assigned to each node, as shown in Fig. 1a. Figure 1b shows that the route

robustness process is to reduce the energy loss and network congestion caused by flooding broadcast through additional local links. This stage is mainly to establish the hierarchical structure of the whole network.

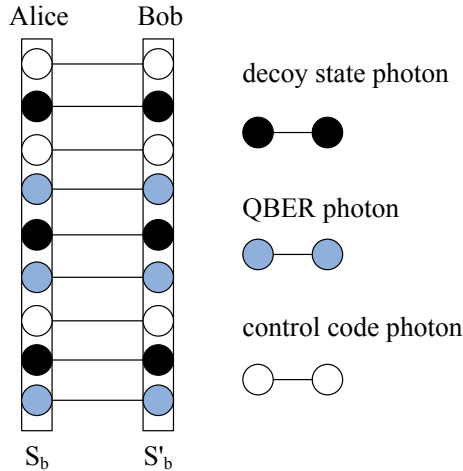


Fig. 1. Schematic diagram of information transmission

3 Quantum Bit Error Rate

Quantum bit error rate is a ratio that takes into account the received error quantum bits and the total quantum bits, it can be abbreviated as QBER (Formula 1), R_{sift} is quantum key bit rate with data filters (include correct codes and error codes), R_{error} is error bit rate in quantum channel.

$$QBER = \frac{N_{wrong}}{N_{right} + N_{wrong}} = \frac{R_{error}}{R_{sift}} \tag{1}$$

The highly attenuated laser pulses with single-photon sources can be regarded as a statistic process obeying Poisson distribution, and the theoretical model of the quantum bit error rate caused by the background light is established and the expressions of the quantum bit error rate are given by Formula 2, and every pulse photon number is μ .

$$P(n, \mu) = \frac{\mu^n}{n!} \exp(-\mu) \tag{2}$$

Transmission/Reverberation of single photon obeys the binomial distribution in the optical link, and the probability which receiver at least detects one photon is given by Formula 3, T_{app} is system devices transmission rate, η_{det} is quantum efficiency of single-photon detector.

$$\begin{aligned}
 P_{n \geq 1} &= \sum_{k=1}^n \binom{n}{k} (T_{cha}P_{aeq}T_{app}\eta_{det}F_{mea})^k (1 - T_{cha}P_{aeq}T_{app}\eta_{det}F_{mea})^{n-k} \\
 &= 1 - (1 - T_{cha}P_{aeq}T_{app}\eta_{det}F_{mea})^n
 \end{aligned} \tag{3}$$

According to Formulas 2 and 3, the original quantum key bit rate is represented as Formula 4, R_{rep} is the transmitter pulse repetition rate.

$$\begin{aligned}
 R_{raw} &= R_{rep} \sum_{n=1}^N P(n, \mu) P_{n \geq 1} = R_{rep} \sum_{n=1}^N \frac{\mu^n}{n!} \exp(-\mu) [1 - (1 - T_{cha}P_{aeq}T_{app}\eta_{det}F_{mea})^n] \\
 &= R_{rep} [1 - \exp(-\mu T_{cha}P_{aeq}T_{app}\eta_{det}F_{mea})^n]
 \end{aligned} \tag{4}$$

Quantum key bit rate with data screening is Formula 5

$$R_{sift} = F_{sift}R_{raw} = F_{sift}R_{rep} [1 - \exp(-\mu T_{cha}P_{aeq}T_{app}\eta_{det}F_{mea})] \tag{5}$$

The main factors that cause the quantum bit error rate are the optical elements, the detector dark noise and space optical environment in free-space quantum key distribution. The three factors that cause the quantum bit error rate are R_{opt} , R_{det} , and R_{env} , so the overall $QBER$ is given by Formula 6, $QBER_{opt}$, $QBER_{det}$, and $QBER_{env}$, respectively, correspond to the $QBER$ from the optical elements, the detector dark noise, and space optical environment.

$$QBER = (R_{opt} + R_{det} + R_{env})/R_{sift} = QBER_{opt} + QBER_{det} + QBER_{env} \tag{6}$$

3.1 Analyze Quantum Bit Error Rate Caused by Optical Element

Optical element takes on nonideal response and imperfect alignment for the quantum direct communication system, so the photons cannot reach the correct detector, thus cause error codes. The probability of assuming a photon to arrive at an error detector is P_{opt} , quantum bit error rate caused by optical elements is given by Formulas 7 and 8.

$$R_{opt} = R_{sift}P_{opt} = F_{sift}R_{rep} [1 - \exp(-\mu T_{cha}P_{aeq}T_{app}\eta_{det}F_{mea})^n] P_{opt} \tag{7}$$

$$QBER_{opt} = R_{opt}/R_{sift} \tag{8}$$

3.2 Analyze Quantum Bit Error Rate Caused by Dark Count Probability for Single-Photon Detector

Dark count probability for single-photon detector always exists to cause error codes in the free space quantum key distribution system. Assume that the dark count probability for single-photon detector is R_{dark} and the time window is τ , the dark count probability

for single photon is described as Formulas 9 and 10 in every detector and every time window.

$$R_{dark} = R_{dark}N\tau \quad (9)$$

$$R_{dark} = F_{sift}R_{rep}R_{dark}N/2 \quad (10)$$

Number of detectors is N , the probability of correct codes and error codes both are $1/2$, so quantum bit error rate caused by dark count probability for single-photon detector is Formula 11.

$$QBER_{det} = R_{det}/R_{sift} = \frac{1}{2} \frac{R_{dark}N\tau}{1 - \exp(-\mu T_{cha}P_{aeq}T_{app}\eta_{det}F_{mea})} \quad (11)$$

3.3 Analyze Quantum Bit Error Rate Caused by Optical Environment

The background photon arrives at receiver to cause error codes in the background radiation of free space. The background is taken as a homogeneous radiation source in the background radiation model, and background rate is Formula 12. Background radiation spectrum is $W(\lambda)$ and the bandwidth nearby wavelength λ is $\Delta\lambda$, the area of the receiver is A , the viewing angle of the receiver is Ψ_{fv} , the solid angle of background radiation measured by the receiver is Ψ_s , so R_{env} is given by Formula 13.

$$R_{bac} = \begin{cases} W(\lambda)(\Delta\lambda)\Omega_{fv}A & (\Omega_{fv} < \Omega_s) \\ W(\lambda)(\Delta\lambda)\Omega_sA & (\Omega_{fv} > \Omega_s) \end{cases} \quad (12)$$

$$R_{env} = F_{sift}R_{rep}R_{bac}N\tau/2 \quad (13)$$

Quantum bit error rate caused by the optical environment is Formula 14.

$$QBER_{env} = R_{env}/R_{sift} = \frac{1}{2} \frac{R_{bac}N\tau}{1 - \exp(-\mu T_{cha}P_{aeq}T_{app}\eta_{det}F_{mea})} \quad (14)$$

The general expression of quantum bit error rate is finally confirmed as Formula 15 after taking Formulas 8, 11, and 14 into Formula 6.

$$QBER_{env} = P_{opt} + \frac{N\tau}{2} \frac{R_{dark} + R_{bac}}{1 - \exp(-\mu T_{cha}P_{aeq}T_{cpp}T_{det}F_{mea})} \quad (15)$$

4 Rate of Decoy-State Photon Pulse

The single-photon sources used in most practical QSDC systems are an attenuated weak laser pulse added with decoy-state photons. Eve intercepts all single-photon pulses, and saves and repeats multi-photon pulse. Eve cannot distinguish decoy-state photons with others, and the transmission rate of decoy-state photon pulse equal to the transmission rate of multiple photon, so the transmitter can compare the transmission rate of decoy-state photon pulse with the transmission rate of multiple photon to determine whether there are PNS attack and imperfect light attack.

The mean intensity of a uniform coherent light source is μ , and photon numbers obey the poisson distribution, Formula 16 gives the pulse transmission rate Q_μ . When Alice sends n photon pulses, detection probability of Bob is Y_n , total transmission rate of the channel is Q_μ . Transmission rate of decoy-state photons is Q'_μ when there is not PNS attack, and transmission rate of decoy-state photons is Q''_μ when there is PNS attack, transmission rate are given by Formulas 17 and 18. During two formulas, the efficiency of the detector is η_B ; dark count probability from Bob is Y_n ; the transmission rate of multiple photon permitted by Eve is p when there is not PNS attack. Alice and Bob can detect whether the PNS attack exists or not based on analyzing Q_μ , Q'_μ , and Q''_μ (Table 1).

$$Q_\mu = \sum_n Y_n p_n(\mu) = Y_0 e^{-\mu} + \sum_{n=1}^\infty Y_n e^{-\mu} \frac{\mu^n}{n!} \tag{16}$$

$$Q'_\mu = Y_0 + 1 - (1 + \mu)e^{-\mu} - \frac{e^{-\mu(1-\eta_B)}}{1 - \eta_B} \left(e^{-\mu(1-\eta_B)} - 1 - \mu + \mu\eta_B \right) \tag{17}$$

$$Q''_\mu = (Y_0 + p - (1 + \mu)pe^{-\mu}) - \frac{e^{-\mu}}{1 - \eta_B} \left(e^{-\mu(1-\eta_B)} - 1 - \mu + \mu\eta_B \right) \tag{18}$$

Table 1. PNS attack judgement condition

No PNS	$Q_\mu \gg Q'_\mu$	$Q'_\mu = Q''_\mu$
Existing PNS	$Q_\mu \leq Q_\mu$	$Q_\mu > Q''_\mu$

5 Experimental Data Analysis

A simulation experiment is designed according to the research content and transmission step. Based on block transmission mechanism of control code, simulation program divides the secret information into a group of m quantum bits, so the control code contains m quantum bits. Due to the randomness of the detection base, only about 50% quantum photons can choose the same detection base, so the quantum sequence length doubles initial quantum keys. Further, the quantum sequence length is identified

quadruple the control codes because of some initial quantum keys which Alice and Bob choose to compare the information. Incoherent light source attack and imperfect light source attack are led into the experiment, and the experiment is carried 1000 times. The threshold value is defined 0.13 with the analysis of quantum bit error rate based on Table 2, and the transmission rate of decoy-state photon pulse is calculated such as Table 3. So the simulation experiment can effectively detect attacks to improve security.

Table 2. Analysis of quantum bit error rate

Quantum bit error rate	Below 0.13	0.13–0.21	0.21–0.25	0.25–0.29	0.29–0.33	Above 0.33
Times	0	3	18	17	51	911

Table 3. PNS attack judgement result

No PNS	$Q_{\mu}(0.96) \gg Q'_{\mu}(0.08)$	$Q'_{\mu} = Q''_{\mu} = 0.08$
Existing PNS	$Q_{\mu}(0.05) \leq Q'_{\mu}(0.83)$	$Q_{\mu}(0.73) \gg Q''_{\mu}(0.09)$

6 Conclusion

Quantum secure direct communication has been different from quantum key distribution; quantum secure direct communication makes the cryptograph to be transmitted directly in the quantum channel. Its outstanding characteristics are as follows: super velocity of light, high power capacity, high security, and anti-interference, but the imperfection of optical devices can leave loopholes for the quantum attack. Based on free space, this paper puts forward one kind of attack detection method in the field of quantum secure direct communication, the main innovative points as follows: establish one theory model which put the control code into the block to select the polarization conjugate base, introduce quantum bit error rate analysis and decoy-state photon transmission rate analysis for designing security detection process; according to the analysis of the simulation result, this paper proves that the process can effectively detect the eavesdropping and can improve security.

References

1. Göpfert, F., Vredendaal, C.V., Wunderer T.: A hybrid lattice basis reduction and quantum search attack on LWE. In: International Workshop on Post-Quantum Cryptography, pp. 184–202. Springer, Cham (2017)
2. Wei, K., Liu, H., Ma, H.: Feasible attack on detector-device-independent quantum key distribution. *Sci. Rep.* **7**, 449 (2017)
3. Gyongyosi, L., Imre, S.: Geometrical estimation of information-theoretical impacts of incoherent attacks for quantum cryptography. *Int. Rev. Phys.* (2010)

4. Gisin, N., Fasel, S., Kraus, B.: Trojan-horse attacks on quantum-key-distribution systems. *Phys. Rev. A* **73**, 457–460 (2006)
5. Kulik, S.P., Molotkov, S.N.: Decoy state method for quantum cryptography based on phase coding into faint laser pulses. *Laser Phys. Lett.* **14**, 125205 (2017)
6. Popov, M.V., Kondrat'Ev, V.I., Altunim, V.I.: Parameters of microstructure and noiselike intensity fluctuation in pulsar radio emission measured with submicrosecond time resolution provided by the S2 VLBI recording playback system, vol. 202, p. 179 (2016)
7. Xu, B., Peng, X., Guo, H.: The security of SARG04 protocol in plug and play QKD system with an untrusted source. *Quantum Inf. Comput.* **12**, 630–647 (2012)
8. Chen, J., Wang, J., Qin, X.: Nonlinear active phase compensation attack based on multiple operators for measurement. *Laser Optoelectron. Prog.* **53**, 072701 (2016)
9. Zhao-Jun, G.U., Bo, H.E.: Multi-source attack graph intrusion detection algorithm based on suspicious queue. *Comput. Eng. Des.* (2017)
10. Sun, S.H., Jiang, M.S., Liang, L.M.: Passive Faraday-mirror attack in a practical two-way quantum-key-distribution system. *Phys. Rev. A* **83**, 161–164 (2011)



Research on Model and Application of Elevator Safety Remaining Service Life

Qi Li, Zhenfeng Shi^(✉), and Shang Sun

School of Science, Harbin Institute of Technology, 92 West Dazhi Street,
Nan Gang District, Harbin, Heilongjiang, China
szf@hit.edu.cn

Abstract. In view of the lack of maintenance of elevators in China, this paper builds a statistical model of elevators' remaining service life based on elevator operation and maintenance management data, plots the remaining service life curve of elevators, and implements the prediction of preventive maintenance intervals from time to time and scrap time (i.e., the safety remaining service life) for elevators, and verifies the reliability of the model through empirical analysis.

Keywords: Maintenance interval · Safety remaining service life · BPNN · FCM · Slope threshold · Techno-economic indicator

1 Introduction

With the development of urbanization, the demand for elevators has risen, and elevator safety issues have caused public panic. As a loss-type mechanical equipment, the elevator implements a regular maintenance system. When the elevator enters the depletion period, if the system of regular maintenance is still implemented, not only is there a huge potential safety risk, it will also damage economic efficiency. In this context, the subject establishes a model for the remaining service life of elevators.

Although there have been relatively significant research results in the field of mechanical equipment maintenance intervals and remaining service life, but most of the studies are to solve the optimal time intervals and maintenance times within a given time zone (Wang Yvjun, Liu Qinming, Mailk et al. have achieved remarkable results in this field) or directly predict the remaining life of the equipment (Yang Zichun, Zhu Q. S., Mulubrhan, Shien-Tsung Chen and others made outstanding contributions to this field). Therefore, this project hopes to build a model for the remaining service life of elevators which can determine the maintenance time interval and predict the remaining service life on the basis of existing researches, making the model more practical.

2 Elevator Safety Remaining Life Model Without Maintenance

After the elevator equipment enters the depletion period, the reliability of the equipment is continuously declining. When it drops to a certain threshold, the first preventive maintenance of the equipment during the depletion period is required. The purpose of this section is to establish a model for the safety remaining life of elevators without maintenance under the first preventive maintenance and to determine the time interval from the current time to the first preventive maintenance. According to the relevant literature, the indicator system that can reflect the healthy operation state of the elevator is determined, including 4 subsystems: elevator equipment car system, traction machine system, door system, and drag system, a total of 17 specific indicators. The model establishment process is as follows.

2.1 Select Related Indicators

First perform principal component analysis of each subsystem's original index separately. The comprehensive influence weights of the original index on the evaluation object can be determined based on the results of the principal component analysis. Through the weighted average, four comprehensive indicators X_1, X_2, X_3, X_4 , which represent healthy operation status of four subsystems of elevator equipment are obtained.

2.2 Classification of Healthy Operation Levels of Elevators

The FCM is used to classify the monitoring sample data into three categories: 'excellent', 'good', and 'middle'. In order to reduce the number of iterations and improve efficiency, the initial clustering center is determined by the traditional system clustering method.

2.3 Construct Elevator Life Index

The elevator life index is constructed based on the rank sum ratio (RSR) comprehensive evaluation method, and the elevator life index corresponding to each observation sample is calculated. At the same time, the elevator life index corresponding to the three cluster centers of fuzzy clustering can also be obtained. The specific method of constructing the elevator life index is as follows:

- (1) Write sample rank according to the nature of the index: there are known n observation samples $x_i = (x_{i1}, x_{i2}, \dots, x_{im}), i = 1, 2, \dots, n$, each sample is m -dimensional. Finally, the rank matrix $R = (r_{ij})_{n \times m}$ is obtained.
- (2) Calculate the rank sum ratio statistics: when the impact weights of the indicators on the evaluation objects are not equal, the weighted rank sum ratio statistics WRSR need to be calculated, and

$$WRSR_i = \frac{1}{n} \sum_{j=1}^m w_j r_{ij}, i = 1, 2, \dots, n$$

where w_j is the influence weight of the j -th indicator on the evaluation object.

In fuzzy clustering, the monitoring samples have been divided into three categories: ‘excellent’, ‘good’, and ‘middle’. In this section, the correlation coefficient between each index and the category variable is used to determine the weight of each evaluation index, and $|r_j^*|$ is the absolute value of the correlation coefficient between the j th evaluation index and the categorical variable. So the elevator life index Ra is calculated as

$$Ra_i = \sum_{j=1}^m r_{ij}w_j = \sum_{j=1}^m r_{ij}(|r_j^*| / \sum_{j=1}^m |r_j^*|), (i = 1, 2, \dots, n)$$

2.4 Prediction Based on BP Neural Network

First use BP neural network algorithm to complete the prediction of the elevator life index at the next moment, and set the limit conditions to determine the time interval for the first preventive maintenance. Assume that there are n monitoring samples x_1, x_2, \dots, x_n . The specific steps are as follows:

- (1) Take the sample monitoring value of the elevator for any consecutive p times as the input node, and the sample detection value of the $p + 1$ th time as the output node, set the number of nodes in the hidden layer to n , and train the BP neural network;
- (2) Use the trained BP neural network model to predict the healthy running status of the elevator in the future;
- (3) Calculate the distance between the prediction sample and the three types of cluster centers: ‘excellent’, ‘good’, and ‘middle’, and classify the health status of the new prediction sample;
- (4) The health status of the elevator will decay with time. If the prediction sample has been classified as ‘middle’, set the distance threshold ε_1 to determine whether the elevator’s running status has entered the ‘bad’ category, once the distance from the prediction sample to the ‘middle’ class exceeds ε_1 , stop forecasting. And it is considered that the elevator’s operating status grade has been classified as bad, which is the elevator’s scrap point or end-of-life point. The specific setting method of ε_1 is as follows: Using the results of fuzzy clustering, calculate the distance $d_{2i} (i = 1, 2, \dots, n)$ between all monitoring samples in the ‘middle’ category and cluster center of this type, and can be defined as

$$\varepsilon_1 = \max\{d_{2i}\}, i = 1, 2, \dots, n$$

2.5 Fit the Remaining Service Life Curve of the Elevator Without Fitting Maintenance

Use the curve fitting the elevator life index from the beginning of monitoring to the end of the prediction, it is considered that the attenuation law of the elevator

life index obeys the Weibull distribution. Finally, the function form of the safety remaining service life curve without maintenance is determined as

$$Ra(t) = e^{-\left(\frac{t}{\beta}\right)^\alpha}$$

where α and β are all unknown parameters, which can be determined by specific elevator operation data.

3 Elevator Safety Remaining Life Model with Maintenance

Any kind of equipment that is scrapped without maintenance is very uneconomical and causes great waste. Therefore, the model of the remaining service life of the elevator without maintenance is the basis for the establishment of the model under maintenance. Assume that based on the model of elevator safety remaining life without maintenance, the function form of the elevator’s remaining service life curve under non-maintenance conditions has been determined as

$$R_{a_1}(t) = e^{-\left(\frac{t}{\beta_1}\right)^{\alpha_1}}$$

$R_{a_1}(t)$ can be regarded as the elevator’s remaining safety life model before the first maintenance, where α_1, β_1 represent the model parameters of the elevator’s safety remaining service life curve before the first maintenance and can be obtained by the elevator operating data. The process of establishing the safety remaining service life model for elevators under maintenance conditions is summarized below.

3.1 Determine the Slope Threshold of Elevator Safety Remaining Service Life Model

According to the model $R_{a_1}(t)$ of elevator safety remaining life without maintenance, the time point for the first time h_1^{\max} needing preventive maintenance can be obtained. The slope of $R_{a_1}(t)$ at this point is the given elevator performance degradation rate threshold ε_2 . There are

$$\varepsilon_2 = R_{a_1}'(h_1^{\max}) = -\alpha_1 \frac{(h_1^{\max})^{\alpha_1 - 1}}{\beta_1^{\alpha_1}} e^{-\left(\frac{h_1^{\max}}{\beta_1}\right)^{\alpha_1}}$$

3.2 Determine the Elevator Performance After Preventive Maintenance Based on Performance Restoration Factors

Assuming that the i th preventive maintenance has been completed for the elevator, the elevator performance recovery factor after maintenance is

$$\eta_i = (bC_{pi})^{ci}$$

where $C_{pi} = C_{pf} + \Delta t_i C_{pv} = C_{pf} + s^{i-1} \Delta t_1 C_{pv}$ is the cost of the i th preventive maintenance, which is related to the fixed maintenance cost C_{pf} , variable maintenance cost C_{pv} , and the time consumed Δt_i during the i th preventive maintenance. b, c are adjustment parameters for preventive maintenance costs and maintenance times, which can be artificially given. Assuming that the elevator life index after the i th preventive maintenance is restored to $R_{a_{i+1}}(h_{i+1}^{\min})$, its formula is

$$Ra_{i+1}(h_{i+1}^{\min}) = Ra_i(h_i^{\max}) + \eta_i(Ra_i(h_i^{\min}) - Ra_i(h_i^{\max}))$$

where $Ra_i(t)$ and $Ra_{i+1}(t)$ represent the elevator safety remaining life model before and after the i th preventive maintenance, respectively, and the specific form of $Ra_i(t)$ is known. According to the given slope threshold, using $Ra_i(h_i^{\max}) = \varepsilon_2$ it can solve the h_i^{\max} ,

$$\begin{aligned} h_i^{\min} &= h_{i-1}^{\max} + \Delta t_{i-1} \\ &= h_{i-1}^{\max} + s^{i-2} \Delta t_1 \\ \Delta T_i &= h_i^{\max} - h_i^{\min} \end{aligned}$$

in the same way, h_{i-1}^{\max} is reverse solution using $Ra_i(h_{i-1}^{\max}) = \varepsilon_2$. Here, the starting point of the elevator's remaining service life model after the i -th repair is $(h_{i+1}^{\min}, Ra_{i+1}(h_{i+1}^{\min}))$.

3.3 Determine the Safety Remaining Service Life Model After the i th Preventive Maintenance

It is known that the time consumed for the i th service is Δt_i , and the time interval from the end of the last maintenance to the i th service is denoted as ΔT_i . And the decay rate of the elevator life index after maintenance is greater by δ times than the decay rate under the same life index of the elevator before the maintenance, and the elevator safety remaining life model $Ra_{i+1}(t)$ after the i th maintenance can be determined according to the following formula:

$$\begin{aligned} Ra_{i+1}(\sum_{k=1}^i \Delta T_k + \sum_{k=1}^i \Delta t_k) &= Ra_i(\sum_{k=1}^i \Delta T_k + \sum_{k=1}^{i-1} \Delta t_k) + \eta_i(Ra_i(\sum_{k=1}^{i-1} \Delta T_k + \sum_{k=1}^{i-1} \Delta t_k) \\ &\quad - Ra_i(\sum_{k=1}^i \Delta T_k + \sum_{k=1}^{i-1} \Delta t_k)) \\ Ra_{i+1}(t) &= \int_0^t \delta^i Ra_1'(u - \sum_{k=1}^i (\Delta t_k + \Delta T_k)) du + D_i \\ &= \delta^i Ra_1(t - \sum_{k=1}^i (\Delta t_k + \Delta T_k)) + D_i \end{aligned}$$

where $i = 2, 3, \dots$ and $Ra_{a_{i+1}}'(h_{i+1}^{\max}) = \varepsilon_2, i = 1, 2, 3, \dots$

3.4 Judge Whether the i th Maintenance Should be Performed Based on the Techno-Economic Indicator

Assuming that the elevator’s safety remaining service life model after the i th maintenance has been determined, the following judges whether the $(i+1)$ th preventive maintenance should be performed on the elevator based on the threshold ε_3 of the techno-economic indicator K . The calculation formula for the maintenance cost K per unit of time before the $(i+1)$ th maintenance is

$$K_{i+1} = \frac{\sum_{k=1}^{i+1} C_{mr} F_k + \sum_{k=1}^i C_{pk} + \sum_{k=1}^i C_{lk}}{\sum_{k=1}^{i+1} \Delta T_k + \sum_{k=1}^i \Delta t_k} = \frac{\sum_{k=1}^{i+1} C_{mr} (\ln Ra_k (\sum_{i=1}^{k-1} \Delta T_i + \sum_{i=1}^{k-1} \Delta t_i))}{\sum_{k=1}^{i+1} \Delta T_k + \sum_{k=1}^i \Delta t_k} - \frac{\ln Ra_k (\sum_{i=1}^k \Delta T_i + \sum_{i=1}^{k-1} \Delta t_i)}{\sum_{k=1}^{i+1} \Delta T_k + \sum_{k=1}^i \Delta t_k} + \frac{\sum_{k=1}^i (C_{pf} + s^{k-1} \Delta t_1 C_{pv}) + \sum_{k=1}^i s^{k-1} \Delta t_1 C_{lu}}{\sum_{k=1}^{i+1} \Delta T_k + \sum_{k=1}^i \Delta t_k}$$

where C_{mr} is the fixed maintenance cost for each failure; C_{pf} , C_{pv} are the fixed maintenance cost and the variable maintenance cost for each preventive maintenance; C_{lu} is the loss cost per unit time during the preventive maintenance of the elevator; α_k , β_k are the model parameters in the remaining service life model $R_{\alpha_k}(t)$ of the elevator before the k -th maintenance; s is the adjustment coefficient of maintenance time, which is a constant greater than 1; Δt_1 is the time spent on the first preventive maintenance.

When $K_{i+1} > \varepsilon_3$, it means that the $(i + 1)$ th maintenance should not be performed again. And the elevator life ends, it should be scrapped, that is, the elevator reaches the inflection point of retirement time; if $K_{i+1} < \varepsilon_3$, then go back to Sect. 3.2. It shows that the maintenance of the elevator at this time is more economical than direct scrapping.

4 Empirical Analysis

4.1 Sources of Data

Collected 1000 sets of sample data through an elevator operation monitoring system of a company in Harbin. Each set of data includes 17 indicators. Using the models established in Sects. 2 and 3, complete the forecast of the unscheduled maintenance intervals and scrap time points of the elevators in operation.

4.2 The Specific Results

The specific results are as follows: (1) 17 elevator life index values are predicted based on the BP neural network. The predicted termination point is the first preventive maintenance time point of the elevator. The sampling is performed every 2 h. Therefore, starting from the current monitoring time, the first preventive maintenance is required after 34 h. Based on the 1017 sample data and the related software, the specific function form of $R_{a_1}(t)$ is obtained:

$$R_{a_1}(t) = \exp(-(\frac{t}{1318})^{1.449})$$

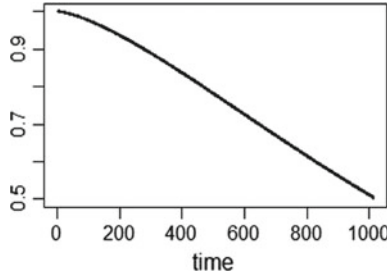


Fig. 1. Elevator safety remaining service life curve without maintenance

where $\alpha_1 = 1.449$, $\beta_1 = 1318$, and the fitting degree of the model reaches 0.9975, and the fitting effect is better. The safety remaining service life curve of elevator without maintenance is shown in Fig. 1.

(2) Then, under the premise of artificially given some empirical values, by giving a slope threshold and a technical threshold, the elevator will once again arrive at the point of preventive maintenance after 8 preventive maintenance. That is, the time point before the ninth preventive maintenance of the elevator, which is the elevator scrap point, and accurately gives the irregular maintenance interval between two adjacent preventive maintenance (as shown in Table 1);

Table 1. Distance from previous preventive maintenance and techno-economic indicator before this repair

Maintenance times	Distance from last maintenance (<i>h</i>)	Techno-economic indicator before maintenance	Maintenance time (<i>h</i>)
1	2032.000	0.674	2.000
2	415.850	1.585	2.400
3	320.489	2.439	2.880
4	251.215	3.338	3.456
5	198.908	4.339	4.147
6	158.509	5.485	4.977
7	126.862	6.819	5.972
8	101.832	8.386	7.166
9	81.913	10.238 > 10	—

Then, if the last sample monitoring time is assumed to be the current time, under the premise of normal maintenance, the elevator reaches the inflection point of retirement after 1722 h from the current moment, that is, the elevator’s safety remaining service life is 72 days. After 72 days, the elevator needs to be scrapped or replaced with a large part. According to the simulation results, the model for the safety remaining service life of the elevator under maintenance conditions is depicted, as shown in Fig. 2.

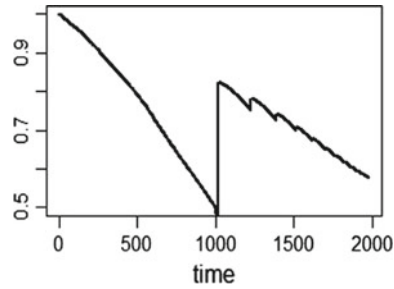


Fig. 2. Elevator safety remaining service life curve with maintenance

As can be seen from Fig. 2, with the increase in the number of preventive maintenance, the elevator performance after preventive maintenance has been restored, but the recovery value is getting lower and lower; moreover, the performance of the elevator is decaying faster and faster. Experience is consistent.

(3) Finally, by comparing with the results of the regular maintenance model, it is shown that the elevator's safety remaining life model established in this project has obvious advantages both in terms of economic benefits and risk prevention, and the model has practicality and reliability.

5 Conclusion

Based on the existing research, this topic establishes a model of the elevator's safety remaining service life which can determine the maintenance interval and predict the remaining service life. The main research results can be divided into the following aspects: the model of the remaining service life of the elevator without maintenance is established, the model of the remaining service life of the elevator under maintenance conditions is established, and the application of the model is completed.

References

1. Yujun, W.: Application of bathtub curve theory in civil aviation equipment management. *Commun. World* **17**, 280–281 (2017)
2. Liu, Q., Dong, M., Chen, F.F.: Single-machine-based joint optimization of predictive maintenance planning and production scheduling. *Robot. Comput.-Integr. Manuf.* **51**, 238–247 (2018)
3. Malik, M.A.K.: Reliable preventive maintenance scheduling. *AIIE Trans.* **11**(3), 221–228 (1979)
4. Jiangjiang, Z., Zichun, Y.: 3D Thermal fluid-structure coupling stress-strain analysis and life prediction of ship gas turbine blades. *Chin. Ship Chuan Res.* **5**(5), 64–68 (2010)
5. Zhu, Q.S., Gao, F., Ma, H.C., et al.: Failure behavior of flip chip solder joint under coupling condition of thermal cycling and electrical current. *J. Mater. Sci. Mater. Electron.* **1**, 1–9 (2017)

6. Mulubrhan, F., Mokhtar, A.A., Muhammad, M.: Fuzzy Activity Based Life Cycle Costing For Repairable Equipment 00007 (2016)
7. Chen, S.-T., Pao-Shan, Y.: Real time probabilistic forecasting of flood stages. J. Hydrol. **30**(1), 63–77 (2007)



Safety Evaluation Model of Intelligent Elevator Cloud Management Platform

Shang Sun, Zhenfeng Shi^(✉), and Qi Li

School of Science, Harbin Institute of Technology, 92 West Dazhi Street,
Nan Gang District, Harbin, Heilongjiang, China
szf@hit.edu.cn

Abstract. With the development of social economy and the improvement of people's living condition, elevator has become an indispensable tool because of its fast and convenient features. But the safety problem caused by elevator has brought great threat to people's life and property security, which aroused wide concern of the society. Therefore, developing the research of elevator safety evaluation has important social significance. The research of this paper mainly include the following three aspects: (1) The research of elevator safety evaluation index system. According to the operation data and the existing research results, we used the analytic hierarchy process to determine the multilevel structure of elevator safety evaluation index system. And both expert scoring method and entropy weight method were used to calculate the weight of index system, then the weight was combined by the integrated optimal method, to obtain the optimal weight of each index. (2) The research of elevator safety evaluation methods. First, the safety grade of whole elevator system and each index were divided. Because of the complexity and fuzziness of elevator system, we used the fuzzy mathematics and extension method, combining the matter-element theory. Then we used the correlation function to assess each index's security state. Finally, we established an extension fuzzy comprehensive evaluation model for elevator safety. (3) Example analysis of elevator safety evaluation methods. We took operation data of elevator in M city, for instance, analysis and MATLAB to establish safety evaluation program. The results show that the safety evaluation method is operable and effective.

Keywords: The elevator · Safety evaluation · Integrated optimal weighting method · Extension theory · Fuzzy mathematics

1 Introduction

With the rapid development of economy, elevator has been widely used as an indispensable vertical transportation facility because of its convenience and relative safety. At present, China's elevator market has already been at the top of the world. By the end of 2016, China's elevator ownership has reached 4.9369

million, with year-on-year growth of 15.9%. However, the safety problems of elevator have great threat to public safety, arousing wide concern of the society. According to the statistics of ACSIQ, there were 692 elevator accidents from 2002 to 2015, causing 527 deaths and 358 injured. And the average annual death result from elevator accident was about 40. Therefore, developing the new elevator safety evaluation method is of great significance to reduce the risk and provide elevator management scientific decision-making.

With the development of the safety standards and technical specifications, some scholars, institutions and universities carried out the research about safety evaluation method applying in elevator. K. A. M. Moinuddin and I. D. Bennetts established the elevator mechanical model, which test's and analyzes the strength of the structure of elevator steel components. It not only assessed some indexes such as elevator steel components, steel wire wheel force, and elevator shaft bearing, but also carried out risk assessment for the overall structure of elevator, which clarified the specification of elevator construction safety. I. A. Papazoglou and O. N. Aneziris build up the risk assessment model of elevator fallen, which introduced the method of quantitative analysis into the safety analysis of elevator safety accidents.

For different parts of elevator system, K. Kolowrocki established multiple reliability functions to calculate the value of the safe operation of the elevator. Then he used the maximum and minimum of elevator safely running to assess the elevator. Gu xuyi compared and summarized the safety evaluation method at home and abroad, and put forward the method based on risk principle which combined the accident tree analysis, safety checklist, and other methods. At the same time, he realized the function of the elevator safety comprehensive evaluation by using database technology to design the expert system. Zhang hao and Rui Yannian proposed new method to evaluate the actual performance of elevator using the set-valued statistics iterative method to establish the weight of the elevator evaluation system, which combined fuzzy mathematics method with extension theory. Also, they clearly divided the performance grade and index threshold value. Fang Yuchi and Yu feng considered the research about safety prevention and waning was less. Therefore, they put forward the method to assess elevators security levels by making use of fuzzy theory. It can not only objectively reflect the difficulty of elevator control, but also improve the practical application value of fuzzy theory. Finally, the improved fuzzy normalization model was used to establish the optimal control strategy for elevator.

2 The Research of Elevator Safety Evaluation Index System

The elevator is a complex and fuzzy dynamic system, whose safety performance is closely related to the subsystems in elevator. Therefore, the construction of safety evaluation index system is the basis of safety evaluation.

2.1 Establish the Index System

It is difficult to take all the information of the evaluation index into account when establishing the safety evaluation index system. Therefore, we establish a multilevel comprehensive evaluation index system containing 17 indexes at 3 levels with analytic hierarchy process.

2.2 Confirm Index Weight

Considering the difference that the contribution of each index to the safety of elevator, it is necessary to discuss and study the valuation method of the weight of index system. Therefore, we use the integrated optimal method combining the weight obtained by expert scoring method and entropy weight methods.

1. Determine the optimal weight form of integration. Assuming that there are l methods to determine the weight. The weight vector obtained by combining all the weighting methods is

$$W_c = \varphi_1 W_1 + \varphi_2 W_2 + \dots + \varphi_l W_l$$

Let $W = (w_1, w_2, \dots, w_l)^T$, $\psi = (\varphi_1, \varphi_2, \dots, \varphi_l)^T$, then $W_c = W\psi$.

2. Define comprehensive evaluation matrix. Defining the comprehensive evaluation matrix is $D = R \times W_c$, where R represents standardized data matrix, $D_i = \sum_{j=1}^m r_{ij} w_{cj}$, $i = 1, 2, \dots, n$.

3. Construct the sum of squares. Because its main idea is to maximize the total sum of squares so that making the comprehensive evaluation value the most dispersed, we denote the sum of squares between i th record and others as

$$v_i(W_c) = \sum_{i_1=1}^n \left[\sum_{j=1}^m (r_{ij} - r_{i_1 j}) w_{cj} \right]^2$$

Then, the total sum of squares is denoted as

$$\begin{aligned} J(W_c) &= \sum_{i=1}^n v_i(W_c) = \sum_{j_1=1}^m \sum_{j_2=1}^m \left[\sum_{i=1}^n \sum_{i_1=1}^n (r_{ij_1} - r_{i_1 j_1})(r_{ij_2} - r_{i_1 j_2}) \right] w_{cj_1} w_{cj_2} \\ &= W_c^T R_1 W_c = \psi^T W^T R_1 W \psi \end{aligned}$$

It is proved that the unit eigenvector ψ^* corresponding to the maximum eigenvalue of $W^T R_1 W$ is the optimal solution of the above equation.

4. Calculate the optimal weight

$$W_c^* = W\psi^* = (w_{c1}^*, w_{c2}^*, \dots, w_{cm}^*)$$

2.3 Simulation

In this section, we use the above method to calculate the index weight based on the actual data of the elevator in M city. The following are the simulation results, where $D_1 = (I_1, I_2, I_3, I_4, I_5)$, $D_2 = (I_6, I_7, I_8, I_9)$, $D_3 = (I_{10}, I_{11}, I_{12}, I_{13})$, $D_4 = (I_{14}, I_{15}, I_{16}, I_{17})$ and $D_5 = (C_1, C_2, C_3, C_4)$ (Table 1).

Table 1. The index weight

Index	Expert scoring method	Entropy weight method	Integrated optimal weighting method
D_1	(0.22, 0.22, 0.17, 0.12, 0.28)	(0.22, 0.14, 0.22, 0.22, 0.22)	(0.22, 0.18, 0.19, 0.17, 0.25)
D_2	(0.42, 0.29, 0.12, 0.17)	(0.26, 0.25, 0.25, 0.25)	(0.34, 0.27, 0.18, 0.21)
D_3	(0.37, 0.23, 0.13, 0.28)	(0.25, 0.25, 0.25, 0.25)	(0.31, 0.24, 0.19, 0.27)
D_4	(0.26, 0.19, 0.24, 0.32)	(0.25, 0.26, 0.23, 0.26)	(0.24, 0.19, 0.26, 0.31)
D_5	(0.16, 0.45, 0.11, 0.29)		

3 Research of Elevator Safety Evaluation Methods

At present, many scholars have developed and applied the safety evaluation technology in the elevator field, such as safety inspection table method, accident tree analysis, fuzzy comprehensive evaluation, and so on. Considering the fuzziness and complexity of elevator safety evaluation, we use the fuzzy comprehensive evaluation method combining matter-element model and extension theory to set up an extension fuzzy comprehensive evaluation model.

3.1 Construct Model

1. Safety classification

According to the practical experience, mathematical calculations and relevant standard, we classify the safety level of the elevator as safe, relatively safe, basic safety, relatively dangerous, and dangerous.

2. Determine classical and extensional matter elements

The classical matter elements refer to the range of data taken by each performance level with respect to corresponding features, denoted as

$$R_j = (N_j, c_i, v_{ij}) = \begin{pmatrix} N_j & c_1 & v_{1j} \\ & c_2 & v_{2j} \\ & \vdots & \vdots \\ & c_m & v_{mj} \end{pmatrix} = \begin{pmatrix} N_j & c_1 & [a_{1j}, b_{1j}] \\ & c_2 & [a_{2j}, b_{2j}] \\ & \vdots & \vdots \\ & c_m & [a_{mj}, b_{mj}] \end{pmatrix}$$

where N_j represents m safety level, c_i represents the index under safety level N_j , v_{ij} represents the value range of c_i with respect to N_j . The extensional matter elements is denoted as

$$R_p = (N_p, c_i, v_{ip}) = \begin{pmatrix} N_p & c_1 & v_{1p} \\ & c_2 & v_{2p} \\ & \vdots & \vdots \\ & c_m & v_{mp} \end{pmatrix} = \begin{pmatrix} N_j & c_1 & [a_{1p}, b_{1p}] \\ & c_2 & [a_{2p}, b_{2p}] \\ & \vdots & \vdots \\ & c_m & [a_{mp}, b_{mp}] \end{pmatrix}$$

which refer to the range of values taken by the whole of the performance level with respect to a feature. N_p represents the overall safety level, v_{ip} represents the value range of c_i with respect to N_p .

3. Determine the target element

The matter element to be evaluated is denoted as

$$R_0 = (N_0, c_i, y_i) = \begin{pmatrix} N_j & c_1 & y_1 \\ & c_2 & y_2 \\ & \vdots & \vdots \\ & c_m & y_m \end{pmatrix}$$

4. The index membership

In order to eliminate the influence caused by different dimensions of each index, we divide the elevator indexes into five safety levels, which are, respectively, [0, 0.2], [0.2, 0.4], [0.4, 0.6], [0.6, 0.8], [0.8, 1.0]. Then, the relative membership degree of each index is changed according to the linear relation.

5. Determine the correlation between the evaluation index and the safety level

Correlation degree $K_j(y_i)$ refers to the membership degree of each evaluation index with respect to the safety level j , which is denoted as

$$K_j(y_i) = \begin{cases} \frac{\rho(y_i, v_{ij})}{\rho(y_i, v_{ip}) - \rho(y_i, v_{ij})} & y_i \notin v_{ij} \\ \frac{-\rho(y_i, v_{ij})}{|v_{ij}|} & y_i \in v_{ij} \end{cases}$$

where $\rho(y_i, v_{ij}) = |y_i - \frac{1}{2}(a_{ij} + b_{ij})| - \frac{1}{2}(b_{ij} - a_{ij})$ denotes the distance between y_i and v_{ij} . $\rho(y_i, v_{ip}) = |y_i - \frac{1}{2}(a_{ip} + b_{ip})| - \frac{1}{2}(b_{ip} - a_{ip})$ denotes the distance between y_i and v_{ip} . $|v_{ij}| = |b_{ij} - a_{ij}|$ is called the value interval of the i th index in the j th safety level.

6. Comprehensive evaluation of safety level

According to the index weight calculated by integrated optimal weighting method, the correlation degree of the safety level to be evaluated with respect to the safety level j is denoted as

$$K_j(y) = \sum_{i=1}^m \alpha_i K_j(y_i)$$

If $K_{j_0} = \max(K_j(y))$, Then N_0 is rated as security level j_0 .

We introduce the concept of health degree, which not only reflects the degree of deviation from adjacent categories, but also describe the actual safety performance of elevators as an indicator variable. Denote:

$$\bar{K}_j(y) = \frac{K_j(y) - \min_j K_j(y)}{\max_j K_j(y) - \min_j K_j(y)}$$

Then the health degree of y is denoted as:

$$h = \frac{\sum_{j=1}^m j \cdot \bar{K}_j(y)}{\sum_{j=1}^m \bar{K}_j(y)}$$

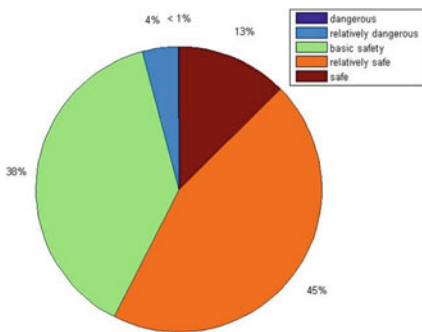
3.2 Simulation

In this section, we use the actual data of the elevator in M city to assess the safety level of elevator, combining the indexes weight calculated by Sect. 2.2. Taking Time 1 as example, we can figure out the safety evaluation information of each subsystem and integral of elevator.

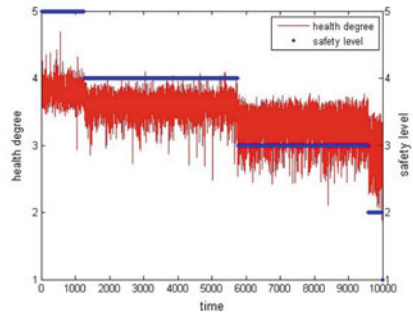
From Table 2, where T is Traction system, D is Door system, L is Lift car system and E is Electric traction system, we can calculate that the safety level of elevator is safe and its health degree is 3.9113 in Time 1. From (a), we can figure out that the distribution of elevator safety level, where the elevator operates normally in 95% time while having safety hazard in 5% time. From (b), we can conclude that the health degree of elevator is descending step by step, which is consistent with the change trend of safety level. In addition, most health degrees fluctuate between 3 and 4, which also verifies the effectiveness of safety rating evaluation results.

Table 2. Distance from previous preventive maintenance and techno-economic indicator before this repair

Index layer	Weight	Correlation matrix					Safety level
		Dangerous	Relatively dangerous	Basic safety	Relatively safe	Safe	
T	0.1644	-0.5435	-0.3764	0.3353	-0.3154	-0.4804	4
D	0.4457	-0.6452	-0.5415	0.5608	-0.5817	-0.4293	3
L	0.1051	-0.4447	-0.2178	0.3233	-0.4550	-0.5899	2
E	0.2848	-0.8636	-0.8247	0.7546	-0.5850	-0.2159	5



(a) The distribution of elevator safety level



(b) The safety level and health degree of elevator

References

1. Aneziris, O.N., Papazoglou, I.A., Baksteen, H., Mud, M., Ale, B.J., Bellamy, L.J., Oh, J.: Quantified risk assessment for fall from height. *Saf. Sci.* **46**(2), 198–220 (2008)
2. Arunraj, N.S., Maiti, J.: Risk-based maintenance techniques and applications. *J. Hazard. Mater.* **142**(3), 653–661 (2007)
3. Bennetts, I.D., Moinuddin, K.A., Goh, C.C., Thomas, I.R.: Testing and factors relevant to the evaluation of the structural adequacy of steel members within fire-resistant elevator shafts. *Fire Saf. J.* **40**(8), 698–727 (2005)
4. Cassa, A.M., Van Zyl, J.E., Laubscher, R.F.: A numerical investigation into the effect of pressure on holes and cracks in water supply pipes. *Urban Water J.* **7**(2), 109–120 (2010)
5. Chang-ming, G.X.Y.Z., Xiao-feng, Z.P.Z.: Research of comprehensive safety assessment method for elevator system. *China Saf. Sci. J. (CSSJ)* **6**, 028 (2008)
6. Chen, W., Xia, J.H.: An optimal weights combination method considering both subjective and objective weight information. *Math. Pract. Theory* **1**, 003 (2007)
7. Kołowrocki, K., : Asymptotic approach to reliability evaluation of rope transportation system. *Reliab. Eng. Syst. Saf.* **71**(1), 57–64 (2001)
8. Liu, D., Zou, Z.: Water quality evaluation based on improved fuzzy matter-element method. *J. Environ. Sci.* **24**(7), 1210–1216 (2012)
9. Saaty, T.L.: Decision making with the analytic hierarchy process. *Int. J. Serv. Sci.* **1**(1), 83–98 (2008)
10. Yao, R., Gao, Y., Xu, J.: Radio channel fingerprint model based on entropy weight method. In: 2016 25th Wireless and Optical Communication Conference (WOCC), pp. 1–4. IEEE (2016)
11. Zhang, H., Yan-nian, R.U.I.: Research on extension-fuzzy-based evaluation used for elevator performance. *Electrotech. Electr.* **8**, 12–15 (2009)
12. Zhang, X.Q., Liang, C.: Application of fuzzy matter-element model based on coefficients of entropy in comprehensive evaluation of water quality. *J. Hydraul. Eng.* **9**, 1057–1061 (2005)



A Chatbot Design Method Using Combined Model for Business Promotion

Jie Zhang, Hao Huang^(✉), and Guan Gui^(✉)

College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
{1017010502, guiguan}@njupt.edu.cn

Abstract. The combination of commercial development and artificial intelligence services becomes more and more important. Chatbot is considered one of the effective techniques by using information retrieval (IR) and natural language processing (NLP). In this paper, we collect a series of business-promoted chat data and conduct a series of cleanup and classification of these data sets. Since the speech of different people is random, the similarity is calculated by using a combination of the retrieval model and the generated model, and then the final answer is generated using long-short term memory (LSTM) training and prediction. Finally, we use the TF-IDF weighting method to improve the dialog. Experimental results show that the proposed method can communicate with humans and answer real-time questions.

Keywords: Chatbot · NLP · LSTM · TF-IDF

1 Introduction

In recent years, the rise in popularity of artificial intelligence (AI) has created a demand for services that help users to skip over some simple and boring works. One of the services is question answering (QA) using natural language processing (NLP) and spoken dialog system (SDS), which is a technique for providing answers to specific questions [1]. This services allows users to ask questions in natural language and will look up the user's question in the database [2] and returns the answer with the highest degree of matching. In addition, a long-short term memory (LSTM) based multi-layer embedding mechanism is used to extract the semantic information of the long sentences by client and customer service to provide recommendations for daily conversation.

The rest of this paper is organized as follows: Sect. 2 introduces the process model. Section 3 presents the experimental setup and results. Finally, Sect. 4 presents the conclusion and future works.

2 Process Model

Frequently asked questions (FAQs) is a common application customer service, in which questionnaires and corresponding answers are predefined in the database [3]. User’s query is compared with questions in existing FAQs to get a related answer to respond to the query. The “retrieval model” can complete this work [4]. However, the predefined database may not cover all the user questions. The “generate model” can respond more flexibly to the question of “open domain” [5]. This method does not need to manually process the database. Nevertheless, the model must be trained by a very large database, and the resulting sentence is difficult to maintain consistency and fluency (Fig. 1).

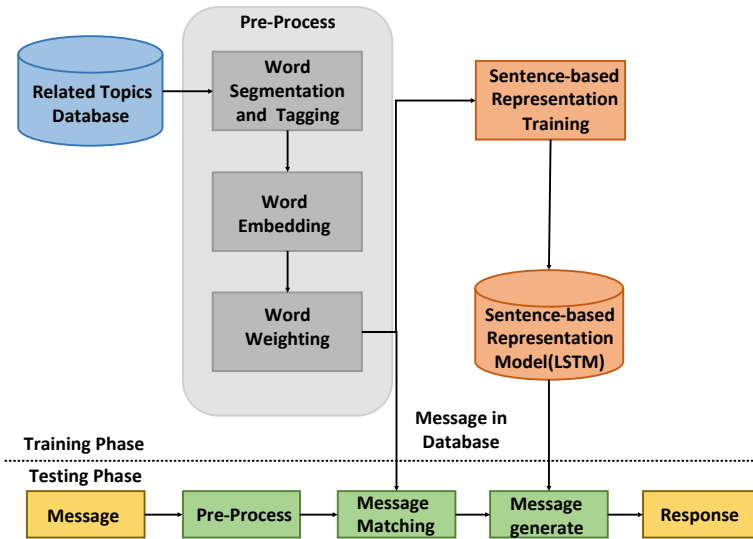


Fig. 1. The proposed framework.

2.1 Data Processing

From the statistical point of view, the input to the word segmentation problem is a string $C = [c_1, c_2, \dots, c_n]^T$ and the output is a word string $S = [w_1, w_2, \dots, w_m]^T, m \leq n$. For a particular string C , there will be multiple segmentation schemes S . The task of segmentation is to find the most probable segmentation scheme among these S , that is, to segment the input string into the most likely words and sequence as

$$\text{Seg}(c) \underset{S \in G}{\text{argmax}} P(S|C) = \underset{S \in G}{\text{argmax}} \frac{P(C|S)P(S)}{P(C)} \tag{1}$$

If there are two segmentation possibilities S_1 and S_2 , calculate the conditional probabilities $P(S_1|C)$ and $P(S_2|C)$, and then a segmentation scheme corresponding to a large probability value is adopted. According to the Bayesian formula

$$P(S|C) = \frac{P(C|S) \times P(S)}{P(C)} \quad (2)$$

where $P(C)$ is the probability that the string appears in the corpus and just a fixed value that is used to normalize. There is only one way of the probability of recovering from a word string to a Chinese character string, so $P(C|S) = 1$. Hence, the sizes of $P(S_1|C)$ and $P(S_2|C)$ are compared to compare the sizes of $P(S_1)$ and $P(S_2)$.

The task of the word segmentation of the probabilistic language model is to find a segmentation scheme S in all the results obtained by the full segmentation so that $P(S)$ is maximized. For simplicity, we assume that the probability between each word is context-independent

$$\begin{aligned} P(S) &= P(w_1, w_2, \dots, w_m) \approx P(w_1) \times P(w_2) \times \dots \times P(w_m) \\ &\propto \log P(w_1) + \log P(w_2) + \dots + \log P(w_m) \end{aligned} \quad (3)$$

For different S , the value of m is not the same, in general, the larger the m , the smaller $P(S)$ will be. In other words, the more words are dropped, the probability will be smaller. This is in line with practical observations, such as the maximum length matching segmentation tends to make m smaller. The probability of calculating any word appears as follows:

$$P(w_i) = \frac{n_{w_i}}{N} \quad (4)$$

where n_{w_i} means the number of occurrences of w_i in the corpus and N means the total number of words in the corpus.

2.2 Word Embedding

Word embedding can transform One-Hot Encoder into low-dimensional continuous values, which is dense vectors, and words with similar meanings will be mapped to similar positions in the vector space [6]. The two most important models used in word embedding are the CBOW and Skip-gram models. Here we introduce these two models separately and a weighting technique to serve later work.

2.2.1 CBOW Model and Skip-Gram Model

The CBOW full name is the Continuous Bag-of-Words Model, which predicts the current word on the premise of knowing the current word context. CBOW model is shown in Fig. 2a. The neural network structure of the CBOW model is designed as follows:

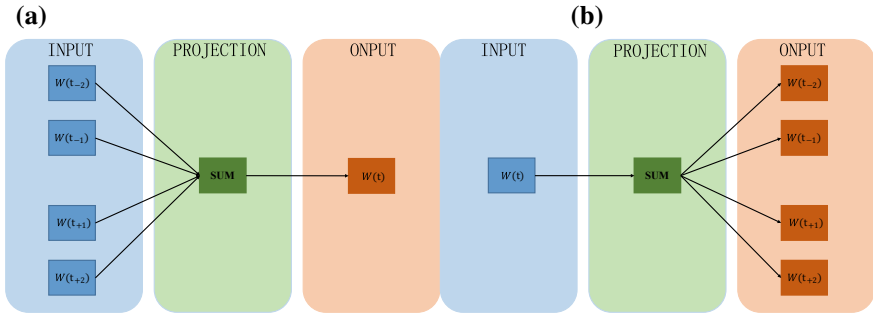


Fig. 2. a CBOW model; b skip-gram model.

Input layer: the context of the word W has $2c$ word vectors.

Projection layer: sum the $2c$ vectors of the input layer.

Output layer: a Huffman tree where the leaf nodes are words that appear in the corpus and the weight is the number of occurrences.

This design changes the end-to-end connection into a summation sum, which reduces the dimensions; it removes the hidden layer, which reduces the amount of calculations.

The Skip-gram full name is the continuous skip-gram Model, which predicts the context when the current word is known. skip-gram model is shown in Fig. 2b. The neural network structure of the skip-gram model is designed as follows:

Input layer: W word vector $v(w)$.

Projection layer: still $v(w)$, just a form.

Output layer: Huffman tree like CBOW.

2.2.2 TF-IDF

TF-IDF is called term frequency-inverse document frequency, which is a commonly used weighting technique for information retrieval and information mining. TF-IDF is a statistical method for assessing the importance of a word for one document set or one of the documents in a corpus. Term frequency (TF) refers to the number of occurrences of a given word in the file. This number is usually normalized.

$$tf_{ij} = \frac{m_i}{M} \tag{5}$$

where tf_{ij} is the number of features appearing in the document, m_i is the number of occurrences of a term w in a class. M is the number of all entries in the class. The main idea of inverse document frequency (IDF) [7] is that if the number of documents containing the term t is smaller and the IDF is larger, indicates that the term has a good category distinguishing ability.

$$idf_i = \log\left(\frac{N}{n_i + 1}\right) \tag{6}$$

where idf_i refers to the reciprocal of the feature item, N is the total number of documents in the corpus and n_i is the number of documents containing the term w . The denominator in the formula is increased by 1 is to avoid that the denominator is 0.

The high frequency of words in a particular file, and the low file frequency of that word in the entire file set, which can produce a high weighted TF-IDF. Therefore, TF-IDF tends to filter out common words and retain important words [8].

$$w_{ij} = tf_{ij} \times idf_i = \frac{m_i}{M} \times \log\left(\frac{N}{n_i + 1}\right) \quad (7)$$

Taking into account that the length of document contents will affect the weight calculation, we use normalized processing and get the following formula:

$$w_{ij} = \frac{tf_{ij} \times \log\left(\frac{N}{n_i + 1}\right)}{\sqrt{\sum_{l=1}^N tf_{ij}^2 \times \log^2\left(\frac{N}{n_i + 1}\right)}} \quad (8)$$

2.3 LSTM Model

In order to solve the problem of disappearance gradient and explosion gradient are the main motivation behind the introduction of a new structure [9] called the storage unit LSTM model. The structure is shown in Fig. 3. The memory cell consists of four main components: the input gate, the neurons with autoregressive connections (connected to itself), the forget gate and the output gate. The autoregressive connection has a weight of 1.0 and ensures that the state of the memory cell can remain constant from a time step to another, with the exception of external disturbances. The gate is used to adjust the interaction between the memory unit itself and its environment. The input gate can allow the input signal to change the state of the memory cell or prevent it. On the other hand, the output gate may allow the state of the memory cell to have an effect on other neurons or prevent it. Finally, the forget gate can adjust the self-reverting connection of the storage unit so that the unit remembers or forgets its previous state as needed.

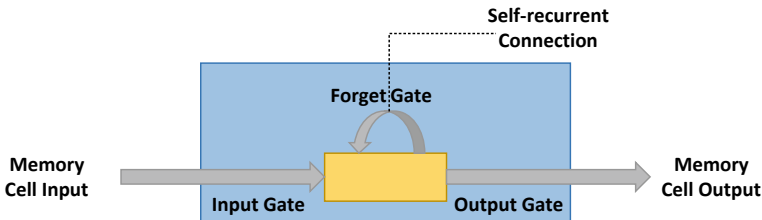


Fig. 3. LSTM storage unit.

The following formula describes how to update a layer of memory cells t at each time step [10]. In these equations, x_t is the input to the memory cell layer, $W_i, W_f, W_c, W_o, U_i, U_f, U_c, U_o$ and V_o are the weight matrix. b_i, b_f, b_c , and b_o are the offset vector. First, we calculate the value i_t of the memory cell state over time, input gate and candidate \widetilde{C}_t :

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (9)$$

$$\widetilde{C}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (10)$$

Second, we calculate the activation value of the forget gate at time in the memory cell f_t

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (11)$$

Given the input gate activation value i_t , forgetting the gate activation value f_t and the candidate state value \widetilde{C}_t , we can calculate the new state C_t of the memory cell over time:

$$C_t = i_t \times \widetilde{C}_t + f_t \times C_{t-1} \quad (12)$$

With the new state of memory cells, we can calculate the values of the output gates and the subsequent output

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + V_o C_t + b_o) \quad (13)$$

$$h_t = o_t \times \tanh(C_t) \quad (14)$$

3 Experimental Results

The purpose of the experiment is to complete a commercial automated Chatbot. This system can give accurate answers to common questions asked by users and respond quickly, besides can generates answers for specific sentences.

3.1 Performance with Combined Model

After many experiments, we found that when the input sentence finds the corresponding sentence with a matching degree of 0.8 or more in the database through the retrieval model, the answer can be output as the final answer (Fig. 4).

```

input sentence: 你好
text 1205: 你好, score : 1.0
您好我是南京首屏的您考虑做百度推广吗

input sentence: 你们这个怎么做
text 98: 你们 是 怎么 做 的, score : 0.842167612636
就是把您公司的产品或者服务展示在百度首页上, 比如您是做搬家的, 客户在百度上搜索“搬家”这个词, 就可以将您公司展示在首页

input sentence: 请问怎么收费
text 1093: 怎么 收费, score : 0.870075646043
每年服务费两千四最低充值六千一共八千四

```

Fig. 4. Retrieval model answer some questions.

(a)

```

input sentence: 还有其他服务吗
text 307: 免费 吗, score : 0.482983659661
text 2461: 还有 周末 口碑, score : 0.445892061182
text 1897: 你们 提供 哪些 服务, score : 0.436170974398
text 1108: 有关键词 包年 推广 的 服务 吗, score : 0.435558778408
text 2696: 有关键词 包年 推广 的 服务 吗, score : 0.435558778408
text 3054: 那个 代理 公司 吗, score : 0.397825736313
text 119: 效果 还行 吗, score : 0.390580910414
text 1779: 我们 其他 的 合作, score : 0.375337055096
text 2762: 我们 其他 的 合作, score : 0.375337055096
text 41: 百度 还是 其他, score : 0.366241641594
text 1112: 百度 还分 区域 吗, score : 0.365098741322
text 2598: 百度 还分 区域 吗, score : 0.365098741322
text 82: 客服 态度 不好 账户 维护 服务 跟不上, score : 0.363500712151
text 1: 我们 可以 先 试用 吗, score : 0.362124595777
text 1894: 其他 平台 你们 也 做 啊, score : 0.341639488525
text 2527: 其他 平台 你们 也 做 啊, score : 0.341639488525
text 103: 做 了 推广 能 保障 给 我们 带来 的 收益 吗, score : 0.32891184662
text 1110: 按年 收费 吗, score : 0.32055590827
text 121: 效果 怎么样, score : 0.31805803675
text 122: 效果 怎么样, score : 0.31805803675
text 1019: 包年 什么 价格 呢, score : 0.313327412003
text 120: 效果 好 吗 访问量 大不大, score : 0.309514198686
text 1073: 我们 是 预 交 钱 吗, score : 0.307271883626
text 2670: 我们 是 预 交 钱 吗, score : 0.307271883626
text 3178: 我们 里面 还有 钱 的 吧, score : 0.307010179036

```

(b)

```

input sentence: 还有其他服务吗
不好意思 我们 现在 没有 免费 试用

```

Fig. 5. **a** The degree of matching of a particular question in a document; **b** generated model gives a similar answer.

When the input matching degree is lower than 0.8, the generation model is used to complete the question and answer. The questions shown are ranked from high to low (Fig. 5).

3.2 Performance with TF-IDF

We have found that the answers to certain questions are not ideal and feature extraction is insufficient. TF-IDF is used to improve the performance (Fig. 6).

(a)

```

input sentence: 我需要
text 629: 我 不 需要, score : 0.868709975511
好的后面有需要的再联系

```

(b)

```

input sentence: 我需要
text 8815: 需要 什么, score : 0.817195085923
我们南京首屏主要是做百度推广的

```

Fig. 6. **a** The answer given under the usual weighting method; **b** the answer given under the TF-IDF.

For the traditional weighting method, the word “need” is identified as “don’t need” in the database with a large number of occurrences, thereby outputting an error reply. TFIDF technology improves this phenomenon.

4 Conclusion

The main task of this paper is to propose a combination of retrieval model and generative model, and to improve the weighting aspect makes the characteristics of each sentence more obvious. This model is mainly used in Chatbot for commercial promotion. Then the model will continue to be refined, making the response more persuasive and adaptable to more scenarios.

References

1. Vargas-Vera, M., Lytras, M.D.: AQUA: hybrid architecture for question answering services. *IET Softw.* **4**(6), 418–433 (2010)
2. Abdul-kader, S.A.: Question answer system for online feedable new born chatbot. In: 2017 Intelligent Systems Conference (IntelliSys), pp. 863–869 (2017)
3. Su, M.-H., Wu, C.-H., Huang, K.-Y., Hong, Q.-B., Wang, H.-M.: A chatbot using LSTM-based multi-layer embedding for elderly care. In: 2017 International Conference on Orange Technologies (ICOT), pp. 70–74 (2017)
4. Hu, B., Lu, Z., Li, H., Chen, Q.: Convolutional neural network architectures for matching natural language sentences. In: *Advances in Neural Information Processing Systems*, vol. 27, pp. 2042–2050 (2014)
5. Quarteroni, S., Manandhar, S.: Designing an interactive open-domain question answering system. *Nat. Lang. Eng.* **15**(1), 73–95 (2009)
6. Grangier, D., Keshet, J., Bengio, S.: Discriminative keyword spotting. *Speech Commun.* **51**(4), 317–329 (2009)
7. Spärck Jones, K.: A statistical interpretation of term specificity and its retrieval. *J. Doc.* **28**(1), 11–21 (1972)
8. Guo, A., Yang, T.: Research and improvement of feature words weight based on TFIDF algorithm. In: *Proceedings of 2016 IEEE Information Technology, Networking, Electronic and Automation Control Conference (ITNEC 2016)*, pp. 415–419 (2016)
9. Hochreiter, S., Schmidhuber, J.J.: Long short-term memory. *Neural Comput.* **9**(8), 1735–1780 (1997)
10. Graves, A., Jaitly, N.: Towards end-to-end speech recognition with recurrent neural networks. In: *International Conference on Machine Learning*, pp. 1764–1772 (2014)



Design of Monitoring and Warning System for Dangerous Gases in Oil Tank

Yuelan Ji^(✉), Yongjie Yang^(✉), and Zhongxing Huo

School of Electronics and Information, Nantong University, Nantong 226019, China

1179121631@qq.com, yang.yj@ntu.edu.cn

Abstract. Aiming at the hidden danger of harmful gas leakage and insufficient oxygen supply in the operating environment of oil tanks, this paper designs and implements monitoring and warning system for dangerous gas in the oil tank. The system uses the multilayer distributed structure, combines with various gas detection sensor technology and wireless communication technology, and adopts the STM32 microcontroller based on ARM Cortex-M3, with LoRa spread spectrum MESH ad hoc network module, 4G module, and OLED display module, and so on, to realize the collection and transmission of the gas concentration data in the oil tanker and display them on the upper computer software in real time. According to the field test of Zhongyuan shipping Automation Co., Ltd., the system meets the design requirements. It has the characteristics of low cost, flexible distribution, safe and practical and so on, which has a very good value for promotion.

Keywords: Oil tank monitoring · Gas collection · Wireless ad hoc network · Internet of things · MCU · UCOSII

1 Introduction

In the tanker cabins, there are always some harmful gases, such as carbon monoxide, hydrogen sulfide, and CH_4 , that will directly affect the safe operation of workers and endanger their health. In addition, due to the chemical reaction such as metal corrosion, bulkhead layer aging, and other chemical reactions, the natural respiration of grain and other goods will consume oxygen. In the closed oil cabins, air does not circulate for a long time, resulting in a lack of oxygen, which is also an important cause of casualties [1–3]. At present, most of the hand-held oxygen and explosive detectors on the market are expensive, requiring more manpower for a large space, which can not meet the requirement of timeliness. And the dangerous gas monitoring and warning system designed in this paper is not only low cost, saving manpower, but also can be monitored remotely, which is safe and reliable.

The monitoring and warning system for dangerous gas designed in this paper, which can monitor the concentration of dangerous gas and oxygen concentration in the operating environment in real time. When the concentration of dangerous gas exceeds the warning value and the oxygen concentration is lower than the warning value, the on-site operation personnel and the monitoring center can receive the alarm signal at

the same time in order to avoid the accident, which is also a highlight of this paper. In addition, the system can also display the gas concentration data of each cabin node on the upper computer. When the gas concentration is abnormal, it will immediately issue an alarm and display the alarm cabin node number and gas concentration value.

2 Overall Design of System

2.1 The Composition and Structure of the System

The architecture of the monitoring and warning system for dangerous gas in the oil tank is shown in Fig. 1. The system adopts the multi-layer distributed structure. From the bottom up, they are the terminal layer, the transmission layer, and the monitoring layer. Among them, the transmission layer is the communication bridge of the whole system, which plays the role of coordinating the data transmission. The terminal layer is composed of multiple nodes distributed in different compartments. Each monitoring node is responsible for collecting the concentration of carbon monoxide, hydrogen sulfide, oxygen, and flammable gas in the current cabin and the power of the nodes, then transmit the current data frame through the LoRa wireless transmission system based on 433 MHz [4]. The oil tank's data relay node transmits data through 4G to the server in order to realize uploading data in real time. After the server parses the received data frames, the data will be stored in the database. The upper computer software is designed based on Qt, which can graphically display the data of each node, historical query, and set thresholds.

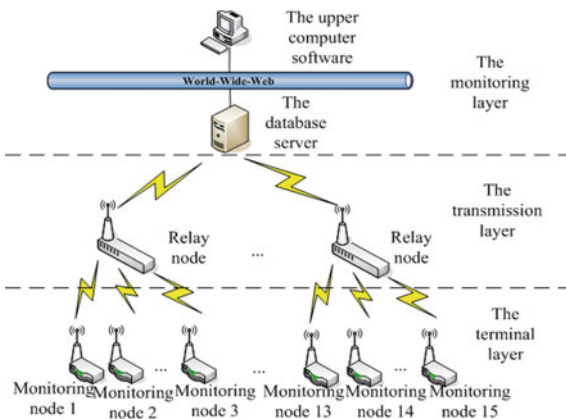


Fig. 1. The overall structure of system

3 Hardware Design of System

The hardware of monitoring and early warning system for dangerous gas in the oil tank includes monitoring nodes and relay nodes. Their hardware design mainly includes the hardware design of the main control unit, the hardware design of the data acquisition unit, and the hardware design of the data transmission unit.

3.1 Hardware Design of the Main Control Unit

Both monitoring nodes and relay nodes choose STM32F103ZET6 based on Cortex-M3 of ST company, whose working frequency is as high as 72 MHz [5]. And it has 51 common fast I/O ports. The chip's code execution efficiency is also high. So its configuration satisfies the hardware requirements of the peripheral circuit. Its peripheral hardware circuit includes power supply circuit, reset circuit, and download circuit. Both monitoring nodes and relay nodes choose 12 V voltage power supply. The voltage of 3.3 V and 5 V should be supplied to each module, respectively. The voltage of 5 V is obtained by LM2596S-5.0, and then the 3.3 V voltage is obtained through the conversion of AMS117-3.3 chip. The reset circuit is reset by key and the low level is effective.

3.2 Hardware Design of the Data Acquisition Unit

3.2.1 Hardware Design of Sensor Acquisition

The monitoring node is equipped with four gas sensors, which are carbon monoxide sensor, hydrogen sulfide sensor, oxygen sensor, and methane sensor, and they are responsible for collecting the gas concentration of carbon monoxide, hydrogen sulfide, oxygen, and CH₄ in the cabin [6, 7]. The accuracy of the collected gas concentration information is related to the performance of the sensor, and the selection of the gas sensors and the characteristics of the products are illustrated as shown in Table 1. The gas reacts chemically on the electrodes of sensors, resulting in generating charges, that can realize gas detection. The power supply, current-voltage conversion and compensation circuits of each sensor are integrated on a circuit board so that they all have excellent electromagnetic interference resistance. When the concentration is 0, the output voltage is 0, and the gas concentration in the linear range is proportional to the voltage output.

Table 1. The selection of the gas sensors and the characteristics of the products

Name	Model	Measurement range	Resolution	Response time (s)
Carbon monoxide sensor	CLE-0052-400	0–500 ppm	1 ppm	≤30
Hydrogen sulfide sensor	CLE-0112-400	0–100 ppm	0.1 ppm	≤30
Oxygen sensor	40XV	1–25% vol	0.2% vol	<15
CH ₄ sensor	4P90C	0–100% LEL	1% LEL	<20

The microcontroller obtains the voltage signal of each sensor through AD conversion [8]. Through the corresponding transformation formula of each sensor, the voltage is converted to the actual value. The schematic diagram of the sensor acquisition circuit is shown in Fig. 2.

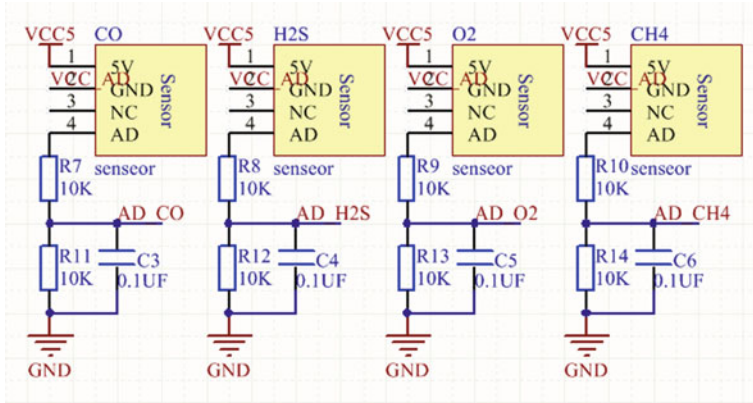


Fig. 2. The schematic diagram of the sensor acquisition circuit

3.2.2 Hardware Design of Power Acquisition

The schematic diagram of node power acquisition is shown in Fig. 3. The 12 V voltage is divided into 3.3 V through the resistance. This method is not accurate enough but simple and easy; because of collecting the division of voltage, there may be voltage division exceeding the safe voltage of microcontroller. So the division of voltage should be passed through LM358 that is used as a voltage follower to further protect the pin of the microcomputer.

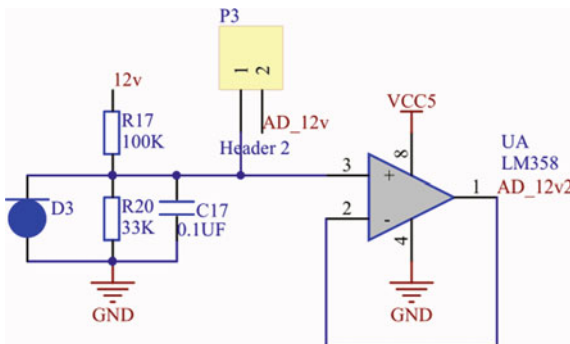


Fig. 3. The schematic diagram of node power acquisition

3.3 Hardware Design of Data Transmission Unit

The data transmission between the monitoring node and the relay node chooses YL-800MN-100 mW wireless serial port module. It is a micro-power wireless MESH networking module with high-performance, low power, long distance based on Sx1278 LoRa spread frequency hopping technology. It is embedded with wireless MESH ad hoc network protocol and it has excellent network self-healing, stability, and excellent data throughput. It supports seven-level routing, and the network coverage is more than 7 km (Air speed is 12 kbps). And in the environment of an oil tank, the penetration of 433 MHz is better. The module has a sensitivity of -148 dBm and a maximum transmitting power of $+20$ dBm, which can automatically adjust according to the actual situation [9–11]. Its working voltage is 4.5–5.5 V and the power consumption is lower than 100 MW. It is suitable for battery-powered occasions. The microcontroller configures the module through the serial port to support data transmission through transmission mode.

4 Software Design of System

4.1 Software Program Design of Nodes

The program of monitor nodes and relay nodes are based on UCOSII operating system. It is a real-time operation system which is tailored, preemptive and with real-time multitasking kernel that can run on ROM [12, 13]. It has high portability and is especially suitable for a microprocessor with limited space. After the node is on power, the main function calls the `setup()` function to complete the initialization of each module, invokes `OSInit()` to initialize the UCOSII, creates the task by calling the `OSTaskCreate()` function, and finally calls `OSStart()` to start the multitask scheduling of the UCOSII. The program flowchart of the monitoring node is shown in Fig. 4. The program flowchart of the relay node is shown in Fig. 5.

4.2 The Software Design of the Upper Computer

The upper computer software runs on the Windows system, takes Qt as the development environment, and adopts the C++ language to develop and design. The upper computer software monitors the real-time data information of each node, analyzes the received data information through Json, enters the SQLITE database after analysis, and then sends it back to every node on the interface. The functions of the upper computer's displaying interface are divided into:

1. Real-time data display: The gas concentration and node power of each node can be more intuitively understood through the column diagram;
2. Set Threshold: The staff can individually set any gas concentration of each node individually;
3. Curve display: It is possible to observe the trend of data information;
4. Historical data query: The staff can view and analyze historical data of the last 6 months analysis in the generated excel table.

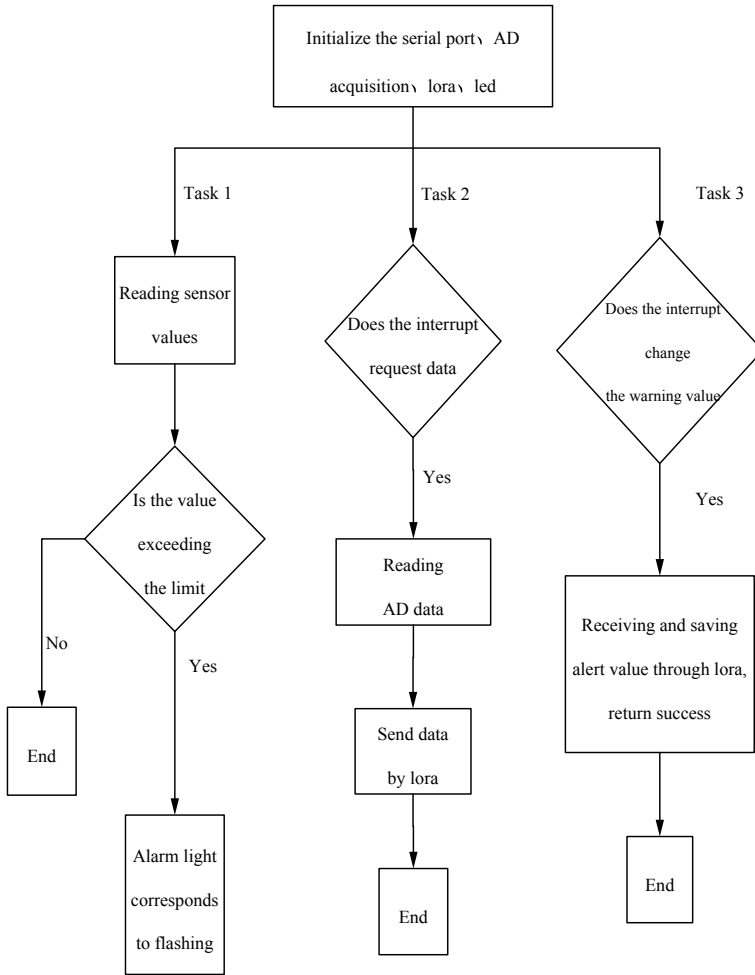


Fig. 4. The program flowchart of monitoring node

4.3 MESH Communication Data Frames

The relay node sends the task of requesting data to the monitoring node, thus the monitoring node sends data to the relay node [14]. The MESH data frame format sent by relay node is shown in Table 2.

The frame type is used to identify different types of application frames. 05 is the application data frame, which is used to the interface of the networking protocol application layer. The frame number is fixed to 00. The command type 82 represents the response of request. The load length refers to the length of the frame load part in the frame format, and the load length of the frame is 7 byte. In the data section, 00 01 represents the number of relay node, power indicates the power of the current relay node sending data frames, and 03 “r” “e” “q” indicates that 3 bytes of data are “r” “e”

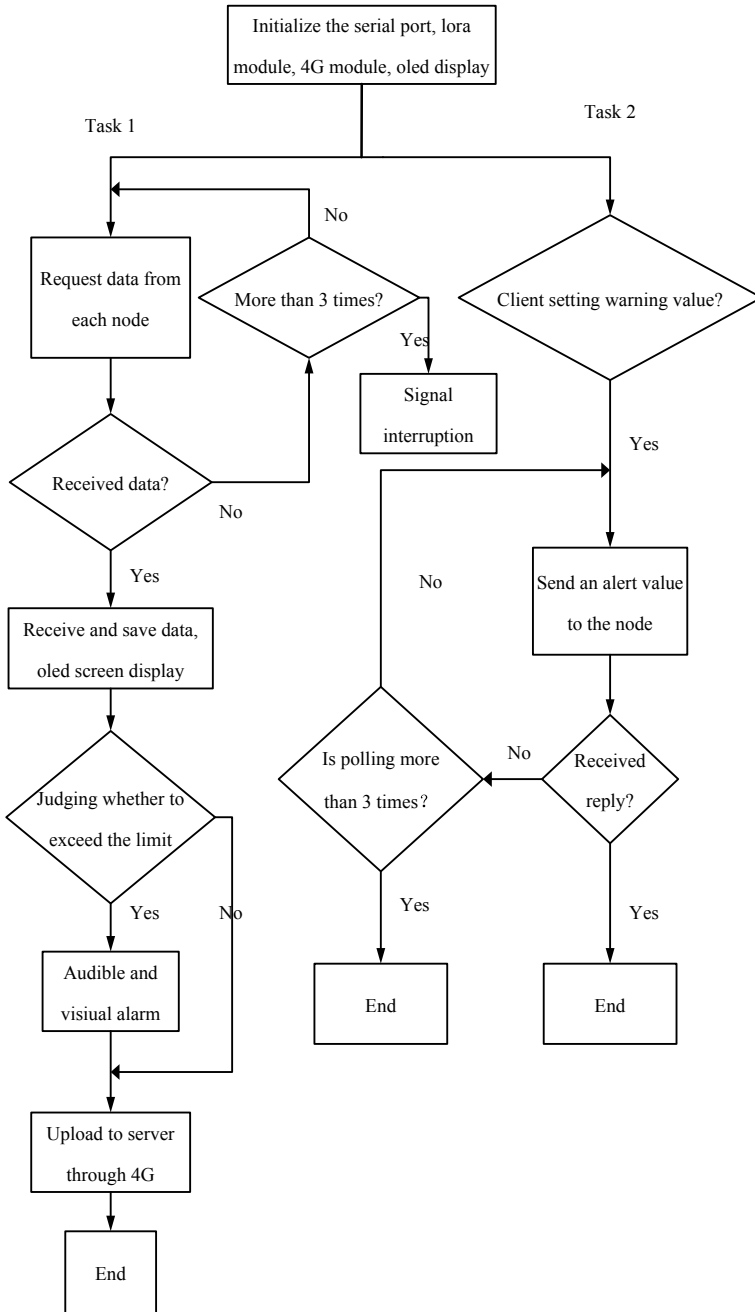


Fig. 5. The program flowchart of relay node

Table 2. MESH data frame format for relay transmission

Frame type	Frame number	Command type	Load length	Data	Check bit
1 byte	1 byte	1 byte	1 byte	N byte	1 byte
05	00	82	07	00 01 power 03 “r” “e” “q”	Check bit

“q”, “r” “e” “q” used to identify request sensor values. The check bit is the XOR result of all bytes starting from the first byte of the frame (frame type byte) to the check byte.

The data frames returned are: 05 00 01 12 00 01 00 07 01 0C “d” “s” CO H2S O2 CH4 ele stat check bit.

Among them, 05 means sending user data; 00 is a fixed value; 01 represents an application data transmission request; 12 refers to the number of bytes behind the load length to the byte ahead of check bit; 00 01 means that the target address is 00 01; 00 means no ACK response; 07 means the maximum level jump is 7; 01 indicates that routing is automatic routing; 0C indicates that there are 12 bytes of data to be sent; “d”, “s”, co, h2s, o2, ch4, ele, stat are user data, they are a total of 12 bytes; “d” “s” is used to identify sending sensor values, and CO H2S O2 CH4 occupies 2 bytes, respectively.

5 The Test of System

5.1 The Test Results

The system platform is shown in Fig. 6, with three monitoring nodes on the left and relay node on the right. The system is tested in a shipping Automation Co., Ltd. in Nantong. The monitoring nodes are placed in three corners of the tanker. The relay node is placed at the cabin door, and the magnet alarm lamp is absorbed on the cabin door. The upper computer is monitored and managed by professionals. The display interface is shown in Fig. 7. From Fig. 7, the monitoring nodes of the system can work stably with no interference between each other and can upload the gas concentration and power of each node to the server. The system has achieved the requirements and realized the function of the system, which is a successful application of intelligent data acquisition in the environment of an oil tank.

5.2 Early Warning Test

Taking node 1 as an example, the oxygen concentration around the oxygen sensor head is changed by a lighter. From the OLED display on the relay node, the oxygen measurement of node 1 is 16 and flashing display. The magnet-type warning light makes a sound and light alarm, as shown in Fig. 8. The display interface of the upper computer software is shown in Fig. 9, and the column graph of the O₂ of node 1 becomes red, while the instant status shows the oxygen abnormality of the node 1.

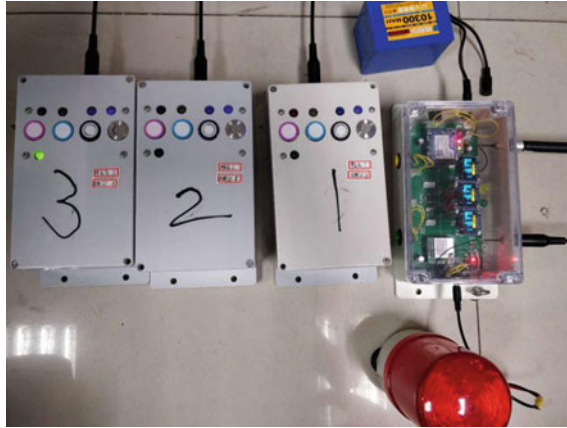


Fig. 6. The system platform

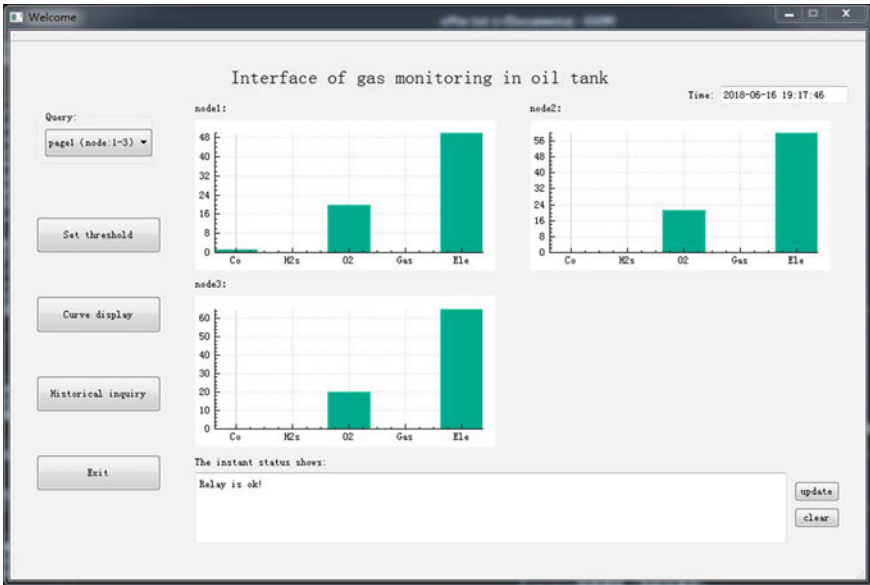


Fig. 7. The display interface

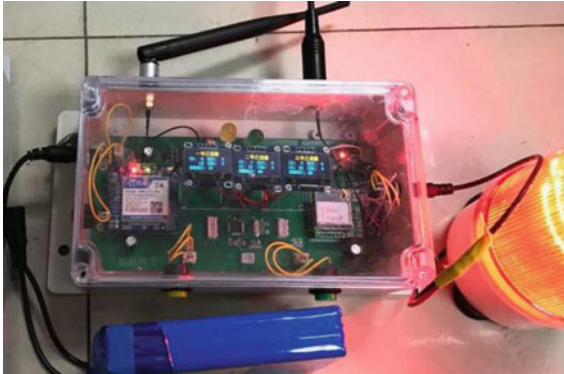


Fig. 8. Alarm chart of relay node

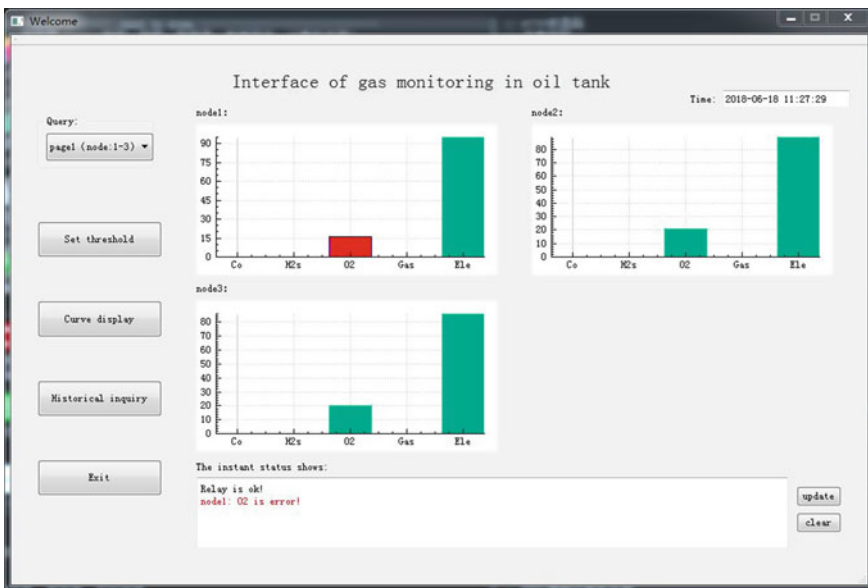


Fig. 9. The display interface of the upper computer

6 Conclusions

The monitoring and early warning system for dangerous gases in oil tank designed in this paper realize the integration of embedded technology, sensor technology, LoRa wireless technology, and upper computer software technology. The hardware and software system design of the monitoring node and relay node adopt low power processor STM32 as the core. Through repeated debugging of the system, the design of the monitoring and early warning system for dangerous gases in the oil tank is successful. The structure of the system is simple and reliable, and the arrangement of the nodes is

flexible. The system adopts wireless communication to transmit data to realize the intelligent early warning function. The system can not only be applied into the cabin gas monitoring but also can be widely used in coal mines, chemical plants, oil refining and other scenes, which have wide application prospects.

Acknowledgements. This work was the first phase project of Jiangsu University Brand Specialty Construction Project (PPZY2015B135). In addition, it was completed under the support of Nantong University-Nantong Intelligent Information Technology Joint Research Center Open Topic (KFKT2017B05). The authors thank the 3 anonymous reviewers for their helpful suggestions.

References

1. Weli, V.E., Itam, N.I.: Impact of crude oil storage tank emissions and gas flaring on air/trainwater quality and weather conditions in Bonny Industrial Island, Nigeria. *Open J. Air Pollut.* **05**(2), 44–54 (2016)
2. Xu, Y., Yang, K., Luo, Y., et al.: Cyanobacteria bloom monitoring and early warning system based on GIS and WSNs—a case study in Dianchi Lake. In: *International Conference on Geoinformatics*, pp. 1–8. IEEE (2016)
3. Sendari, S., Rahmawati, Y., Kamdi, W., et al.: Internet-based monitoring and warning system of methane gas generated in garbage center, 2018:012078 (2018)
4. Zhou, L., Sun, S., Zhang, Y., et al.: Long-distance running test system based on 433 MHz wireless module. In: *IEEE International Conference on Communication Technology*, pp. 339–343. IEEE (2016)
5. Yang, B.: Design and implementation of intelligent home wireless gateway based on STM32. In: *International Conference on Information Science and Control Engineering*, pp. 258–260. IEEE Computer Society (2017)
6. Sultangazin, A., Kusmangaliyev, J., Aitkulov, A., et al.: Design of a smartphone plastic optical fiber chemical sensor for hydrogen sulfide detection. *IEEE Sens. J.* **PP**(99), 1–1 (2017)
7. Robinson, D.B., Lorenzo, A.P., Macrygeorgos, C.A.: The carbon dioxide-hydrogen sulphide-methane system: Part II. Phase behavior at 40°f and 160°f. *Can. J. Chem. Eng.* **37**(6), 212–217 (2015)
8. Minhas, S., Khanum, A., Riaz, F., et al.: A non parametric approach for mild cognitive impairment to AD conversion prediction: results on longitudinal data. *IEEE J. Biomed. Health Inform.* **PP**(99), 1–1 (2016)
9. Salameh, H.B.: Spread spectrum-based coordination design for spectrum-agile wireless ad hoc networks. *J. Netw. Comput. Appl.* **57**(C), 192–201 (2015)
10. Wan, X.F., Yang, Y., Du, X., et al.: Design of propagation test node for LoRa based wireless underground sensor networks. In: *Progress in Electromagnetics Research Symposium—Fall*, pp. 579–583 (2017)
11. Usmonov, M., Gregoretti, F.: Design and implementation of a LoRa based wireless control for drip irrigation systems. In: *International Conference on Robotics and Automation Engineering*, pp. 248–253. IEEE (2018)
12. Zheng, S., Wang, D., Zheng, R., et al.: Research on the detection system for impact energy of pneumatic drill based on stress wave technique. In: *IEEE International Conference on Electronic Measurement & Instruments*, pp. 1415–1419. IEEE (2016)

13. Cai, G.H., Liang, C., Huang, Q., et al.: A remote intelligent control system based on zigbee wireless technology. *Appl. Mech. Mater.* **644–650**, 71–75 (2014)
14. Pradhan, S.C., Mallik, K.K.: Minimization of overhead using minislots allocation algorithm in IEEE 802.16 mesh network. In: *Fifth International Conference on Eco-Friendly Computing and Communication Systems*, pp. 68–72. IEEE (2017)



Design of Standard Cell for Anti-radiation

Bei Cao^(✉), Pengfei Wu, and Danyang Qin

Electronic Engineering School, Heilongjiang University, Harbin 150080, China
{caobei, qindanyang}@hlju.edu.cn, wpf76842172@163.com

Abstract. With the advancement of the aerospace industry, the reliability of integrated circuits that can overcome the impact of the radiation environment continues to increase. In this paper, two kinds of radiative effects of integrated circuits are focused on the total ionizing dose effect and the single-event latch-up effect. In response to these two effects, the anti-radiation reinforcement technology of the logic gate standard cell was adopted at the corresponding design level. The design of the schematic and the circuit simulation are performed, including transient simulation and static simulation in this paper. In addition, the standard cell's logic information is verified and integrated into the standard cell library. A guard ring is designed on the layout of the standard cell so that it has radiation resistance. The layout of the standard cell is abstracted and the file containing the standard cell physical information is exported. The verification results and physical information of the standard cell demonstrate that the proposed technique can achieve the effect of anti-radiation.

Keywords: Standard cell · Radiation effect · Anti-radiation · Logic information · Physical information

1 Introduction

After entering the twenty-first century, mankind's enthusiasm for exploration into the world of outer space has increased. In order to meet human desires for knowledge, all countries with aerospace strength have launched a series of aerospace projects accordingly [1]. The number of satellites is also growing at a rapid rate. However, in the face of harsh radiation environments in outer space, ensuring the radiation resistance of satellites has also become a key technology in current aerospace [2]. So the anti-radiation design of standard cells is of great significance that the integrated circuit chip's working status is more stable.

From 1960s to 1980s, about 2,000 errors and malfunctions occurred in the spacecraft launched by the United States [3]. Whether it is a bipolar or MOS type circuit device, a parasitic transistor structure is present in the circuit structure. After the parasitic transistor structure is bombarded by high-energy particles, the integrated circuit will operate in a latched state. And the metal wiring area of the MOS type integrated circuit contains parasitic transistors, which generate leakage current after radiation. Meanwhile, it is important that the integrated circuit still has high reliability after being exposed to radiation. It plays an extremely crucial role in the normal operation of the system [4]. Therefore, it is of great significance to adopt the anti-

radiation design for the standard cell. After years of exploration and research, the research of anti-radiation integrated circuits has achieved results. In particular, research on the reinforcement of the radiation damage mechanism, the craftwork, layout, circuit structure and chip system design has developed [5]. At present, with the continuous enhancement of integrated circuit design and the increasing research on the space industry, the technology of chip design applied to aerospace is also increasing [6]. Satellite applications have a long time since becoming an indispensable method in economic and social development.

In this paper, it will focus on the total ionizing dose effect (TID) and the single-event latch-up effect (SEL). The impact of the space radiation environment on the circuit components is also described. The circuit of the anti-radiation standard cell is designed based on the Smic 0.18 μm process library. Then, the circuit simulation is carried out. Finally, the logical information of standard cells are integrated. The guard ring are added to the layout of the standard cell, and the physical information of the standard cell is extracted.

2 Design Theory of Anti-radiation Standard Cell

The radiative effect of integrated circuits mainly depends on the radiation environment in which the integrated circuit is located. The spacecraft's integrated circuits are mainly exposed to the radiation of the space environment. So a brief introduction of the radiation effects is described, as well as the analysis of the TID and SEL subjected to the integrated circuit and the corresponding reinforcement methods.

2.1 Radiation Effects

When the spacecraft is in a radiation environment for a long time and is bombarded by high-energy particles, the electrical parameters of the integrated circuit in the spacecraft will be changed accordingly so that it will result in device failure. This is the total ionizing dose effect (TID). The positive and negative fixed space electric charge generated in the SiO_2 layer of the MOS device due to the influence of radiation is the source of the TID generated by the MOS device. When the MOS device is exposed to ionizing radiation, both the positive charge induced in the oxide layer and the trapped charge trapped in the interface state will change, and it will result in a shift in the threshold voltage of the MOS transistor. At the same time, the transconductance of the MOS transistor will also be reduced.

In addition, when the MOS device is subject to single-particle incidence, the parasitic transistor in the MOS transistor is in a forward biased state, so a latch is formed and result in a short circuit between the power supply and the ground so that the MOS device will operate at low voltage and high current. It is a single-event latch-up effect (SEL). At this time, even if the bias voltage of the gate is reduced, the leakage current of the CMOS circuit will be very large.

2.2 Anti-radiation Effect Reinforcement Method

With the development of the deep submicron process of integrated circuits, the influence of the threshold voltage drift can be basically ignored. Therefore, the anti-total ionizing dose effect mainly starts from eliminating the increased off-state leakage current. Mainly the following two methods can be used:

1. Prevent leakage from the drain to the source of the edge of the leakage path. In the layout design of the standard cell, when the layout of the MOS transistor is drawn, the ring-gate structure can effectively prevent the drain-to-source path.
2. Prevent leakage paths below the SiO₂ layer. By adding a guard ring to the layout of the standard cell, parasitic transistors are difficult to turn on, thereby suppressing leakage under the field oxide layer.

Two methods to suppress single-event latch-up effects

1. Make contact holes as many as possible on the layout. The more contact holes, the smaller the parallel resistance of the well and the substrate.
2. A P+ guard ring is added around the N-channel MOS transistor and an N+ guard ring is added around the P-channel MOS transistor.

3 Circuit Design and Simulation of Anti-radiation Standard Cell

When a standard cell is designed, the schematic of the standard cell should be first designed. The standard cell circuit design has a simple logic gate design and a complex standard cell circuit design. The width and length of the MOS transistors for each standard cell in the standard cell library need to be determined, and each standard cell has a different drive circuit structure.

3.1 Standard Cell Circuit Design

In the Smic 0.18 μm process, the schematic circuit diagram of the standard cell is drawn. The width of the PMOS transistor of the INV 1 time drive force is 0.9 μm , the width of the NMOS transistor is 0.6 μm . The width of a PMOS transistor is 1.8 μm , the NMOS transistor is 1.2 μm INV is an inverter with two times drive capability. A 2 times drive INV has a larger output, but its input capacitance is larger than the input capacitance of the 1-time drive inverter. In addition, the width-to-length ratio of the MOS transistor in the standard cell is increased, and the on-resistance of the MOS transistor is also reduced, so the speed of the logic gate is faster.

Among them, the port and name of the symbol map must match the internal port name of the schematic. The schematics of INVX1 and INVX2 are shown in Figs. 1 and 2.

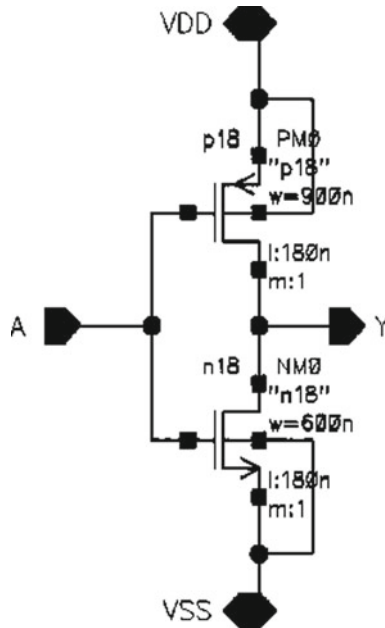


Fig. 1. INVM1 circuit schematic

3.2 Transient Simulation of Standard Cells

In many standard cells, the inverter is the simplest CMOS combinatorial logic cell, but it is also the core cell of digital integrated circuit design. Among the 34 anti-radiation standard cells designed in this paper, the following will mainly use INV as an example to introduce the simulation of anti-radiation standard cells and extract the logical information of standard cells.

At room temperature, with 100% VDD, transient simulations of INV were performed and the waveforms shown in Fig. 3 were obtained. It can be seen that the A-side signal is opposite to the Y-side signal, and it indicates that INV can implement the inversion function.

3.3 Timing Simulation of Standard Cells

In this paper, the measured standard cell's propagation delay is referenced to 50% VDD, and the transmission time information is referenced to 10% VDD and 90% VDD. In the Smic 0.18 μm standard cell library, the input timing of the input signal and the external load capacitance of the combinatorial logic cell are stored in a two-dimensional lookup table. The definition of the two-dimensional lookup table for the ascending propagation delay in the .lib file is shown in Fig. 4.

Index_1 and index_2 represent the clock skew of the input signal and the external load capacitance, respectively, by changing these two variables to determine the propagation delay and transmission time of the standard cell. As shown in Fig. 5, the

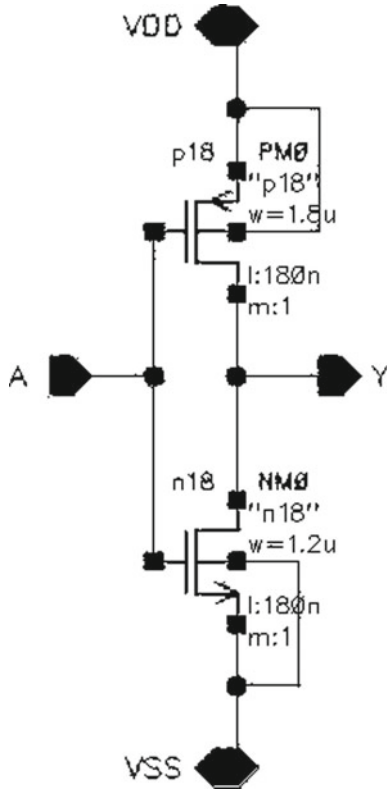


Fig. 2. INVX2 circuit schematic

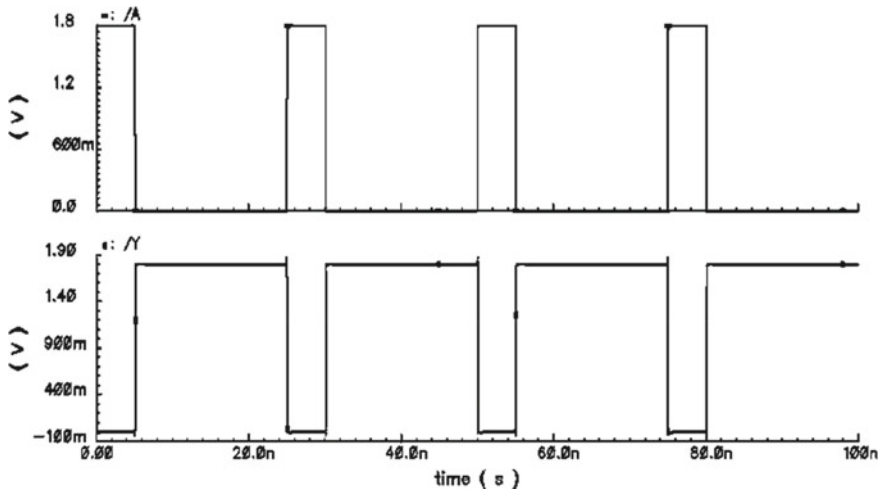


Fig. 3. INV transient simulation waveform

```

timing() {
  related_pin : "A";
  timing_sense : negative_unate;
  cell_rise(delay_template_7x7) {
    index_1 ("0.03, 0.1, 0.4, 0.9, 1.5, 2.2, 3");
    index_2 ("0.00035, 0.021, 0.0385, 0.084, 0.147, 0.231, 0.3115");
  }
}
    
```

Fig. 4. Lookup table definition for timing information in the .lib file

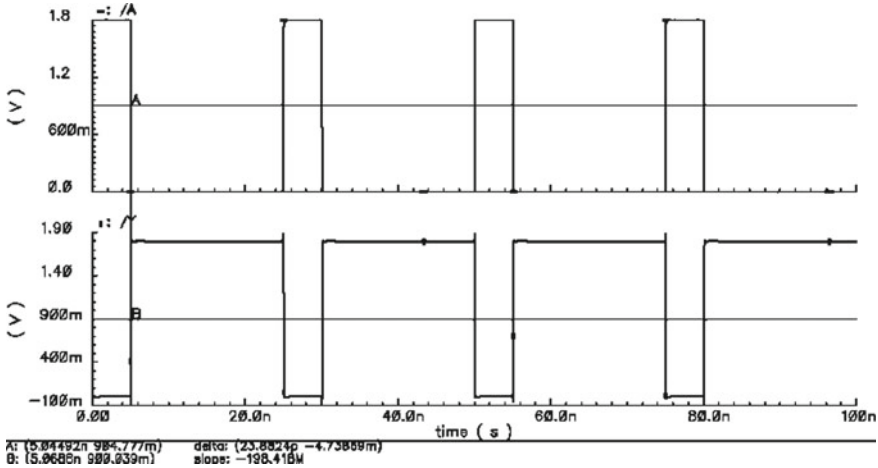


Fig. 5. Rising propagation delay measurement of INVX1

rise propagation delay of INVX1 can be measured as 23.8824 ps, because each propagation delay and transmission time has 49 kinds of situations. Meanwhile, the script method can be written to carry on the time sequence simulation to the standard cell, and it can save a lot of time. After obtaining the timing information of the standard cells, the timing information is integrated into the format of a two-dimensional lookup table.

3.4 Standard Cell Power Simulation

When the output of the standard cell is logically inverted, the energy consumed during the simultaneous conduction of the pull-up network and the pull-down network is referred to as internal short circuit power consumption. The power consumption of the internal short circuit within a time interval t is shown in Eq. (3.1). The external load capacitor power dissipation is shown in Eq. (3.2).

$$E_{short} = \int_0^t I_{DD} * VDD dt \tag{3.1}$$

$$E_{load} = \frac{1}{2} * VDD^2 * C_L \quad (3.2)$$

When the internal rising edge power consumption is measured, the external load capacitor is charged. At this time, the measured internal short circuit power consumption includes the external load capacitor charging power consumption. So the external load capacitor power consumption should be subtracted. When the internal falling edge power consumption is measured, the external load capacitor is discharged and the external load capacitance does not affect the internal short circuit power consumption. In this case, the external load capacitor power consumption does not need to be subtracted.

The clock skew is set to 0.03 ns in the first group, and the output load capacitance is set to 0.00035 pF in the first group. When the input signal A is inverted from high to low, the output signal Y goes from low to high. When the level is reversed the transient waveform of the supply current is shown in Fig. 6. Calculated from Eq. (3.1), the internal drop power is 0.03337 pW. A total of 49 internal power consumption measurements need to be measured and integrated into a two-dimensional lookup table format.

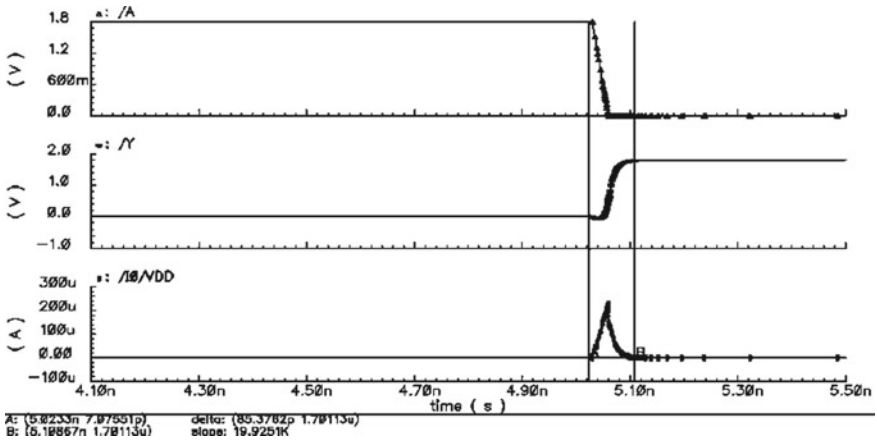


Fig. 6. INVX1 internal falling power measurement

When the MOS transistor is turned off, the insulating property of the insulating material used in the process of fabricating the MOS transistor is insufficient, so that it will result in that the MOS transistor cannot be completely turned off, and excess power consumption. That is leakage power consumption. Leakage power P_{static} is shown in Eq. (3.3).

$$P_{static} = VDD * I_{leakage} \quad (3.3)$$

As shown in Fig. 7, when the input voltage is 0 V, the supply current is 23.6577 pA. By Eq. (3.3), the leakage power consumption of INVX1 is 42.58386 pW.

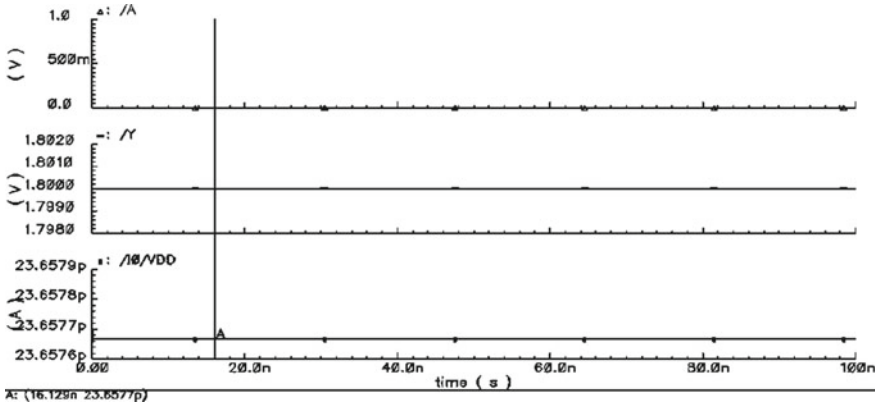


Fig. 7. Measurement of INVX1 leakage power consumption simulation waveform

3.5 Input Capacitance Measurement of Standard Cell

The extraction of the capacitance of the standard cell mainly refers to the extraction of the input pin capacitance. However, the input capacitance value of the combinational logic cell cannot be measured directly, and the input capacitance can only be measured through an indirect method. The specific circuit is shown in Fig. 8. At this time, the INVX8's output load capacitance is the input pin capacitance of INVX1.

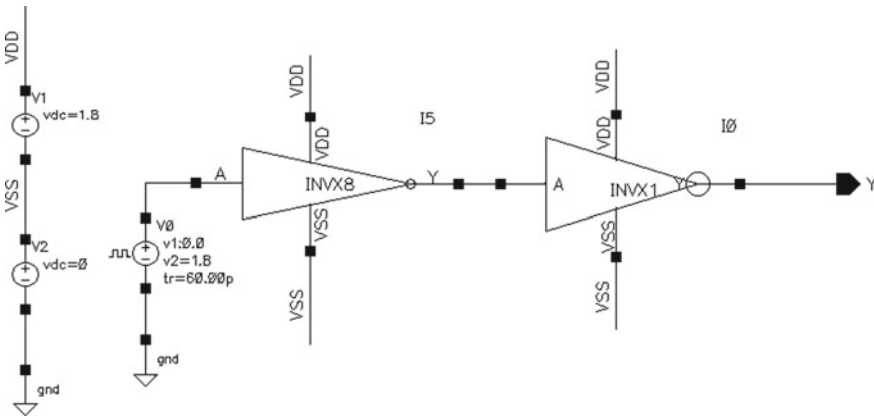


Fig. 8. Circuit for measuring input capacitance

X is the transmission time of INVX8 input port. Y is the output load capacitance of INVX8. Z is the propagation delay of INVX8. And the relationship between the three variables of X , Y , and Z can be obtained by the interpolation algorithm. See Eq. (3.4).

$$Z = A + B * X + C * Y + D * X * Y \quad (3.4)$$

According to the data in the propagation delay lookup table of INVX8, let $X = 0.06$ ns, $Y = 0.168$ pF, $Z = 0.1198$ ns. Taking four sampling points near X , Y , and Z and substituting these four sampling points into Eq. (3.4). Then, four linear equations are obtained. At this point, the values of A , B , C , and D can be calculated. The relation between the propagation delay of INVX8 and the output load capacitance is shown in Eq. (3.5).

$$Z = 0.014431 + 0.358059 * Y \quad (3.5)$$

As can be seen in Fig. 9, the falling propagation delay of INVX8 is 0.0158625 ns. Substituting into Eq. (3.5), the INVX1 input pin capacitance is 0.003997 pF.

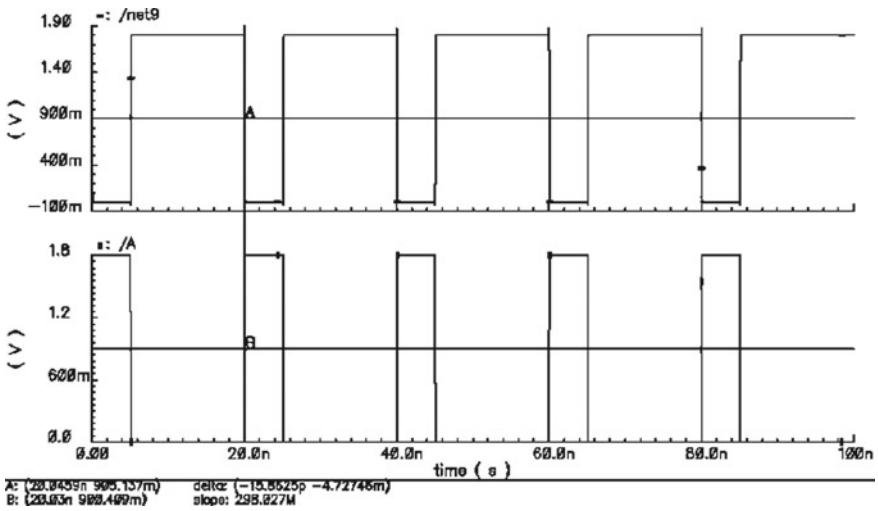


Fig. 9. Measurement input pin capacitance simulation

3.6 Logical Information Verification of Standard Cells

After the synthesis process is completed, the synthesis results are written in a structure-level Verilog netlist file. Part of the netlist structure is shown in Fig. 10. As can be seen from the figure, the synthesis tool has successfully synthesized the 34 standard cells designed in this paper. It indicates that the logic information of the anti-radiation standard cell designed in this paper is correct.


```

NOR2XL U24453 ( .A(n24406), .B(n24405), .Y(N57) );
NAND2XL U24454 ( .A(clk_in_num[10]), .B(n24406), .Y(n24409) );
NAND2XL U24455 ( .A(n24409), .B(n24404), .Y(n24408) );
NOR2XL U24456 ( .A(clk_in_num[10]), .B(n24406), .Y(n24407) );
NOR2XL U24457 ( .A(n24408), .B(n24407), .Y(N58) );
INVX1 U24458 ( .A(n24409), .Y(n24410) );
NAND2XL U24459 ( .A(clk_in_num[11]), .B(n24410), .Y(n24413) );
NAND2XL U24460 ( .A(n24413), .B(n24404), .Y(n24412) );
NOR2XL U24461 ( .A(clk_in_num[11]), .B(n24410), .Y(n24411) );
NOR2XL U24462 ( .A(n24412), .B(n24411), .Y(N59) );
INVX1 U24463 ( .A(n24413), .Y(n24414) );
NOR2XL U24464 ( .A(clk_in_num[12]), .B(n24414), .Y(n24415) );
NOR2XL U24465 ( .A(n24416), .B(n24415), .Y(N60) );
endmodule

```

Fig. 10. Integrated part of the netlist

4 Layout Design and Abstraction of Anti-radiation Standard Cell

Design rules play an interface between circuit designers and process engineers. In this paper, the design of anti-radiation standard cells is mainly focused on layout design, and 34 key combinations of the dual-loop protection structure are used to design the combinational logic cell layout.

4.1 Anti-radiation Standard Cell's Plate Drawing Method and Verification

Since the Smic 0.18 μm process is used to design standard cells, the design rules of the Smic 0.18 μm process must be understood before the layout is drawn, including the minimum linewidth, minimum spacing, and minimum extension. According to the layout design rule of the Smic 0.18 μm process, the standard cell layout height is set to 5.04 μm and the width is a multiple of 0.66 μm . Different from common commercial standard cells, the layout of the standard cell designed in this paper incorporates a P+ guard ring and an N+ guard ring. It will make it has a larger area than the common standard cell, as shown in Figs. 11 and 12.

After verification by the DRC, the density of the five-layer metal except the metal 1 in the standard cell designed in this paper is correct, and the density problem can be ignored here. The LVS inspections of the standard cells of radiation resistance designed in this paper are all matched with each other.

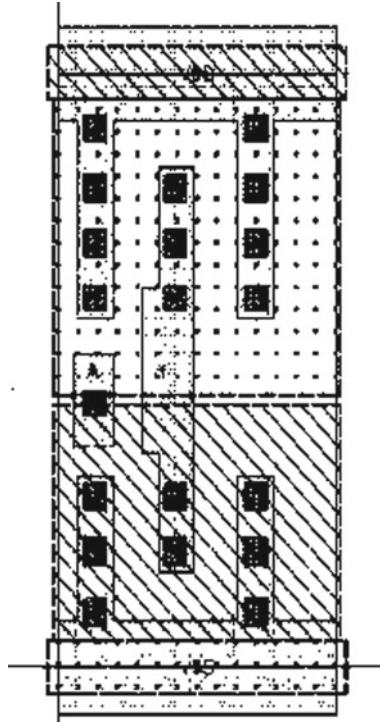


Fig. 11. INVX4 layout without a guard ring

Parasitic parameters of standard cells were extracted by PEX, and a netlist file containing parasitic capacitance and parasitic resistance information was obtained. The netlist is replaced with the parasitic parameters in the netlist part of the input .scs file generated by the previous simulation and post-simulation is performed on the NAND2X2. The simulation waveform is shown in Fig. 13. The waveform is similar to the previous simulation, so it indicates that the NAND2X2 is functionally correct with parasitic parameters.

4.2 Anti-radiation Standard Cell Layout Abstraction

In order for the place-and-route tool to perform the layout work more efficiently, some of the physical information of the standard cell should be extracted and a .lef file needs to be generated. For example, the bounding box of each standard cell, the power/ground connection point of the standard cell, and the input/output signal connection ports, wiring barriers, and other information. A view of a standard cell containing only this information is called an abstract view of the standard cell. An abstract view of NAND2X1 is shown in Fig. 14.

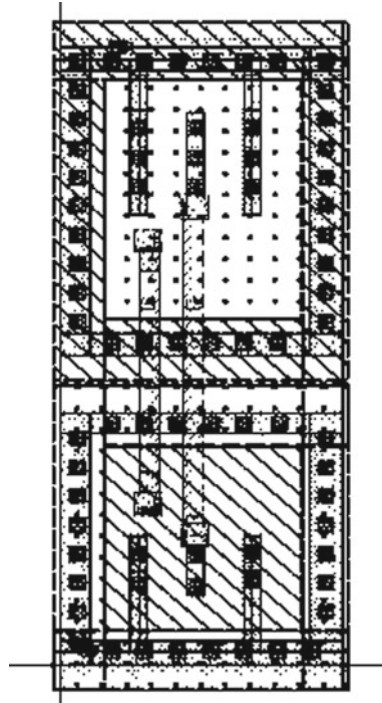


Fig. 12. INVX4 layout with guard ring

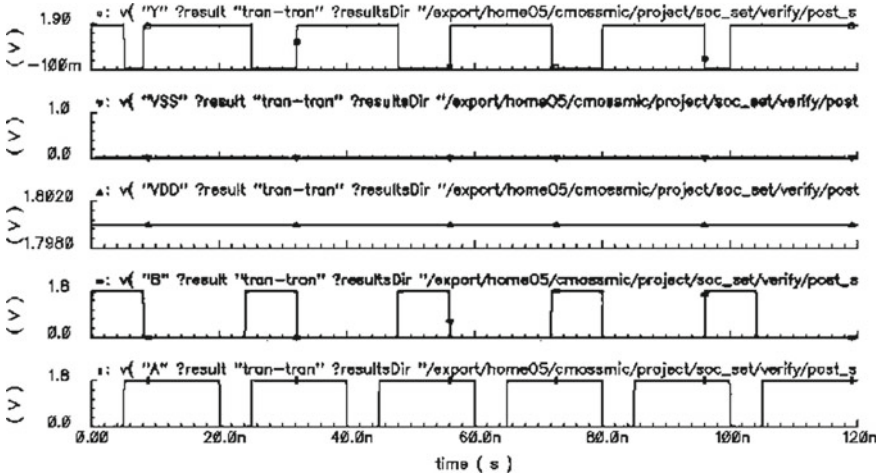


Fig. 13. Post-simulation waveform of NAND2X2

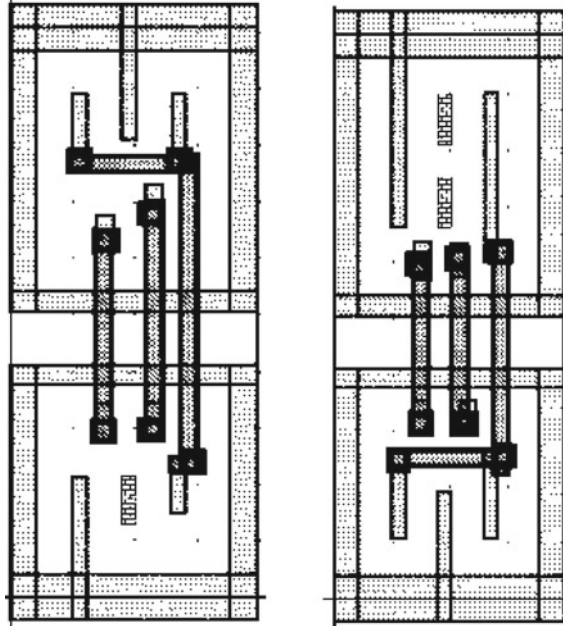


Fig. 14. Abstract view of NAND2X1 and NOR2X2

In this step of the layout abstraction, the standard cell Site is mainly set, and the other setting fields are kept at the default values. When the abstraction of the standard cell is completed, the output of the .lef file is completed. The geometric details of the standard cell layout are abstracted in the abstraction process. The place-and-route tool is provided with the layout and information related to the process for the wiring. And the writing efficiency will be improved.

The .lef file containing physical information from the layout is a library exchange format file. The header file section gives the length and width of the standard cell, and the rest shows the standard cell port position, layer definition, and other information. The exported .lef file for INVX1 is shown in Fig. 15.

After the placement and routing are completed, some standard cells views are shown in Fig. 16. The standard cell designed in this paper is a combinational logic cell. There is no error indication in the layout and routing process, and the final generated standard cell layout is verified by DRC and LVS. The display shows no error. The 34 standard cells designed in this paper can be recognized by the digital integrated circuit tools and have the correct function.

```

MACRO INX1
  CLASS CORE ;
  FOREIGN INX1 0 0 ;
  ORIGIN 0.0000 0.0000 ;
  SIZE 2.9700 BY 10.0800 ;
  SYMMETRY X Y ;
  SITE INX1site ;
  PIN A
    DIRECTION INPUT ;
    PORT
      LAYER VIA12 ;
      RECT 1.1050 2.0550 1.3650 2.3150 ;
      RECT 1.1050 7.5250 1.3650 7.7850 ;
      LAYER METAL1 ;
      RECT 1.0450 1.9950 1.4250 2.3750 ;
      RECT 1.0950 1.9950 1.3750 2.7050 ;
      RECT 1.0450 7.4650 1.4250 7.8450 ;
      RECT 1.0950 7.1650 1.3750 7.8450 ;
      LAYER METAL2 ;
      RECT 1.0450 7.4650 1.4250 7.8450 ;
      RECT 1.0450 1.9950 1.4250 2.3750 ;
      RECT 1.0950 1.9950 1.3750 7.8450 ;
    END
  END A

```

Fig. 15. INX1 .lef file

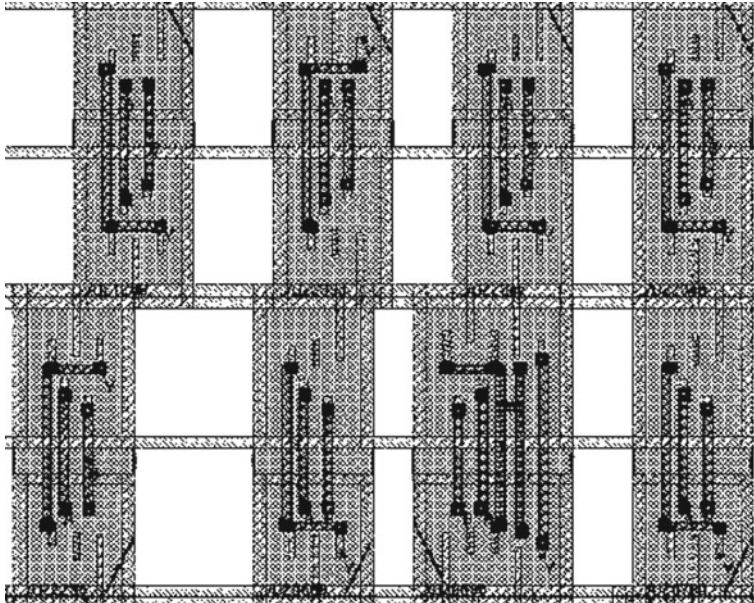


Fig. 16. A view of some combinational logic cells after placement and routing

5 Conclusions

Based on the theoretical analysis of total ionizing dose effect and single-event latch-up effect, 34 combinatorial logic cells with anti-radiation properties were designed in this paper. The standard cell was subjected to circuit simulation and the standard cell's propagation delay and transmission time were measured. According to the clock skew and external load information in the lookup table, the standard cell's internal power consumption and leakage power are measured. An INVX8 is added to the input of the standard cell circuit and the input capacitance of the standard cell is calculated by using the interpolation method. In the layout design, P+ guard ring and N+ guard ring were added to the layout of the standard cell, so that the standard cell has radiation resistance performance. After the layout verification, the layout of the standard cell is abstracted and a .lef file containing the standard cell physical information is obtained. The verification of logical information and physical information was carried out. From this, the anti-radiation standard cell designed in this paper has availability.

Acknowledgements. I am grateful to have the opportunity to enter the HIT microelectronics center. They gave me great encouragement and support during the study. At the same time, I also became interested in digital integrated circuits. Meanwhile, I would like to express my gratitude to my classmates for teaching me the experience and methods in my graduation design work. Their active, optimistic, progressive, hardworking research spirit inspired me at all times.

References

1. Walker, J.A., Sinnott, R., Stewart, G., Hilder, J.A., Tyrrell, A.M.: Optimizing electronic standard cell libraries for variability tolerance through the nano-CMOS grid. *Philos. Trans. R. Soc. A* **368**(1925) (2010)
2. Jianhua, J., Man, L., Lei, W., Yumei, Z.: An effective timing characterization method for an accuracy-proved VLSI standard cell library. *J. Semicond.* **35**(2) (2014)
3. Qi, C., Xiao, L., Guo, J., Wang, T.: Low cost and highly reliable radiation hardened latch design in 65 nm CMOS technology. *Microelectron. Reliab.* **55**(6) (2015)
4. Kung, D.S., Puri, R.: Optimal P/N Width Ratio Selection for Standard Cell Libraries, pp. 276–283. *IEEE* (1999)
5. Brusamarello, L., Wirth, G.I., Roussel, P., Miranda, M.: Fast and accurate statistical characterization of standard cell libraries. *Microelectron. Reliab.* (12) (2011)
6. Tao, H.: Design and Verification of High Reliability Standard Cell Library. National University of Defense Technology, Hefei (2011)



Design and Implementation of Intelligent Crib-Based on Android

Junjiao Zhang^(✉), Yongjie Yang^(✉), and Zhongxing Huo

School of Electronics and Information, Nantong University, Nantong 226019,
China

1047891636@qq.com, yang.yj@ntu.edu.cn

Abstract. The traditional crib has the simple function of carrying the baby to sleep and play, and the baby falling asleep also needs the parents to shake the crib manually. According to the above situation, the intelligent crib-based on Android is designed and implemented. The system takes STM32 MCU as the core, combines sensor technology, wireless communication technology, and video monitoring technology, uses MQTT server, and develops an app with Android studio. Parents can not only view the baby's condition by remote video, but also check the baby's fever, bed wetting, and control music play, shake the crib through the mobile phone interface. The design is convenient and effective to lighten the burden on parents to take care of infants and has high application value.

Keywords: Android · Intelligent · STM32 · MQTT

1 Introduction

With the improvement of living standards, there is a growing demand for the function of the crib. Most of the existing cribs are for babies to sleep and play, and parents still need to spend a lot of time and energy in caring for their babies [1]. In order to solve these problems effectively, an intelligent crib is developed in this paper. The system uses multiple sensors to collect the baby's body temperature and wetting the bed in real time. This monitoring information can be displayed on the app interface, so as to realize the effective care of infants.

2 Overall Design of System

The system is divided into a management layer, transmission layer, and acquisition layer. The overall architecture of the system is shown in Fig. 1.

The acquisition layer is mainly data acquisition terminal, with STM32F103C8T6 as the main control chip, carrying temperature and humidity sensors, sound acquisition

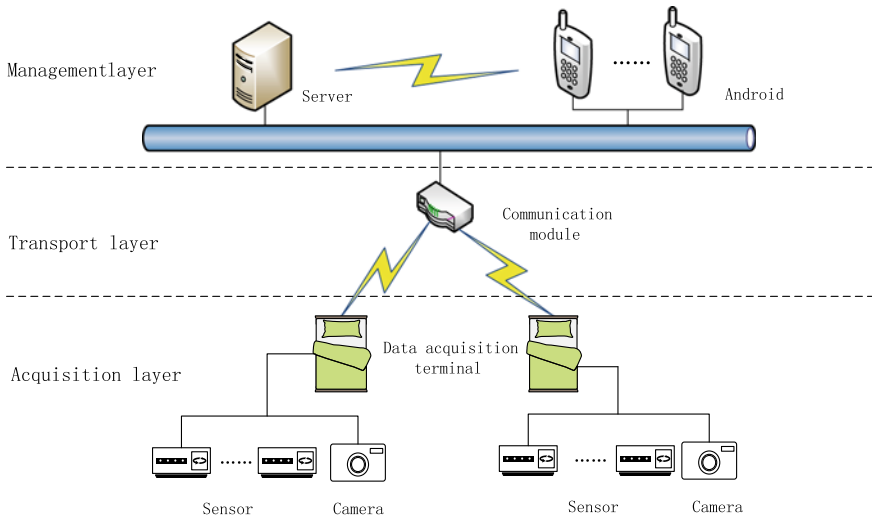


Fig. 1. System structure diagram

sensors and MP3 player. Among them, the temperature and humidity sensor is responsible for collecting the temperature and humidity information in the crib, the audio acquisition sensor detects the baby crying and the MP3 player plays the soothing baby music when the baby cries [2]. In addition, the camera provides real-time baby pictures for caregivers.

As an important part of the whole system, the communication module of the transport layer realizes two-way communication between the acquisition layer and the management layer. The communication module uses a WiFi module for wireless transmission, loading mature 802.11n protocol, to ensure the stability and reliability of data transmission [3].

The management layer consists of the MQTT server and the Android side. The Android terminal connects to the server to obtain the status of the device in the acquisition layer and updates the data displayed by the app interface in real time [4]. The app can accept the user's operation and send the data instructions to the acquisition device through the MQTT server.

3 Hardware Design

Hardware design is the basis of the system function, which mainly includes: a power module, a sensor module, drive motor and communication module. The structure is shown in Fig. 2. The hardware system has the characteristics of reliable data transmission, stable performance and a high degree of automation.

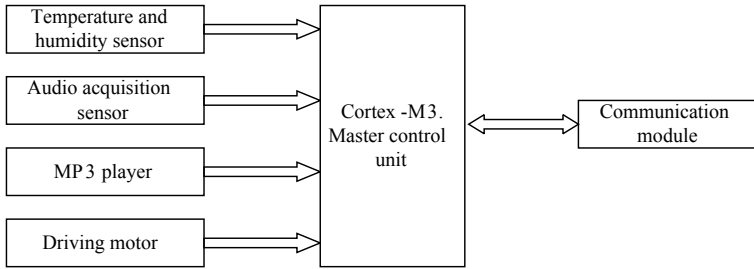


Fig. 2. Hardware design structure diagram

3.1 Main Control Chip

Considering the cost, performance and power consumption, the main chip of the data acquisition system is STM32F103C8T6 microprocessor [5]. The CPU based on arm Cortex-M3 kernel can work at a maximum frequency of 72 MHz, with a single cycle multiplication instruction and hardware division, as well as a programmable priority interrupt system. It has 64 kB Flash memory and 20 kB SRAM memory and integrates a wealth of on-chip peripherals. The master chip has the characteristics of low power consumption, low cost, and fully meets the design requirements of the system.

3.2 Power Supply Module

In order to maintain the cooling function of the circuit board, the power module adopts the external 12 V switching power input. In view of the needs of 5 and 3.3 V power supply in each module of the hardware design, the power supply step-down circuit shown in Fig. 3 is designed.

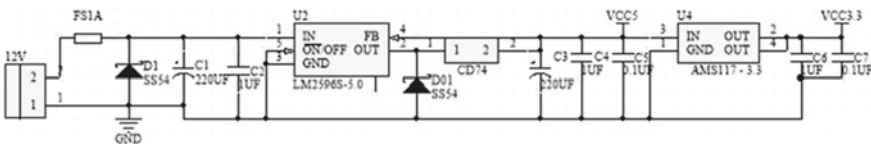


Fig. 3. Circuit diagram of power module

The power input end is equipped with a 1 A self-recovery fuse and diode for power reverse connection protection, preventing the wrong operation from burning the circuit board. The whole circuit adopts the multi-stage step-down method to avoid the problem of heating the power chip due to excessive voltage drop. A power supply is added to filter before and after each step of the step-down. The 5 V voltage generated by LM2596-5.0 supplies power to the sensor and drive motor, while the 3.3 V voltage generated by AMS1117-3.3 supplies power to the main control chip and the communication module.

3.3 Sensor Module

Sensor technology has become an important means to obtain the information about the physical world. SHT10 temperature and humidity sensor, audio acquisition sensor and MP3 player are selected to detect the temperature, and humidity in the crib and crying information, respectively [6].

Here we will mainly introduce the audio acquisition circuit, which can collect baby’s crying through the electret microphone. The baby’s crying frequency is generally within the range of 400–1200 Hz, so it is necessary to design a band-pass filter with a lower frequency of 400 Hz and an upper limit of 1200 Hz to filter the interference signal. In addition, the module supports the reading and writing of memory cards, and can download and change the music on its own.

3.4 Communication Module

The communication module of the system adopts WiFi wireless transmission [7]. This module is USB-C215 WiFi module developed by Jinan company. It has the advantages of ultra-small size, an ultra-low power consumption of serial port, and supports TTL serial port, WiFi wireless network and two-way transparent transmission of Ethernet port.

The application circuit for the USB-C215 WiFi module is shown in Fig. 4. After WIFI initialization, the green light is on; data acquisition terminal successfully connected to the server, the yellow light is on.

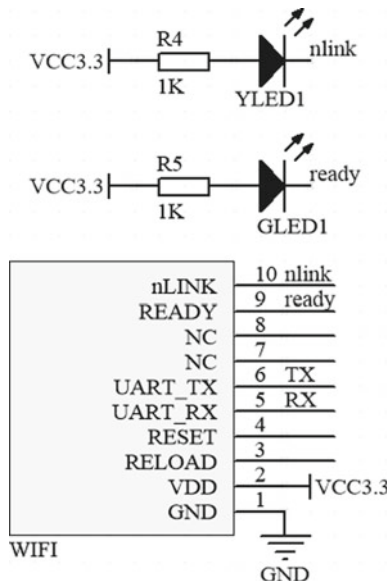


Fig. 4. Application circuit diagram for USB-C215 WiFi module

4 Software Design

The software design of the system is mainly divided into two parts: the software design of the acquisition terminal and the software design of Android. It uses Youren transmission cloud to communicate with these two parts. Youren transmission cloud is a server developed based on the MQTT protocol [8]. MQTT is a publish/subscribe protocol based on TCP/IP for the mobile terminal device. The publish/subscribe message pattern means that the publisher and the subscriber do not interact with the data directly through the MQTT protocol, but interact with the message broker indirectly through the MQTT protocol [9].

4.1 The Software Design of the Acquisition Terminal

The software design flow of the acquisition terminal is shown in Fig. 5.

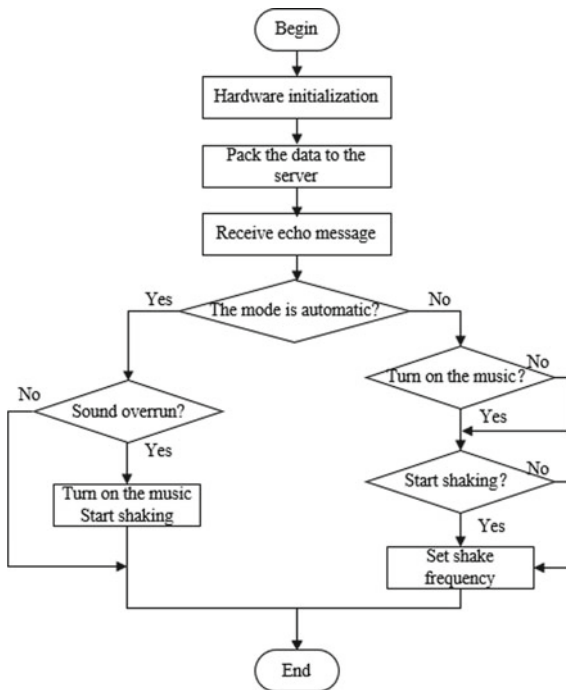


Fig. 5. Flow chart of software design for acquisition terminal

After the circuit is powered on, the program first needs to initialize the modules on the hardware circuit board, including the clock frequency and interrupt vector initialization of STM32 MCU, and so on. JSON (JavaScript Object Notation), a communication protocol in JSON format for data transmission, is a lightweight data representation method, which records data in the way of key: value.

Before sending the sensor information to the server, the acquisition terminal assembles the collected data into the following format:

{“temperature”: 30, “humidity”: 60, “noise”: 30, “auto”: 1, “music”: 1, “turn”: 1, “frequency”: 10}.

Key and the corresponding value are separated by a colon “:”. Key values in turn represent temperature (unit: °C), humidity (unit: % RH), baby sound (unit: dB), crib mode (0: auto 1: manual), music switch state (0: open 1: close), swing switch state (0: open: close), crib mode (0: open: close), rocking frequency of crib (range 1–100).

4.2 The Software Design of the Android

The main program of software design at the Android end mainly includes server connection, data transfer, XML page layout, video monitoring, etc. Flow Chart of Software Design at Android is shown in Fig. 6 [10].

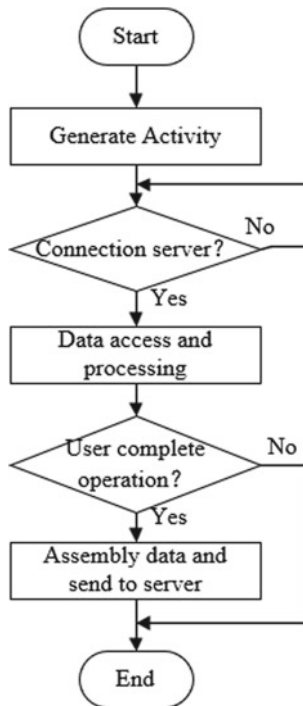


Fig. 6. Flow chart of software design at Android

When the Android terminal connects with the transparent cloud server, the data callback function is initialized. App receives data, the callback function is automatically invoked. It is only when the callback function's return code is 0x02 that the Android side is connected to the server successfully [11].

OnReceiveEvent function and publishForuName function are used in data access and upload respectively. The former converts the received data array into a string, then parses it through JSONObject, and then app updates the display content according to the parsed data. If the user clicks the switch button, the program sends the assembled JSON string to the collection device through the publishForuName function [12].

XML-layouts combine multiple layout managers and use text boxes, radio buttons, scrollbars, and other controls to generate two Activity interfaces: the login interface, as shown in Fig. 7, and the monitoring interface, as shown in Fig. 8.

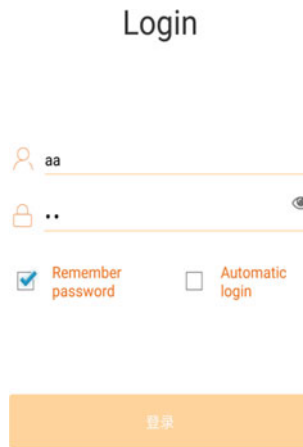


Fig. 7. Login interface

Enter the correct user name and password in the login interface before entering the monitoring interface, which shows the temperature and humidity of the baby bed, the size of baby crying, and the state of the crib in real time. Crib mode is divided into automatic mode and manual mode.

Video images are captured using fluorite cloud C6C wireless webcam. It uses EZOPEN protocol, which is a URL protocol specially designed for developers, and currently supports video preview and playback functions.

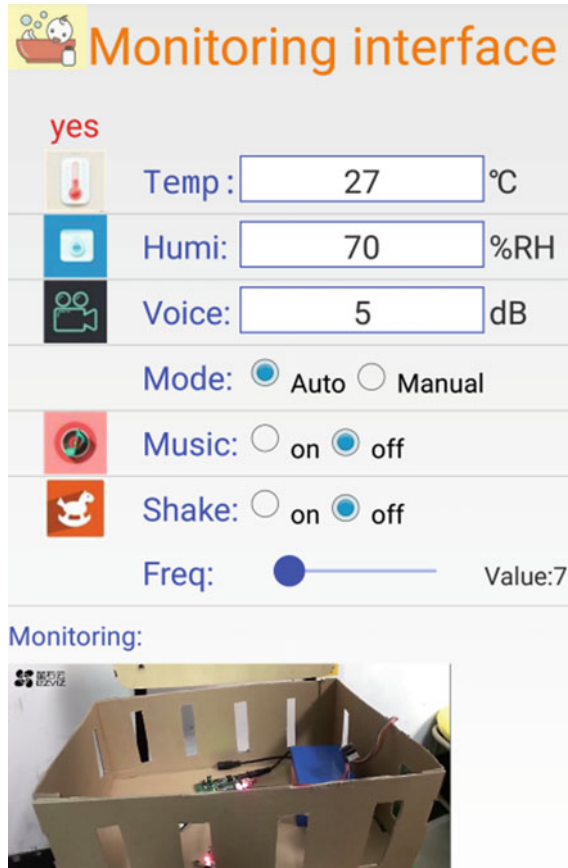


Fig. 8. Monitoring interface

5 Overall Design of System

In order to test the effect, the crib model was designed and simulated. Prior to the test, a Corfu HTD8818D temperature gun, a pen probe hygrometer, and a Sima AR824 decibel meter were prepared. The accuracy of the temperature gun is ± 0.3 °C, that of the hygrometer is $\pm 1\%$ RH, and that of the decibel meter is ± 1.5 dB. By selecting three infants to be put into crib, the temperature, humidity and sound of app interface and temperature, and humidity measuring instrument and decibel meter were compared, as shown in Table 1. It can be seen from the data in the table that the error of temperature is within 0.3° , the error of humidity is within 4% RH, and the error of sound value is between 0 and 4 dB.

Table 1. Comparison table of temperature, humidity and sound

Num	APP: temperature (°C)	APP: humidity	APP: sound (dB)	Device: temperature	Device: humidity	Device: sound (dB)
1	37.2	77% RH	40	37.3	74% RH	43.6
2	37.4	75% RH	26	37.2	70% RH	25.3
3	37.1	72% RH	13	37.4	73% RH	15.1

6 Conclusions

According to the present situation and demand of crib development, an intelligent crib management system based on Android is developed in this paper which is based on the classic three-layer Internet of things architecture model, and combined with systematic hardware design of data acquisition and software modular design. The test results verify the rationality and feasibility of the design. The whole system focuses on the humanization of taking care of infants, aiming at alleviating the pressure of parents, and realizing the highly intelligent crib. At the same time, the utility model makes up for the shortage of the existing crib function and has certain practicability.

Acknowledgements. This paper is supported by Nantong University-Nantong Intelligent Information Technology Joint Research Center Open Topic (KFKT2017B05) and the first phase project of Jiangsu University Brand Specialty Construction Project (PPZY2015B135).

References

1. Li, G., Zhong, L.J., Zhang, R.: Design and development of portable folding baby bed. *Appl. Mech. Mater.* **443**, 65–68 (2014)
2. Chen, G.: Design of automatic baby nursing trolley based on intelligent sensing technology. *Appl. Mech. Mater.* **686**(8), 258–262 (2014)
3. Sowah, R.A., Ofoli, A.R., Krakani, S.N., et al.: Hardware design and web-based communication modules of a real-time multisensor fire detection and notification system using fuzzy logic. *IEEE Trans. Ind. Appl.* **53**(1), 559–566 (2017)
4. Li-Jing, B.I., Jiang, B.C., Wang, X.L., et al.: Real-time vehicle monitoring system based on android. *Comput. Knowl. Technol.* (2017)
5. Dai, C.G., Ren, D.J., Jiang, T., et al.: Design on remote signals acquisition instrument based on STM32 microprocessor. *Mech. Electr. Eng. Technol.* (2013)
6. Guo-Cheng, L.I., Chen, J.M., Wei-Lin, L.I., et al.: The design of baby real-time monitoring system that based on the Internet of things. *Electron. Des. Eng.* (2015)
7. Xiong, J., Wang, T.: Design of multi-parameter portable water quality detector based on wireless network. *Wirel. Pers. Commun.* **8**, 1–14 (2018)
8. Guan, Q.Y., Hong-Bin, L.I., Bo, Y.U.: Research and application of the MQTT protocol on the android platform. *Comput. Syst. Appl.* (2014)
9. Chooruang, K., Mangkalakeeree, P.: Wireless heart rate monitoring system using MQTT ☆. *Procedia Comput. Sci.* **86**, 160–163 (2016)

10. Long, W., Xie, Y.: Design and development of health care based on android. *Comput. Knowl. Technol.* (2016)
11. Naik, N.: Choice of effective messaging protocols for IoT systems: MQTT, CoAP, AMQP and HTTP. In: *Systems Engineering Symposium*, pp. 1–7. IEEE (2017)
12. Nurseitov, N., Paulson, M., Reynolds, R., et al.: Comparison of JSON and XML data interchange formats: a case study. In: *ISCA, International Conference on Computer Applications in Industry and Engineering, Caine 2009, November 4–6, 2009, Hilton San Francisco Fisherman’s Wharf, San Francisco, California, USA. DBLP, 2009*, pp. 157–162



Big Data-Based Attack Scenario Reconstruction Architecture in Smart Grid

Liang Guo^{1(✉)}, Qianqian Jin^{1(✉)}, Ying Liu², Yuanyi Xia³, and Han Hu⁴

¹ NARI Group Corporation/State Grid Electric Power Research Institute, Nanjing, China

baoxingchuan@126.com, jinqianqian@sgepri.sgcc.com.cn

² State Grid Corporation of China, Beijing, China

³ State Grid Jiangsu Information and Telecommunication Company, Nanjing, China

⁴ School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing, China

Abstract. The intelligence of power grids has made the relationship between distribution networks and the Internet more and more compact. Therefore, in order to cope with the various threats in the situation of smart grid, it is necessary to study from multiple perspectives. Among them, attack scenario reconstruction is a more effective method of network security defense. However, the existing attack scenario reconstruction technology is not combined with the actual situation of the power grid. In this paper, we proposed a grid-based attack scenario reconstruction framework which is based on big data. The framework consists of KNN-based attack data classification and state machine-based attack scenario restoration. In addition, we also implemented prototypes and evaluated the effectiveness and availability of databases provided by IDS in China Grid Corporation. The results show that the framework proposed in this paper improves the efficiency and accuracy of analyzing attacker strategies.

Keywords: Big data · Attack scenario reconstruction · Smart grid · KNN

1 Introduction

The modern power system has evolved into a complex coupled network system consisting of an information communication system and a physical power system. That is a power information physical fusion system. Studies have shown that failure of devices in the information and communication networks or in the power

system itself or malicious attacks can endanger the safe operation of the entire coupled network system and cause serious consequences [1]. In recent years, there have been incidents of power system attacks. On December 23, 2015, the Ukrainian power grid was attacked by hackers, resulting in widespread blackouts in Ukraine. On January 25, 2016, the Israel Electric Power Bureau suffered a heavy cyberattack and a large number of infected hosts were temporarily shut down, severely affecting the normal power business [2].

In the face of various forms and constantly developing cyberattacks, traditional security event detection methods are increasingly unable to meet the needs of security management. Researching on how to analyze and deal with a large number of alarm information generated by the security management device in depth has become a new direction for the development of security event detection methods. Intrusion Attack Scenario Restoration Technology [3], which is a new method of security analysis, integrates and analyzes various types of information, and it reflects attackers' intent of attack as a whole IT provides threat prediction information and enhances the efficiency of network security management and formulates effective security planning and supervision [4]. The strategy provides the scientific basis. The intrusion attack scenario refers to the process in which the attacker uses the vulnerability of the target system to gradually increase his control authority and finally achieve the attack target [5]. The attack scenario consists of a series of single-step attacks. The intrusion attack scenario restoration technology is to correlate and analyze the original security events generated by security devices such as IDS to restore the attacker's process of attacking and infiltrating the entire network space, and then feedback such high-level scene information to the administrator. The correlation analysis of intrusion alarms and the construction of intrusion scenarios is a new direction for the development of security analysis technology. Because of the difficulty of effectively extracting and analyzing the massive data generated by security events, we urgently need more efficient classification algorithms to achieve accurate extraction of attack signature data. In the field of data mining, there are many urgent learning methods, such as Bayesian classification, decision tree classification, and neural network classification, and analogy-based inertial learning methods such as KNN [6–8]. Because of its simplicity and operability, it is well applied in the field of data classification and pattern recognition. The KNN algorithm classifies categories by relying on neighboring samples rather than by discriminating class domains. There are more intersections to be divided into sample sets; KNN algorithm has greater advantages than other classification algorithms. Therefore, we introduced the KNN classification algorithm when classifying attack data, but the accuracy of the basic KNN classification data has room for improvement [9]. Many scholars have proposed various methods to improve the algorithm. The core of attack scenario reconstruction is to perform scene correlation. In recent years, researchers have proposed many alarm correlation methods, which can be roughly divided into three types. Causality, association, and data mining methods are based on similarity [10]. The basic idea of causality is to associate these two alarms when there is a causal relationship

between the previous alarm and the next alarm. Of some researchers, Peng et al. [11] applied causal methods, but had different prerequisites and consequences. In this approach, artificially defined rules increase the probability of false associations, especially on those that are weakly related. The key to the similarity approach is to associate alerts with similar attributes through clustering and aggregation. Similar alerts are considered the same threat behavior. The challenge with this approach is that the alarms are still not very strongly correlated. Data mining methods apply data mining techniques to alarm databases to mine the characteristics of alarms and group them by similar common features [12]. When dealing with a large number of alarms, improving the performance efficiency of data mining algorithms in alarm-related methods has become popular in the study [13, 14]. However, the existing intrusion attack scenario restoration technologies in smart grid applications are technically faced with the following difficulties: massive original alarm data, abruptness of alarm data, heterogeneity of original alarm data, and lack of a security model. In order to solve the above problems, we propose a new attack scenario restoration framework. The main contributions of this paper are as follows:

- A KNN-based attack data classification algorithm to reduce the number of missed or missed scores of attack data and add data expansion reorganization feedback based on expert rules.
- Implementations of attack scenario reconstruction based on state machine, with the causality correlation method and the similarity method.
- Simulation results based on the database provided by state grid company IDS verify the efficiency and accuracy of our framework.

The remaining of this paper is organized as follows. Section 2 constructs our attack scenario system model. Then, we mainly discuss the proposed rule certainty updating strategy. We will show experimental results in Sect. 4, and conclude the paper in Sect. 5.

2 System Model

2.1 Definitions of Attack Scenario

Alarm instance $A = \{a_1, a_2, \dots, a_n\}$, where a is the object associated with the specific alarm log. Analyzing the original attack logs through expert-defined regular expressions detects and creates each alert instance. Each a consists of typename, facts, level, pre, con, where typename represents the name of the alert type, and different alert types are defined by the same regular expression and marked with a unique name, such as SQL Injection; fact = $\{srcIP, srcPort, destIP, destPort, timestramp\}$, which represent the source IP of the alarm and its port ID, the attribute value of the order of the target IP and its port ID and timestamp, the level is the attack level of the corresponding alarm type, and pre indicates that the trigger can be triggered. The type of alarm for the current alarm type and con indicates the type of alarm that can be

triggered by the current alarm type. The alarm correlation graph $CG = (N, E)$ is a connected directed acyclic graph. The node set N represents a super alarm set, each pair of nodes $n_1, n_2 \in N$ in N , if and only if in E when there is a n_1 to n_2 edge, n_1 prepares for n_2 . Attack scenario is invasion process formed by a number of different attack behaviors in a certain order. Mainly in the form of an alarm association graph, an attack scenario may include one or more alarm association graphs. Attack scenario tree is a formal representation of the attack scenario. It is a static tree data structure defined by XML. It describes the process of multistep attacks. Each node in the tree is a matchable rule that shows the characteristics of an attack step and defines the attributes of events that can be matched by the rule [15]. The relationship between parent and child nodes is “and”, which represents the progression of the time series of attack steps. The relationship between siblings is “or” and it represents different options for the next attack step. The growth of the tree is the development of multistep attacks.

2.2 Attack Scenario Reconstruction Framework

As shown in Fig. 1, the most critical part of the entire framework is the intrusion analysis knowledge base, which is used by all other parts of the framework. The intrusion analysis knowledge base is a machine-readable knowledge base composed of intrusion ontology and knowledge base. The domain of intrusion analysis describes the structure and semantics of the concepts of these concepts. Complementary knowledge is a collection of different intrusion analyses related knowledge, assessment indicators, alarm verification and classification rules, and semantic inference rules. The specific attack scenario restoration process is as follows:

- 1. According to the knowledge base, verify the data after standardization and remove the false alarm data.
- 2. Classify KNN-based attack data and normal data according to the knowledge base, and reduce the data redundancy.
- 3. Use automatic state machines to correlate the pre- and post-attack behaviors based on semantic relationships and causality.
- 4. Finally, build attack patterns and attack scenarios based on the associated attack behaviors.

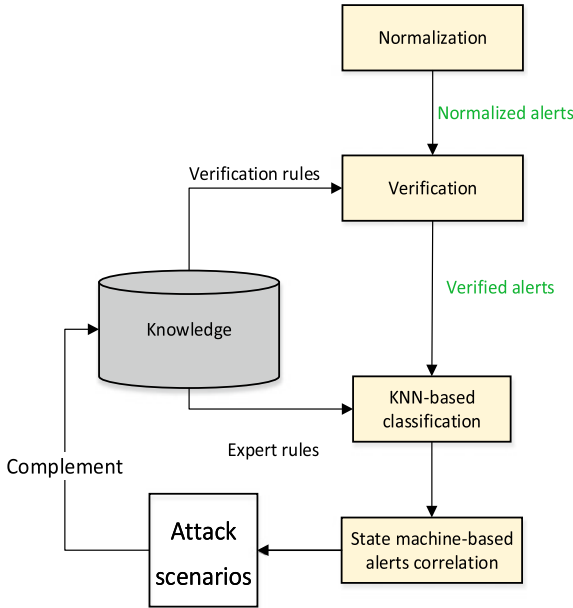


Fig. 1. Attack scenario reconstruction framework.

3 KNN-Based Data Classification

K-nearest neighbor KNN algorithm is a relatively mature data classification method in theory. The idea of the method is that there is a sample data set, also known as the training sample set, and each data sample concentration is tags, which know each data sample concentration and the corresponding relation of the category. Entering the new classification of sample data, the new sample data of each feature attribute will be corresponding characteristic value and the training sample set data set, and then classification algorithm to extract the sample concentration features the most similar data (most) adjacent category labels. When classifying, frequently choose only sample data sets before the highest number of K similarity. According to this, then choose K most similar data set in the classification of the most frequently as a new data set classification. In this article, it is known that there are N attack scenarios, and $K = N + 1$. There is still one type of normal data. The basic idea of this algorithm is to use the weighted KNN algorithm to classify historical alarm data. However, in the actual network attack scenario reconstruction process, other data associated with the attacked port are also often involved, so the normal data needs to be classified twice according to expert rules. The KNN attack data classification procedure based on expert knowledge is as follows: First, find the k-nearest neighbors: Calculate the Euclidean distance between two alerts. $A_1 = (a_{11}, a_{12}, \dots, a_{1n})$, $A_2 = (a_{21}, a_{22}, \dots, a_{2n})$. The Euclidean distance between two points is

$$dist(A1, A2) = \sqrt{\sum_{i=1}^n (a_{i1} - a_{i2})^2} \quad (1)$$

In actual operation, in order to prevent a large weight difference of features having different initial value ranges, the feature values are normalized as follows:

$$v' = \frac{v - \min_D}{\max_D - \min_D} \quad (2)$$

\max_D and \min_D represent the maximum and minimum values of feature A, respectively. v' represents the normalized feature value. In the second step, the classification is based on the categories of k nearby points:

$$C_X = \arg \max_{j \in I} \sum_{y \in A_k} I(C_y = j) \quad (3)$$

A_k is the k -nearest neighbor of y , and C is the label. $A_k = 1$ when the label values of y and j are the same. Otherwise, it is 0.

In the third step, at this time, we have classified the data into K categories based on the classical KNN algorithm. Among them, $K - 1$ attacks and normal data, and for the normal data set P , we defined the classification of attack data with i -dimensional features.

$$Disc_i = 1 - (pre_i - pre_t) \quad (4)$$

The pre_t represents the average accuracy of the classical KNN algorithm, and the pre_i represents the accuracy in the absence of i eigenvalues. According to expert rules, we increase or decrease the i -dimensionality of attack data features. When $pre_i - pre_t < 0$, namely, $Disc_i > 1$, it means that this feature is beneficial to improve the correctness of the final classification result to a certain extent, so as to increase the corresponding weight of the feature on the basis of the original value. This is reasonable. Otherwise, it will reduce the relevant weights. Distinguish $Disc_i$ and treat it as the weight of i -dimensional attack data characteristics.

$$dist(X1, X2) = \sqrt{\sum_{i=1}^n w_i (x_{1i} - x_{2i})^2} \quad (5)$$

Take this Euclidean distance as a new treatment and return to the second step. In the fourth step, the above operation is repeated. When pre_i is greater than a certain threshold, the process ends.

4 Attack Scene Reconstruction Based on Automatic State Machine

The state machine-based attack scene reconstruction technique proposed in this paper combines cluster analysis and causality association analysis to maintain the online detection of security events by maintaining associated state machine

queues, thereby reducing latency. First of all, we need to determine the existing attack scene tree and intrusion scenario library according to expert rules. Then the association algorithm will match the newly appeared security alarm with the attack scene tree. If the matching is successful, the new intrusion process can be restored. Each attack scenario corresponds to an associated state machine. The state machine queue represents the attack situation of the grid at this time. Each state in the queue corresponds to the state machine where the attack occurred.

Input: State machine queue list

Output: Attack scenario context information ScenarioInfo and update state machine sequence list

1. Read attack data

2. If list=0, find the appropriate attack scenario tree in the intrusion scenario library. If yes, execute 3

3. If the existing state machine K can handle the security event j, output the attack scenario ScenarioInfo, otherwise add the event j to the head of the state machine sequence list, skip to 5

4. When there is no state machine that can be associated with the list, the new attack scenario is added to the queue of the state machine.

5. Associate new attacks

In Algorithm 1, each state machine has more than one state. Therefore, we want to maintain a state set to store all states. When an intrusion event occurs, all the attributes of the current state machine are associated with the alarm event. If there is an association, the time counter is incremented by 1 and some features of the event are entered into the database. If the state of the event is the leaf node of the scene tree, the attack scenario description information is generated, and the leaf state generation description information is generated. The attack scenario description information records the complete path from the root node to the leaf node and records other important data. This associated time delay will be reduced to LRU, i.e., if the current state of the state machine is set to successfully handle the security event, because according to the “proximity method”, the probability that the next event will be trained locally is maximum.

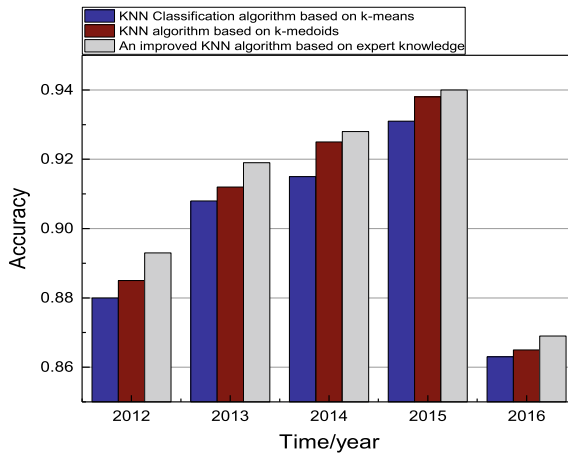
5 Experiment Result Analysis

To verify the effectiveness of the algorithm, experiments were performed on three data sets on the UCI common data set under matlab2016a. Computer-configured experimental environment, Core i3 CPU, memory 2G, operating system Windows 7, experimental data obtained in the experimental environment. The experimental data uses the 2012–2016 IDS dataset provided by the State Grid Corporation of China, as shown in Table 1. These data sets are suitable for experiments on the restoration of multistep attack scenarios. This attack scenario consists of four steps: scan detection, illegal access, installation of DDoS programs, and remote DDoS attacks on the target server. The intrusion data sets in the paper are shown in Table 1.

Table 1. Data sets in our experiments.

Years	Number of samples	Number of attributions
2010	1905	9
2011	6392	14
2012	9567	22
2013	11,294	31
2014	821	31

In the Matlab environment, in order to verify the effectiveness of the improved classification algorithm, two sets of dataset simulation experiments were designed and compared. The criteria for the performance of the classification algorithm are the efficiency of the algorithm and the accuracy of the classification results. In this experiment, the KNN classification algorithm based on expert knowledge proposed in this paper is compared with the other two algorithms (K-means-based KNN classification algorithm and K-Medoids-based KNN algorithm) for comparison of simulation experiments. The first set of data for the simulation experiment we selected is the National Grid IDS data set for 2013. The comparison of the algorithm test results is shown in Fig. 2.

**Fig. 2.** Accuracy results for each algorithm in the 2015 data set.

According to the results of the simulation experiment in Fig. 2, when the amount of data gradually increases, the improved classification algorithm proposed in this paper improves the execution efficiency and classification accuracy of the other two classification algorithms. The results of the second set of experimental simulation data are shown in Fig. 3. The control of the same variable

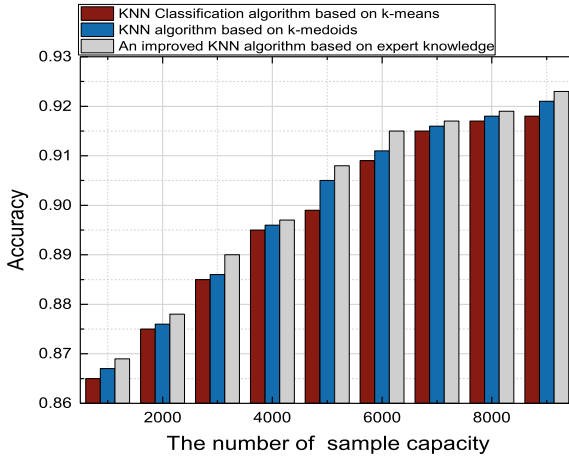


Fig. 3. Accuracy of three types of algorithms on various data sets.

is the same. The algorithm proposed in this paper has been enhanced in data classification accuracy. The experimental results show that our proposed algorithm can greatly improve the efficiency of classification learning without changing the precondition of classification accuracy. Compared with the K-Medoids-based KNN algorithm, the improved algorithm is slightly more efficient. The K-Medoids-based KNN algorithm needs to be calculated step by step until the class converges, which consumes part of the time. Through comparison between the two groups of experiments, the improved algorithm proposed in this paper improves the classification performance compared with the previous two algorithms. Therefore, the method proposed in this paper is effectively for extracting attack data.

6 Conclusions

In the big data environment, grid security is facing new challenges. Therefore, based on the knowledge of big data and machine learning, we propose a new framework for attack scene reconstruction: combine the classical KNN classification algorithm and combine expert rules to improve the accuracy of the classification. Therefore, we propose a new framework for attacking scene which is based on the knowledge of big data and machine learning. In terms of attack scene reconstruction, we use the state machine. The model realizes the correlation between security data and the reconstruction of attack scenarios. Through the analysis of the experimental results, the proposed framework can effectively face new security attacks and provide security for the safe operation of the power system.

Acknowledgements. The work is supported by State Grid Corporation of China Science and Technology Project: Research on Unknown Security Threat Detection Technology Based on Big Data Analysis (No. SGJSXT00JFJS1700101).

References

1. Yu, J., Fang, C., Lu, L., Li, Z.: Mitigating application layer distributed denial of service attacks via effective trust management. *AIET Commun.* **4**(16), 1952–1962 (2010)
2. Wang, K., Li, H., Feng, Y., Tian, G.: Big data analytics for system stability evaluation strategy in the energy Internet. *IEEE Trans. Ind. Inform.* (2017). <https://doi.org/10.1109/TII.2017.2692775>
3. Vimalkumar, K., Radhika, N.: A novel model for detecting application layer DDoS attacks. In: *International Conference on Advances in Computing, Communications and Informatics*, pp. 198–204 (2017)
4. Adhikari, U., Morris, T.H., Pan, S.: A causal event graph for cyber-power system events using synchrophasor. In: *PES General Meeting Conference Exposition*, pp. 1–5 (2017)
5. Hinton, G.E., Osindero, S., Teh, Y.-W.: A fast learning algorithm for deep belief nets. *Neural Comput.* **18**(7), 1527–1554 (2006)
6. Sukhbaatar, S., Makino, T., Aihara, K., Chikayama, T.: Robust generation of dynamical patterns in human motion by a deep belief nets. *J. Mach. Learn. Res.* **20**, 231–246 (2011)
7. Campo, G.L., Cristina, C., de Diego, I.M., Enrique, C.: Detecting denial of service by modeling web-sever behavior. *Comput. Electr. Eng.* **39**(7), 2252–2262 (2013)
8. Prasanna, K., Seetha, M., Siva, A.: CAPriori: conviction based apriori algorithm for discovering frequent determinant patterns from high dimensional datasets. In: *2014 International Conference on Science Engineering and Management Research (ICSEMR)* (2014)
9. Mao, X., Zhao, G., Sun, R.: Naive Bayesian algorithm classification model with local attribute weighted based on KNN. In: *Information Technology, Networking, Electronic and Automation Control Conference*, pp. 904–908. IEEE (2017)
10. Luna, J., Cano, A., Pechenizkiy, M., Ventura, S.: Speeding-up association rule mining with inverted index compression. *IEEE Trans. Cybern.* **46**(12), 3059–3072 (2016)
11. Li, K., Xie, P., Zhai, J., et al.: An improved AdaBoost algorithm for imbalanced data based on weighted KNN. In: *International Conference on Big Data Analysis*, pp. 30–34. IEEE (2017)
12. Ma, X., Liu, F., Qi, Y., Wang, X.: A multiobjective evolutionary algorithm based on decision variable analyses for multiobjective optimization problems with large-scale variables. *IEEE Trans. Evol. Comput.* **20**(2), 275–298 (2016)
13. Ao, W., Song, Y., Wen, C.: Distributed robust attack detection and reconstruction for a class of uncertain nonlinear interconnected CPSs. In: *International Conference on Collaboration Technologies and Systems*, pp. 1819–1824. IEEE (2016)
14. Wang, K., Shao, Y., Shu, L., Han, G., Zhu, C.: LDPA: a local data processing architecture in ambient assisted living communications. *IEEE Commun. Mag.* **53**(1), 56–63 (2015)
15. Zhu, L., Li, M., Zhang, Z., et al.: Big data mining of users energy consumption patterns in the wireless smart grid. *IEEE Wirel. Commun.* **25**(1), 84–89 (2018)



Present Situation Analysis and Future Development of the Command and Control Console

Wenli Jin and Fang Bai^(✉)

The 28th Research Institute of China Electronics Technology Group Corporation,
Nanjing, China
baifang@nuaa.edu.cn

Abstract. In the command and control system, the modeling design of the command and control console occupies an important position. It not only gives people a direct visual experience but also affects the speed and reliability of specific operations. The status of the command and control console was combined. The main design elements were given. The outstanding problems were analyzed. Proposals for development were put forward.

Keywords: The command and control console · The main design elements · Structure design

1 Introduction

The modern command and control system is a system with computer network monitoring as the core. There are many types and quantities of equipment to be configured, and they need to be installed on a centralized central command and control console. The console provides a good man–machine interface. Operators use the keyboard and mouse to conveniently input control management data, display the image processing, and establish a database to achieve centralized and unified monitoring and control of the network system.

In the command and control system, the modeling design of the command and control console occupies an important position. It not only gives people a direct visual experience but also affects the speed and reliability of specific operations. It involves various disciplines such as materials science, forming technology, finishing technology, ergonomics, psychology, aesthetics, and reliability design. The core content of this paper is to analyze the current development trend of the console, grasp many elements in the design, and solve the various contradictions in the design, function, operability, comfort, heat dissipation, and space requirements.

2 Present Situation Analysis

All command and control consoles in the command and control system are special consoles designed to complete the fixed tasks assigned to the system. The consoles have a wide variety, different models, and different types in the aspects of shape and dimension, module division, computer software and hardware platform, human-machine interface, and external interface. Practice has proved that this situation brings great difficulties to production commissioning, operation, and maintenance, and seriously affects the combat effectiveness and availability of the system. Therefore, in the process of the system from centralized and separate to distributed development, modularization, generalization, and lightweight have become the most emphasized elements of structural design [1–3].

The traditional command and control consoles mainly include enclosed cabinet consoles, modular assembly consoles, platform consoles, and lightweight consoles.

The enclosed cabinet command and control console is characterized by modular and custom design, whole assembly, multi-seat/straight line/multiple combinations according to seats and site requirements, metal platens, single hue, large volume of the packing box, and small field integration work. The disadvantage is that the joint is larger due to the side plate (Fig. 1).

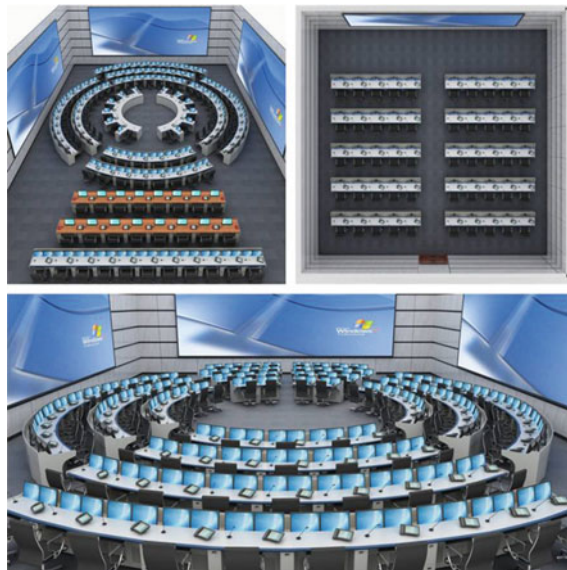


Fig. 1. The structure modular assembly console

The features of the structure modular assembly console are unique modeling, high-density refractory board (warm in winter and cool in summer), chalk-coated white paint, small packing box volume, and large field integration work. For example, the civil aviation airport controller seat is a high-order and custom cabinet. It uses the

skeleton, column, platform, door, lightbox, and bulkhead structure for modular assembly on the spot (Fig. 2).



Fig. 2. The platform-type console

The platform-type console has been preliminarily functional modularized, characterized by a modular assembly and adjustable multiscreen combination. The control area is modular, linear or corner combination, metal platen, two-color collocation, large volume of the packing box, and small field integration work. The latest type of platform-style console is designed through an auxiliary platform and standardized platform to achieve straight line, eight-shape, arc, ring, and other combinations.

The lightweight console is characterized by aluminum plate/aluminum profile welding, lightweight, unique shape, and poor aseismic strength. And the upper shelf, the middle shelf, the underframe, and the platen are each a module.

3 Design Elements

The design of the command and control console needs to meet many factors, such as ergonomic design, reasonable wiring, safety performance, durability, reasonable layout, reasonable heat dissipation, environmental protection standards, environmental adaptability, man-machine environmental engineering design, weight limit, space requirements, airborne/shipboard requirements, and electromagnetic compatibility requirements. It mainly emphasizes ergonomics design, reasonable wiring, and ergonomics engineering design.

3.1 Ergonomic Design

The mode of use should be suitable for the natural form of the human body. When the operator is working, the body and spirit do not need any active adjustment, so as to minimize the fatigue caused by the use of the console. So people are more comfortable, safe, and healthy in the work. A reasonable layout plan, including indoor function, station layout, console structure, and so on, should provide unique design concept.

3.2 Reasonable Wiring

Optimize the design of the cabling system, equipped with a reasonable horizontal and vertical ligation slot, to ensure that the power line and signal line have two-way walk, to ensure that they do not interfere with each other, and to achieve the signal security that strong and weak electricity is separated.

3.3 Human–Machine Environment Engineering Design

The working table height is 700–800 mm, and the depth is guaranteed to be between 350 and 700 mm. There is enough knee and foot space under the working table; the height is not less than 550 mm. The display should be arranged in the operator's best view, and the display surface is perpendicular to the normal line of sight of the operator as far as possible, so that the operator's parallax is the smallest. The effective sight distance of operators to monitors is in the range of 330–650 mm, and 510 mm is the best.

4 Problems

With the progress of science and technology and the further improvement of military and civil demand, the console's problems urgently need to be solved in four aspects: industrial design, three-defense design, structural design, and the application of new materials.

4.1 Industrial Design

Industrial design is a mixture of material and color. The console lacks comprehensive consideration in the aspects of technology, aesthetics, economy, and human nature.

The discovery and use of each new material will produce different methods of forming and processing, which will lead to great changes in product structure, bring new leap to the design of product design, and form a new design style. The comprehension of materials needs to be improved. Designers must understand and grasp the characteristics of materials to maximize the material characteristics and material matching.

The color matching in human–machine interface is related to whether information can be transmitted timely and effectively, and related to the function of products. At present, the design is not aware of the sense of color's temperature.

4.2 Three-Defense Design

The three-prevention design refers to the design of prevention of damp and heat, mold, and salt spray. It is an important measure of environmental adaptability of the console. The implementation of the design and process specifications is not in place, and the results are not good. The reprocessing in assembly is also a factor that affects the three defenses.

4.3 Structural Design

The main problems in structural design are as follows: Currently, there are many customized products and lack of integration. The console is not serialized and modularized, so it cannot adapt to the rapid development of the market. There is a lack of technical accumulation of equipment function institutions and systematic demand research. And the structure of the console is single and requires reserve and accumulation.

4.4 The Application of New Materials

In order to adapt to the characteristics of modern high-tech warfare, new materials (such as composite materials, aluminum alloys, and so on) with lightweight and superior performance must be used. The use of aluminum alloy plates for bending welding or forming of aluminum alloy profiles cannot take into account both the rigidity and the weight of the entire machine. Therefore, as a substitute material, composite material becomes the trend of development. Carbon fiber, as the main material of the ground or airborne command and control console, has a series of unique advantages such as high specific strength, high specific rigidity, good corrosion and fatigue resistance, and strong designability. Therefore, it is necessary to apply it in local areas [4, 5].

5 Development Trends and Suggestion

5.1 Unified Models of Design Pattern

Unified models of design pattern include the thickness of sheet metal, the height range of the table to the ground, the molding/material/thickness of the plate, the size/shape of the heat dissipation hole, the model of hardware, the mechanism of display and control device, the wiring, the panel screw, and so on.

5.2 Coordination with the Command Center Hall Design

The command and control console does not exist alone, and it needs to be placed in a fixed or a mobile command post. Therefore, the color matching design of the console should be bright, clean, warm-hued, and harmonized with the hall environment. Through the application of new materials and new technology, the functions of folding, lifting, rotating, and pitching are realized, and many factors such as intelligence, comfort, and no shield from the head are fully considered.

5.3 Generalization of Structural Forms and Spatial Forms

The console is divided into functional modules. It is mainly divided into display module, control module, and auxiliary module. The display module is inclined Island, the control module is keyboard drawer, and the auxiliary module adopts an integrated rack, a platen, and a decorative side plate.

The seat continues to be combined. The attachments should be standardized. The accessories include partition, guideway, door panel, and fastener. Stacking and progressive utilization of space can reduce height and extend depth. The slanted Island panel is modularized. By fully reducing the height of the display area, taking into account the viewing angle and range, meet the requirements of the boundary, and achieve the generalization of the spatial form.

5.4 Operational Functionalization

The display and control equipment can meet the requirements of electric or manual lifting: manual folding, folding up and down, and vertical lifting. Through the pedals, casters, leveling mechanisms, multifunctional stents, and other mechanisms, the display and control equipment's multifunctional needs can be realized: folding, lifting, pitching, rotating, walking, and so on.

Acknowledgements. Foundation item: Advance research project of the Ministry of equipment development in 13th Five-Year (Grant No.: 41404030101).

References

1. Hongni, H., et al.: Electronic manipulator ergonomic design method for display and control console. *Aviat. Sci. Technol.* **27**(1), 53–58 (2016)
2. Xiaoming, X.: General design of electronic command and control console cable components. *Syst. Solut.* **8**, 33–35 (2013)
3. Hua, C.: Modeling and humanization design of fixed command and control console. *New Technol. Process* 72–76 (2013)
4. Hui, Z.: The connotation and classification methods of command and control console, and the study of spectrum. *Telecommun. Technol.* **52**(10), 1706–1710 (2012)
5. Yi, Y.: The influence of screen layout of the command and control console on user's situational awareness. *J. Taiyuan Univ. Technol.* **47**(3), 399–404 (2016)



Research on Comprehensive Defense Technology of the Emergency Command Vehicle

Fang Bai^(✉), Yangyang Hu, and Ruigang Zhao

The 28th Research Institute of China Electronics Technology Group Corporation,
Nanjing, China
baifang@nuaa.edu.cn

Abstract. In response to the emergency command vehicle's on-site rescue and maintain orders requirements, the research overview, level, and development trend of domestic- and foreign-related technologies about the emergency command vehicle such as anti-sniper, anti-roadside bomb (car bomb), anti-laser radiation, anti-radar radiation, anti-electronic interference, and so on were analyzed. The composition of the comprehensive defense capabilities for vehicles on the highway is that our military's new emergency command vehicle should be proposed.

Keywords: The emergency command vehicle · Comprehensive defense · Anti-sniper · Anti-roadside bomb · Anti-laser radiation · Anti-radar radiation · Anti-electronic interference

1 Overview

From the “5.12” earthquake in Wenchuan to the “Tianjin explosion” incident to the “7.5” incident in Xinjiang, the sudden and catastrophic events caused by natural factors, technical reasons, and human factors have caused major loss of life and property and political influence. In 2005, the State Council reviewed and approved in principle the “Overall Emergency Plan for National Public Safety Emergencies”, various industries and fields have started to build mobile emergency command systems. The mobile emergency command system is a mobile command platform based on a fixed command center to implement mobile rescue command. The system is required to achieve the command and control of the superior command communication system in various non-military war missions and complete major safeguard activities, such as maneuvering to the scene of the accident quickly, organizing on-site rescue, maintaining stability, etc.

The challenge of the emerging security environment requires the rapidity, mobility, emergency, resilience, and accuracy of the mobile command platform, while also placing higher requirements on its own security.

2 Current Development Status at Domestic and Abroad

The comprehensive defense of the mobile command platform is a major issue involving the improvement of vehicles, equipment, tactics, procedures, training, etc. As a key attack target, whether active defense or passive defense is the key consideration for motorized transportation and command. Following the war in Iraq and Afghanistan, the protection technology of the vehicle has developed rapidly, and many countries are ahead of our country in the comprehensive defense for vehicles on the highway, especially western developed countries that have been baptized by war. In the meantime, the same experience and achievements that we can learn from are also abundant. The research overview, level, and development trend of domestic- and foreign-related technologies about the emergency command vehicle, including anti-sniper, anti-roadside bomb (car bomb), anti-laser radiation, anti-radar radiation, and anti-electronic interference, are analyzed one by one as follows [1–3].

2.1 Anti-sniper Reconnaissance

With the increase in the range and accuracy of sniper rifles and the emergence of high-performance viewing equipment, sniper threats are increasing. In combat operations such as conventional operations, peacekeeping operations, and anti-terrorism, especially in urban street fighting and jungle warfare, sniper can shoot the enemy, destroy key equipment, and delay the enemy's actions with high efficiency.

The French company Mtravib developed the anti-sniper sound detection system (PILAR). It can observe and record the bullet's flight trajectory in real time in the environment with a noisy background. Consequently, it can detect, locate, classify, and report the firing position of the small-caliber firearm accurately.

On January 21, 2004, based on the fixed "bullet ear" sniper detection system developed in the mid-90s, BBN Corporation successfully completed Boomerang 1,

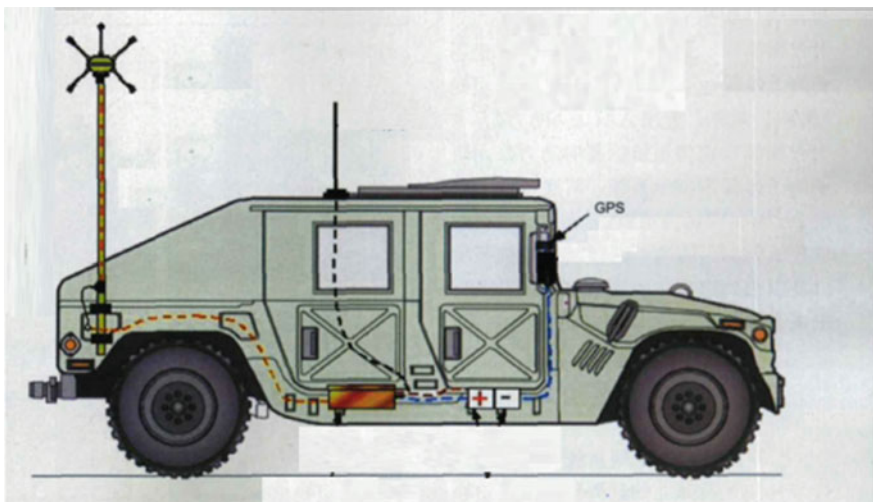


Fig. 1. Vehicle anti-sniper system (Boomerang)

which can work in high-speed vehicles and noisy urban environments. In March 2004, the Boomerang 1 system began to be equipped with the “Hummer” vehicle of the U.S. Marine Corps in Kuwait (Fig. 1).

The Spotlight system, developed by Rafael of Israel, is a high-precision infrared detection system that can accurately detect, identify, and locate multiple light weapons or sniper shooting sources at the same time. The infrared detection system responds quickly, has a high probability of detection, can work day and night, and has fire management capabilities.

French Laser Industry Company (CILAS) has developed the SLD 400, SLD 500 series of anti-sniper laser detection systems. It can detect sniper rifle sight hidden behind camouflage nets or installed with honeycomb panels, and can also detect other optical components such as night vision goggles, rangefinders, and telescopes.

The comprehensive use of various primary and passive detection methods can significantly improve detection accuracy and reduce response time. In February 2007, the U.S. military stationed in Iraq began trials of an anti-sniper detection system called “Red Bird.” The “Red Bird” system is similar in shape to a miniature remote-controlled tank and integrates various detection and positioning systems including acoustic detection systems, central cameras, laser rangefinder, and GPS systems. The French military has also combined the use of advanced SLD 400, SLD 500 series of laser detection systems with the sound detection system (PILAR), which has greatly enhanced actual combat effectiveness.

2.2 Anti-roadside Bombs

In early 2006, the EA-6B equipped with ICAP-III system was deployed to Iraq and Afghanistan. The U.S. military stationed in Iraq were often attacked by roadside bombs. The U.S. military discovered that anti-U.S. armed forces in Iraq and Afghanistan often used mobile phones and commercial networks to transmit intelligence, instruct the target of attacks, and used mobile phones to detonate improvised explosive devices buried on the roadside. Many of the U.S. military soldiers stationed in Iraq were killed by this “homemade bombs” while on patrol. According to incomplete statistics, since the beginning of 2005, Afghanistan’s anti-government armed forces have carried out more than 3,170 suicide attacks or roadside bomb attacks and 274 foreign soldiers and more than 1,500 Afghan civilians have been killed by the Afghan rebels. The intensity and frequency of terrorist attacks in Pakistan have leapt to the top in the world. More than 1,000 people have been killed and 1,720 people have been wounded in more than 70 terrorist attacks in Pakistan within a year. On average, there was a terrorist incident every 5 days, with an average of 16 deaths each time.

In the Iraq War, because the U.S. Army did not have enough electronic warfare technology capabilities, the U.S. Army relied mainly on the electronic equipment of the Air Force and Navy to destroy the enemy’s electronic targets and cleared up obstacles for ground operations. Hostile molecules buried a large number of remote-controlled roadside bombs between urban and rural areas to attack the U.S. Army troops. While using the improvised explosive device (IED) jammer, the U.S. Army has caused

serious interference with its own communications and other RF (radio frequency) systems. The U.S. Army's reliance on the Air Force fighters to detect roadside bombs by transmitting electron beams from an altitude of nearly 10,000 m was like using anti-aircraft guns to fight mosquitoes. The difficulty can be imagined. The Iraqi War made the Army realize that it is urgent to develop and improve its own electronic warfare capabilities without relying on the help of the Air Force and Navy.

In 2006, the U.S. Army set up an electronic warfare officer in the General Staff Operations Department and began to set up courses on anti-improvised explosive devices and electronic warfare for some brigade-level personnel. By the end of 2006, the United States had invested a total of 6.1 billion dollars to study the means to deal with roadside bombs. This figure of money was almost equal to the total funding for the "Manhattan Project" that the United States used to study atomic bombs during World War II.

In addition to installing electronic jammers on soldiers' vehicles to prevent or delay the start of the detonator, the U.S. military used sniffer robots, explosive-handling robots, military dogs, toy cars, and other means to deal with such relatively low-level weapons. In 2008, the United States issued a number of contracts to develop detachable and vehicle-mounted improvised explosive device (IED) countermeasure systems. In the same year, the U.S. Army Detonation Team was equipped with a new type of IED probe arm for detecting and removing IED in Iraq. The U.S. military has also developed new directed energy weapons against IEDs. The system uses three types of directed energy: microwave, laser, and electrical. This system could make it ineffective or destroy the improvised explosive device before vehicles are near the improvised explosive device. The "fire scout" is an improved version of the RQ-8A. It uses the reconnaissance equipment used on the "Global Hawk" and can cope with threats such as land mines and roadside bombs that are common in urban battles. However, the army-type development of the "fire scout" was not smooth, and it was expected that the initial combat capability would have been achieved by 2014 [4, 5].

2.3 Anti-laser Irradiation, Anti-radar Radiation

Optoelectronic confrontation has become a key factor in determining the outcome of modern warfare. The optoelectronic confrontation capability of the participating teams in the war is equivalent to their survivability. Western countries have developed a variety of optoelectronic countermeasure technologies and equipped with a considerable number of optoelectronic countermeasure devices. Currently, they are moving in the direction of broader spectrum, stronger countermeasures, better universality, and higher degree of integration.

Optoelectronic confrontation is a military operation in which the opposing parties use optoelectronic equipment and devices to carry out offensive and defensive operations in the optical band. From the equipment entity point of view, the optoelectronic countermeasure system can be roughly divided into four parts: alarm, stealth, interference, and destruction.

(1) Alarm

Photoelectric alarm technology is continuously developed with the application of light and electricity. The technology was first embodied in radar warnings. The development from the initial single-station ground-based radar to the later multi-based radar network has been going on for decades. After World War II, many countries have done a lot of research in the visible spectrum or in the ultraviolet, infrared, and other spectra.

The phased array radar with multi-target detection and tracking at the same time is currently the focus of radar development in various countries. Phased array technology has emerged as early as the late 1930s. The multi-function phased array radar is used in most of the new-generation mid-range and long-range air defense missile systems being developed. It has become an important symbol of the third-generation mid-range and long-range air defense missile systems.

Infrared alarm is an alarm technology that develops most rapidly after radar alarms. Various developed countries are racing to develop infrared warning systems, such as the TAILAND system in the United Kingdom, the Pirate system in Europe, the OSF system in France, and the IR-OTS system in Sweden. Particularly, the United States' AN/AAS-42 IRST system has been first deployed on the U.S. Navy's F-14D fighter aircraft. Currently, the fighter aircraft has been exported to 17 countries. The above-mentioned developed models are mostly onboard equipment, which can be used on ground vehicles with slight modifications.

Foreign countries attach great importance to the warning work of laser threat and have developed various forms of laser warning receivers. The LAser Homing And Warning System (LAHAWS) in the United States is representative of imaging laser warning devices. 100×100 CCD imaging devices and dual-channel background elimination measures are used in the system. In recent years, China has made considerable progress in the field of laser warning. However, most of the technical indicators have not reached the level of similar products in developed countries.

Compared to the single alarm system, the compound alarm system has many unparalleled advantages. It can fully utilize the advantages of various alarm modes while avoiding their defects. For example, the Multi-Sensor Warning System (MSWS) of South Africa's aviation is a compound warning device that integrates laser warning, UV warning, radar warning, etc. The warning system is a comprehensive threat warning system consisting of a radar warning system, a passive UV missile approach alarm system, and a laser warning system. In the 1990s, the U.S. Revenge Goddess is also a comprehensive integrated device that integrates alarming, tracking, and confrontation. This device is mainly airborne.

(2) Stealth

Stealth technology mainly includes radar stealth technology, infrared stealth technology, visible light stealth technology, etc. Since stealth technology was successfully applied to aircraft and warships and brilliant achievements were created in the war, a new wave of research on stealth technology has been set off internationally.

The initial radar stealth is to reduce the Radar Cross Section (RCS) through the rational design of the aircraft's appearance. Typical applications are the "diamond-

shaped” F-117A bomber and the “flying wing” B-2 Stealth and Strategic Bomber. Some countries, such as the United States and Russia, have now been able to simulate and continuously improve various methods of reducing RCS and establish certain design specifications.

A stealth car developed by the U.S. military has adopted radar absorbing paint on the surface of the car, and the outer cover of the stealth guard plate has adopted carbon fiber material that can absorb radar waves. According to tests, the car will be detected by the radar when it is 500 m away from the radar, while the ordinary car will be detected by the radar at 3000 m.

The appearance of infrared stealth technology can be traced back to the 1950s. For the first time in the United States, U-2 planes used baffles to block the rear of the exhaust system to change the direction of infrared radiation. British “Vickers” F-type main battle tank used infrared stealth materials including polystyrene and polyurethane foam. This kind of foam plastic is light in weight, easy to use, excellent in heat insulation and easy to color. Spraying, brushing, or adhering the foam to the surface of the car body can make the radiation characteristics of the body surface and the background close to or even identical.

At present, the main battle tanks and armored vehicles of various countries in the world are all coated with camouflage, the typical coating is the Temporary Camouflage Coating (TCC) developed by the UK’s Courtaulds Aerospace Company. Camouflage nets also have a better camouflage effect. The Movement Camouflage System (MCS) manufactured by Swedish SAAB company can significantly reduce the possibility of vehicles being found.

2.4 Anti-electronic Interference

The military actions taken by the U.S. military against Iraq, Yugoslavia, and Afghanistan show that the command and control facilities of the enemy in modern warfare have become the primary targets of attack. If the combat unit cannot accept the order and the commander cannot maintain contact with his lower forces, failure will not be avoided. The communication interference system is an electronic system used to generate and transmit interference signals and disturb the enemy’s radio communication. It can destroy the enemy’s voice communications and data links to effectively cut off the command and control commands.

Due to the dependence of various countries on GPS navigation systems, all countries around the world have been paying attention to GPS interference and anti-interference problems. The United States has been seeking various measures to improve the anti-interference ability of GPS systems. In early 2009, Raytheon Company signed a production contract to develop digital anti-jamming antennas for Air Force GPS receivers, and currently, more than 5,000 sets of equipment have been ordered including U.S. government and foreign military users. In order to deal with GPS interference, the United Kingdom introduced a new type of GPS interference monitor. This handheld interference monitor can monitor the interference signal on the L1 channel of GPS satellites and can display unintentional or intentional interference that attempts to interfere with GPS signals.

A radio direction finder is a device that can determine the direction of an enemy's radio radiation source. Typical products include the AN/PDR-12 wireless direction finding system in the United States, the PA005 communication direction finding machine in Germany, the AN/TSQ-114 land-based very high-frequency direction finding system (Trailblazer), and so on. The RC-12 tactical electronic reconnaissance aircraft belonging to the U.S. Army is used to communicate intelligence and measure direction. Its main task is to intercept radio waves emitted by enemy command and control systems, weapon systems, radars, and other electronic equipment, and to conduct positioning. The operating frequency range of Thomson's TRC-290/600 series direction finding system is extended to 0.1–1350 MHz, and the medium-speed frequency-hopping signal can be detected with an accuracy of $\pm 0.5^\circ$ (root mean square value). The WJ-1840 direction finding system of the United States operates at a frequency range of 5 kHz to 18 GHz, with high sensitivity (< -90 dBm) and a direction finding accuracy of 1.5° – 6° .

Domestically, a large number of researches have been conducted on the technical system and performance indicators of the radio direction finder, and a series of work has been carried out on the new spatial spectrum estimation technology, portable direction finding system, and software direction finding system.

3 Conclusion

In summary, the modern battlefield is ever-changing, and the defense requirements on the battlefield are also ever-changing. Different types of equipment require different protections. In the same equipment, due to different applications, the requirements of the protection system are also different. With the development of photoelectric integrated technologies, weapon systems, new material processes, and integrated technologies, various integrated defense systems have also been produced. It is difficult to learn from and purchase foreign defense technologies and equipment. Therefore, it is necessary to develop a new type of comprehensive defense system for vehicles on the highway with independent intellectual property rights.

In response to the threat of partial asymmetrical attacks in logistics support, military transport vehicles are the key target of attack. Military transport vehicles must have the integrated protection capabilities of plug and play, flexible reorganization, fracture resistance, interaction awareness, collaborative high-efficiency, and fighting multi-threat targets to meet future operational requirements. Its main capabilities include anti-sniping, anti-roadside bombs (car bombs), anti-laser illumination, anti-radar radiation, anti-electronic interference (communication between teams: satellite communications or wireless short-wave communication), real-time warning, and smoke camouflage. In the end, these capabilities and operational weapons are integrated to achieve targeted attacks.

Acknowledgements. Foundation item: Advance research project of the Ministry of equipment development in thirteenth Five-Year (Grant No.: 41404030101).

References

1. Deng, Y.X.: Thoughts on electronic protection measures of military targets. *J. Air Force Radar Acad.* **25**(4), 254–257 (2011)
2. Xiaoming, H.: Design of laser-guided bomb control system based on CAN Bus. *Aerodyn. Missile J.* **12**, 58–61 (2012)
3. Jianguo, Z.: Research on countermeasures of terrorist crimes in automobile bombs. *Soc. Sci. Rev.* **3**, 95–97 (2012)
4. Guolong, G.: Demonstration of anti-roadside bomb capability with the new US airborne anti-landmine system. *Infrared* **2**, 45–46 (2011)
5. Bo, Z.: Development of Air Bomb Testing Equipment. Harbin Institute of Technology (2011)



An Introduction of Cognitive Electronic Warfare System

Huaji Zhou^(✉)

Science and Technology on Communication Information Security Control Laboratory,
Jiaxing 314033, China
zhouhuaji1988@sina.com

Abstract. This paper expounds the connotation of cognitive electronic warfare system and development status at home and abroad, analyzes the framework of cognitive electronic warfare system, and introduces three closed-loop processes: cognitive reconnaissance, cognitive countermeasure, and cognitive effectiveness evaluation; finally, to explore the realization approach of cognitive electronic warfare system, proposes a cognitive electronic warfare platform architecture, and focuses on the analysis of the functional requirements of the platform and key technologies.

Keywords: Cognitive electronic warfare · Closed loop · Cognitive reconnaissance · Cognitive countermeasure · Cognitive effectiveness evaluation

1 Introduction

With the continuous advancement of science and technology, the electromagnetic environment in the battlefield is becoming more and more complex, and the level of intelligence in combat objects has been continuously improved, the means of conventional electronic warfare gradually become incompetent, and the combat effectiveness achieved has gradually declined. Therefore, there is an urgent need for a technical means to change the rules of the game to break through the bottleneck and promote the reform of the electronic warfare combat mode. The cognitive electronic warfare technology is generated under such a background. There is no clear concept of cognitive electronic warfare at present. It is generally believed that the cognitive electronic warfare system refers to the introduction of cognitive computing theory in traditional electronic warfare systems; its essence is to extract knowledge from a large number of raw sensor data through the perception of the operational environment, and then deduce the tactics of EW attack optimization strategies in real-time or near-real-time environment, to improve the equipment's self-adaptive ability, and then evaluate its effectiveness, and guide the next attack according to the assessment situation [1].

2 Development of Cognitive Electronic Warfare

The development process of cognitive electronic warfare is not long. Tracing back to the source, cognitive thoughts were first embodied in the field of cognitive radio, the core technology of which is the ability to perceive the surrounding environment and optimize their working parameters according to the environment. Then American scientist Simon Haykins integrated cognitive thinking into the field of radar design and proposed the concept of cognitive radar [2]. Then from 2009, the US military gradually introduced the concept of cognition to electronic warfare equipment in order to improve the cognitive ability and combat effectiveness of active equipment, and this also marks the formation of the concept of cognitive electronic warfare. At present, the US military has launched a series of cognitive electronic warfare projects: Adaptive Electronic Warfare Behavior Learning Project (BLADE, the combat targets are communication and improvised explosive devices) [3–5], Adaptive Radar Countermeasure Project (ARC, the combat targets are radars) [4], Urban Saber Project (Radio Frequency Spectrum Control) [4], Cognitive Jammer Project (CJ, first-generation cognitive jammer architecture) [6], Electronic Warfare Technology Project (based on adaptive machine learning algorithms) [7], Communication under Extreme RF Spectrum Conditions Project (COMMEX, Adaptive Anti-Jamming Communication) [8], and Destroyer SRx system (based on Software Radio) [9]. Some of these projects are progressing smoothly: On May 26, 2016, the US “Aviation Weekly” reported that DARPA is handing over technologies of communications project (CommEx) developed under the extreme radio frequency spectrum to customers, and the Pentagon decided to integrate this technology into Link-16 data link to protect widely used tactical data links from interference; on June 20, 2016, Lockheed Martin announced the success of Lockheed Martin Advanced Technology Laboratory (ATL) and DARPA conducted a series of flight tests at the government test range to demonstrate the ability of the BLADE system to perform spectrum operations more intelligently in face of spectrum challenges. According to market research firm (MarketsandMarkets) in Dallas, the cognitive electronic warfare system is expected to equip US troops at the end of 2017. However, research in the field of cognitive electronic warfare in China still remains at the theoretical stage, which requires domestic counterparts to make more efforts.

3 An Exposition of Cognitive Electronic Warfare System

The cognitive electronic warfare system is an intelligent electronic warfare system, which is formed on the basis of the traditional electronic warfare system by increasing the system cognitive ability such as target recognition, intelligent decision-making, and autonomous learning. The introduction of such techniques as machine learning and intelligent decision enables cognitive EW systems to have the ability to respond effectively to the time-sensitive targets, unknown targets, and cognitive targets. This part will elaborate on the cognitive electronic warfare system from the aspects of the system function framework and the main processing links.

3.1 Cognitive Electronic Warfare System Functional Framework

The functional framework of a typical cognitive EW system is shown in Fig. 1. The modules closely related to the system's cognitive abilities are dynamic knowledge base, cognitive agents, and resource scheduling and function managers. Next, we briefly introduce these three modules.

The dynamic knowledge base contains signal feature library, countermeasure case library, and attack strategy library. Among them, the signal features' library stores the typical technical characteristics of common signals, such as frequency, bandwidth, modulation, and waveform characteristics; countermeasure case library mainly saves typical target samples, effective countermeasures, available countermeasure data, and saves the countermeasure data and samples for specific combat objectives in the form of case samples; the attack strategy library mainly preserves the rules for selecting countermeasure strategies for different combat signal targets, and specific countermeasure strategies which can be used against software resources and attack data resources.

The cognitive agent includes reasoning module, learning module, optimization module, and strategy engine, and is responsible for the realization and coordination of various cognitive functions. Cognitive functions include reasoning, learning, optimization, and strategy generation.

The resource scheduling and function manager is responsible for the scheduling and control of the hardware and software resources of the entire system, and has intersections with all modules.

3.2 Main Processing Links

According to Fig. 1, the cognitive electronic warfare system can be viewed as a closed-loop system consisting of three OODA ("observation, positioning, decision, behavior") cognitive rings. These three closed-loop systems are cognitive reconnaissance loop, cognitive countermeasure loop, and cognitive effectiveness evaluation loop [10], and the dynamic knowledge base provides the corresponding target characteristics, environment, and strategy knowledge for the above three closed-loop systems. For ease of explanation, a simplified block diagram of a cognitive electronic warfare system is given here (Fig. 2).

Cognitive Reconnaissance Loop As the Chinese saying goes, know the enemy and know yourself, and you can fight a hundred battles with no danger of defeat. The cognition of the cognitive electronic warfare system to the environment is obtained through continuous interaction with the environment. Cognitive reconnaissance loop automatically will analyze complex electromagnetic environments (including target electromagnetic information and surrounding environmental information) to obtain electromagnetic environmental analysis results, which provide basis for subsequent processing of targets by cognitive reconnaissance loop. The cognitive reconnaissance loop mainly implements automatic signal classification, including model feature extraction, model matching, and sample training. The results of cognitive reconnaissance loop can be passed

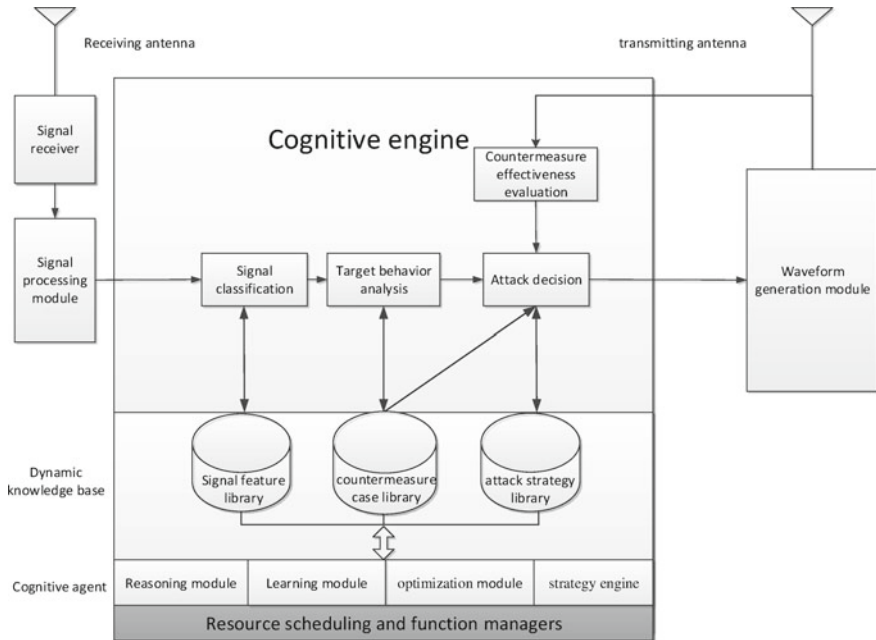


Fig. 1. Cognitive electronic warfare system functional framework.

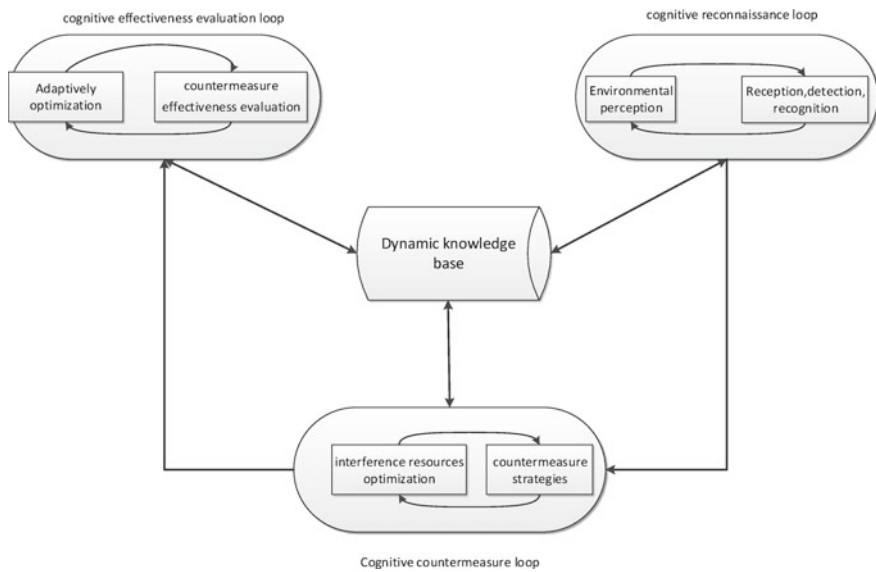


Fig. 2. Block diagram of cognitive electronic warfare system.

to the cognitive countermeasure loop and cognitive effectiveness evaluation loops as data support, and can also be passed to the dynamic knowledge base as accumulating samples.

Cognitive Countermeasure Loop The ability of cognitive electronic warfare systems to cognize and reason the battlefield electromagnetic environment enables cognitive electronic warfare systems to respond more effectively to various types of countermeasure targets and complex battlefield electromagnetic environments than traditional EW systems. Cognitive countermeasure loop continuously updates the decision database by constantly perceiving the surrounding environment. It mainly achieves the optimal countermeasure (interference or attack) against the target, including the generation and optimization of countermeasure strategies, the allocation, and optimization of interference resources. The generated countermeasure strategies can be passed to the cognitive effectiveness evaluation loop for evaluation and can also be passed to the dynamic knowledge base for storage as reasoning cases.

Cognitive Effectiveness Evaluation Loop After confronting the threat signal of the enemy, the cognitive effectiveness evaluation loop evaluates the effectiveness of interference, determine whether to use the electronic counter-countermeasures model, and infer the authenticity of the threat target, by obtaining changes in the threat signal under our interference, radar threat signal as an example: changes in pointing angle, recurrent frequency, and bandwidth. In addition, the cognitive effectiveness evaluation loop can develop targeted electronic countermeasures by optimizing the countermeasure strategies adaptively according to countermeasure effectiveness. The cognitive effectiveness evaluation loop mainly achieves self-adaptive and high-fidelity assessment of combat effectiveness. The evaluation results can be passed to the dynamic knowledge base as accumulating samples.

4 Realization Approach to Cognitive Electronic Warfare System

Compared with traditional electronic warfare systems, cognitive electronic warfare systems have obvious advantages: they can accurately perceive complex electromagnetic environments and have the ability to effectively confront time-sensitive targets, unknown targets, and cognitive targets. So how to achieve cognitive electronic warfare system? This section tries to answer this question in terms of platform architecture and key technologies.

4.1 Platform Architecture

In order to meet the functional requirements of the cognitive electronic warfare system, we propose a platform architecture which is shown in Fig. 3. The system

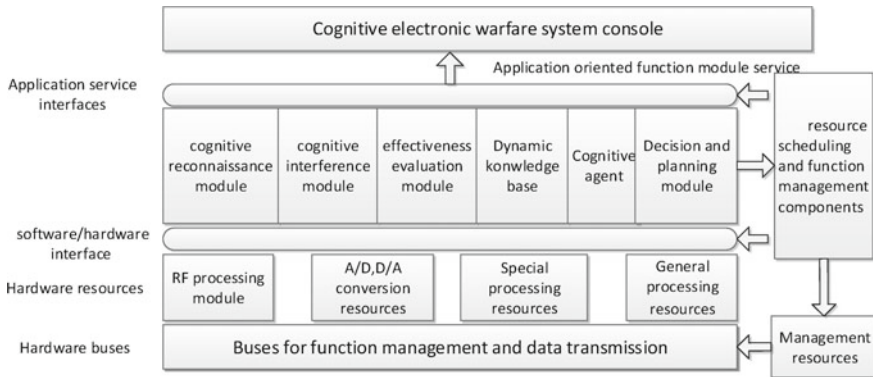


Fig. 3. Platform architecture of cognitive electronic warfare system.

consists of hardware busses, hardware resources, software modules, interfaces, and resource scheduling and function management components.

The hardware bus implements data transmission between hardware resources and provides support for platform monitoring and control as well as resource scheduling and function management. The isolation and mapping between software modules and hardware resources are accomplished through hardware and software interfaces. Connections and interactions between platform function modules are implemented through application service interfaces. The resource scheduling and function management component is responsible for the scheduling and management of software and hardware resources of the entire platform and all interfaces.

The platform architecture was born out of the traditional electronic warfare platform based on software radio technology. It adopts a modular design and is a generalized hardware based, functional software reconfigurable, and developed system architecture. The system uses a developable, scalable, and streamlined hardware as a common platform to implement functions such as cognitive reconnaissance, cognitive interference, and cognitive effectiveness evaluation using reconfigurable and scalable modular software. The system has a generalized structure, the functions are implemented flexibly, and the improvement of the system is improved; at the same time, since the main functions of the system are realized by software, various new signal processing means such as machine learning and deep learning can be conveniently used to improve the performance. And the system can achieve signal processing algorithm updates and upgrades online.

4.2 Key Technologies

According to the functional structure and operational characteristics of the cognitive electronic cognitive system, it is initially concluded that the technical architecture of the cognitive electronic cognitive system is shown in Fig. 4.

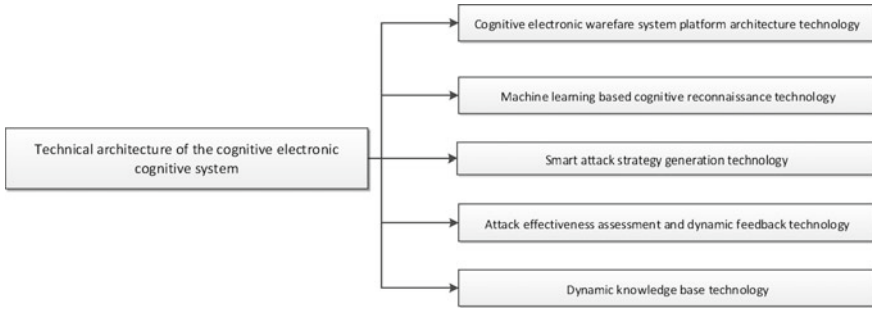


Fig. 4. Technical architecture of the cognitive electronic warfare system.

Cognitive Electronic Warfare System Platform Architecture Technology Cognitive electronic warfare needs to have flexible cognitive countermeasures, which are mainly reflected in related functions such as rapid detection of unknown targets, intelligent attacks and strategy optimization, attack effect evaluation, and dynamic knowledge base management. It is necessary to use software-based radio and theory and adopt standard technologies. Through hardware generalization, functional software, and software component design, an open and reconfigurable system architecture is constructed to ensure the flexibility and scalability of the system. A typical platform architecture is shown in Fig. 3.

Machine Learning-Based Cognitive Reconnaissance Technology In order to better describe the combat intent, movement characteristics, and functional characteristics of combat targets, electronic warfare systems need to acquire more information from complex electromagnetic environments and collect and sort out effective signals based on richer target characteristics. Through the threat feature learning and adaptive signal processing, the target signal characteristics are obtained. After collecting enough unknown signal features, the target signals are classified and identified through machine learning algorithms such as fuzzy clustering and neural networks, thereby constructing reconnaissance technologies with real-time cognitive abilities.

Smart Attack Strategy Generation Technology The technology for the development of the best electronic attack strategy based on the knowledge of the hardware, protocols, operating methods, software, and firmware of the combat target. Specifically, the technology is a kind of interference optimization and decision-making technology. According to the information that is perceived, the optimal interference strategy is given under constraints of system resources and user requirements. Ant colony optimization, neural networks, and other algorithms can be used to solve nonlinear interference resource allocation problems and achieve efficient and robust multi-target attacks.

Attack Effectiveness Assessment and Dynamic Feedback Technology

The battle damage effect of the current attack strategy is evaluated by the technology, and the evaluation result is fed back to the system to judge the performance of the previous attack strategy, and optimization learning is performed by giving a reward value and a penalty value, thereby ensuring system functions evolution. The attack effectiveness assessment model can be constructed through cloud model, game theory, and gray system theory.

Dynamic Knowledge Base Technology The dynamic knowledge base is a key module of the cognitive electronic warfare system that is different from the traditional electronic warfare system, and it is also a strong support for other cognitive modules of the cognitive electronic warfare system to effectively realize the cognitive function. The dynamic knowledge base includes signal signature database, confrontation case database, and attack strategy database, and stores the results of cognitive reconnaissance, cognitive interference, and cognitive effectiveness evaluation in different forms. The dynamic knowledge base is the basis of the system for learning and reasoning; it provides the basis for the function realization and optimization of each module of the system through pre-given prior knowledge and dynamic accumulation. The dynamic knowledge base system can be constructed based on the Oracle database system. And in order to facilitate the retrieval and maintenance of data information, the manual interface of knowledge base needs to be reserved.

5 Conclusion

At present, the development of artificial intelligence technology is in full swing, and the electromagnetic environment in the battlefield is becoming increasingly complex; the intelligentization and recognition of electronic warfare systems are the trends of the day [11]. This article elaborates the connotation of the cognitive EW system and its development status at home and abroad, analyzes the functional framework of the cognitive electronic warfare system, and introduces the three closed-loop processings: cognitive reconnaissance, cognitive confrontation, and cognitive effectiveness assessment. In order to explore the realization of cognitive electronic warfare system, a cognitive electronic warfare platform architecture is proposed, and the functional requirements and key technologies of the platform are emphatically analyzed. Taking into account the upgrade cost and simplicity, the architecture of the cognitive electronic warfare platform proposed in this paper is not significantly different from traditional electronic warfare platforms in hardware infrastructure. After the system compatibility is well coordinated, we can upgrade the software module in the existing equipment directly to achieve the renewal of the electronic warfare system and effectively improve the combat effectiveness of the system.

References

1. Xiangneng, Z., Yongshun, Z., Zewei, H., et al.: Electronic warfare new technology development overview. *Aerosp. Electron. Warf.* 5, 31–34 (2010)
2. Haykin, S.: *Cognitive Dynamic Systems*, pp. 232–235. Cambridge University Press, Cambridge (2012)
3. Zhang, C.: “Cognitive electronic warfare” kicked off—DARPA began to develop “intelligent jammer”. *Commun. Electron. Warf.* 1 (2011)
4. Manz, B.: Cognition: EW Gets Brainy. *J. Electron. Def.* 10 (2012)
5. DARPA.: Notice of Intent to Award Sole Source Contract: Behavioral Learning for Adaptive Electronic Warfare (BLADE) Phase 3. <https://www.fbo.gov/spg/ODA/DARPA/CMO/DARPA-SN-14-24/listing.html>, 19 Feb 2014
6. Air Force.: Cognitive jammer [EB/OL]. <https://www.fbo.gov>, 20 Jan 2010
7. ONR.: Broad Agency Announcement (BAA) NUMBER 13-005 Electronic Warfare Technology. <https://www.fbo.gov>, 19 Nov 2012
8. DARPA.: Broad Agency Announcement: Communications Under Extreme RF Spectrum Conditions (CommEx). Strategic Technology Office/DARPA-BAA, pp. 10–74. <https://www.fbo.gov>, 10 Sept 2010
9. Disruptor SRxTM.: Exelis, 4 Oct 2014
10. Zhang, C., Yang, X.: Research on cognitive EW and cognitive EW system. *J. China Acad. Electron. Inf. Technol.* 9(6), 551–555 (2014)
11. Ni, C., Huang, H.: Research on the composition and key technologies of cognitive electronic warfare system. *Shipboard Electron. Count.Measure* 6(36), 32–35 (2013)



The Application of Geomagnetic Models in Flight Path Planning

Zhiyuan Hang and Zhifang Wang^(✉)

Electronic Engineering College of Heilongjiang University, Harbin 150080,
China

xiaofang_hq@126.com

Abstract. The geomagnetic field is a vector field, and its elements in different regions can match the latitude and longitude uniquely (Fairfield and Mead in *J Geophys Res* 80(4):535–542, 1975) [1]. Therefore, an intuitive understanding of the geomagnetic information on exploration area is vital for geomagnetic navigation or flight path planning (Matteo and Morton in *Radio Sci* 46(4), 2016) [2]. The method proposed in this paper will get the geomagnetic information, which reflects the geomagnetic characteristics more intuitively after visualization. In addition, the characteristics of three geomagnetic models, EMM, WMM, and IGRF, are also compared and analyzed, which will be used for reference when planning the flight path (Kim and Chang in *Adv Space Res* 61(8), 2018) [3].

Keywords: Geomagnetic field · Geomagnetic models · Flight path planning

1 Introduction

The geomagnetic field, as an inherent resource of the earth, and its geomagnetic elements can be applied in the aspects of geological survey, mineral resources exploration, and so on. Geomagnetic model is developed to approximate or describe the temporal and spatial variation of the geomagnetic field.

The geomagnetic models EMM2017, WMM2015, and IGRF12 are used to describe geomagnetic information in this paper, and the characteristics of the three models in describing different regions are compared and analyzed [4]. The magnetic source described by EMM includes the core and the crust of earth, and it is expanded to 790 orders with spherical harmonic function. The magnetic source described by WMM includes only the core of the earth, and it is expanded to 12 orders with spherical harmonic function. IGRF describes a magnetic source that includes only the core of the earth [5].

2 The Application and Comparison of Geomagnetic Models in Flight Path Planning

The geomagnetic information to be compared in this paper is total field, geomagnetic inclination, geomagnetic declination, horizontal gradient, and vertical gradient. The testing regions are Harbin, Beijing, Shanghai, Hainan, Xi'an, and Urumqi. Take Beijing, Hainan, and Urumqi, for instance, the geomagnetic information obtained from different geomagnetic models of the same flight path in each region is as followed, where F indicates the total field, D indicates the magnetic deviation, I indicates the magnetic inclination, G_h indicates the horizontal gradient, and G_v indicates the vertical gradient [6] (Table 1).

Table 1. Geomagnetic information presented by different geomagnetic models in Beijing, Hainan, and Urumqi

		Beijing			Hainan			Urumqi		
		EMM	WMM	IGRF	EMM	WMM	IGRF	EMM	WMM	IGRF
F (nT)	Min	54,525	54,580	54,580	44,180	44,150	44,150	57,075	56,928	56,935
	Max	54,700	54,650	54,650	44,300	44,325	44,325	57,090	56,933	56,941
D (°)	Min	-7.18	-6.92	-6.72	-1.92	-1.90	-1.85	2.75	2.87	2.85
	Max	-7.12	-6.86	-6.86	-1.65	-1.65	-1.65	2.77	2.91	2.90
I (°)	Min	59.00	59.14	59.10	27.66	27.82	27.76	64.72	64.60	64.60
	Max	59.05	59.17	59.14	27.78	27.92	27.84	64.72	64.65	64.62
G_h (pT/m)	Min	0.00	0.00	0.00	-1.00	0.58	0.58	-4.00	0.00	0.00
	Max	2.60	0.22	0.24	-2.00	0.61	0.62	1.00	0.15	0.50
G_v (pT/m)	Min	19.40	26.20	26.17	11.32	12.42	12.37	24.10	28.06	28.06
	Max	19.80	26.21	26.18	11.42	12.42	12.37	24.50	28.08	28.08

2.1 The Visualization Process of Total Field, Geomagnetic Inclination, and Geomagnetic Declination

Two points are determined on the electronic map, and the track determined by two points is captured and interpolated, so that the longitude and latitude sequence of the track is obtained. Then, after the uniform interpolation between the beginning and the termination height of the input, the height sequence is constructed. The GPS information is formed by combining the longitude and latitude sequence with the altitude sequence, and then input the GPS information into the geomagnetic model to obtain the total field, geomagnetic inclination, and geomagnetic declination.

2.2 The Visualization Process of Horizontal Gradient

The method of determining the latitude and longitude sequence of the flight path is the same as above. Then we use the input altitude as the height of the entire track sequence. The GPS track information is formed by combining the longitude and latitude sequence

with the altitude sequence, and the horizontal magnetic field intensity of the track point is obtained by input the GPS information into the geomagnetic model. Then, the spherical distance of the adjacent track points on the sphere of earth can be calculated according to the following formula:

$$d = 2r \arcsin\left(\sqrt{\sin^2\left(\frac{\varphi_2 - \varphi_1}{2}\right) + \cos(\varphi_1) \cos(\varphi_2) \sin^2\left(\frac{\gamma_2 - \gamma_1}{2}\right)}\right) \quad (1)$$

where d indicates the distance between two points on the sphere of earth, r indicates the radius of earth, φ_1 and φ_2 indicate the latitude of two points, and γ_1 and γ_2 indicate the longitude of two points. Longitude and latitude are in radians. According to the distance and the difference value between two adjacent points, the horizontal gradient G_h of this track can be calculated.

2.3 The Visualization Process of Vertical Gradient

The longitude and latitude of a given point are set as the longitude and latitude of the whole track sequence, and the height sequence is generated by interpolation between the given starting and terminating height values. The vertical magnetic field intensity on the track can be calculated by inputting the complete vertical GPS track sequence into the geomagnetic model. According to the distance and the difference value between two adjacent points, the vertical gradient G_v of this track can be calculated.

3 Conclusion

After testing in these regions, most of the total field values described by EMM are slightly larger than those of WMM and IGRF, part of the reason can be attributed to the fact that EMM describes the more geomagnetic source. The values of the three models show a small difference for geomagnetic inclination and geomagnetic declination. Besides, there are some jump values in horizontal gradient calculated by EMM and most of the vertical gradient values described by EMM are smaller than those of WMM and IGRF.

According to the analysis above, when the flight path is planned according to the total field intensity, magnetic inclination, or magnetic declination, three kinds of geomagnetic models are all feasible. When the flight path is planned according to the horizontal gradient or vertical gradient, there are some jump values calculated by EMM, so WMM or IGRF is a better choice.

References

1. Fairfield, D.H., Mead, G.D.: Magnetospheric mapping with a quantitative geomagnetic field model. *J. Geophys. Res.* **80**(4), 535–542 (1975)
2. Matteo, N.A., Morton, Y.T.: Ionosphere geomagnetic field: comparison of IGRF model prediction and satellite measurements 1991–2010. *Radio Sci.* **46**(4) (2016)

3. Kim, J.H., Chang, H.Y.: Statistical analysis of geomagnetic field variations during solar eclipses. *Adv. Space Res.* **61**(8) (2018)
4. Pais, M.A., Morozova, A.L., Schaeffer, N.: Variability modes in core flows inverted from geomagnetic field models. *Geophys. J. Int.* **200**(1), 402–420 (2015)
5. Langlais, B., Manda, M.: An IGRF candidate main geomagnetic field model for epoch 2000 and a secular variation model for 2000–2005. *Earth Planets Space* **52**(12), 1137–1148 (2000)
6. Zhang, T., Gao, D., Zheng, J.: Simplified calculation method of geomagnetic field model. *Chin. J. Space Sci.* (2018)

Author Index

A

Abdussalam, Amr, 121
Alasaarela, Esko, 474
Ali, Safwan, 121
An, Weizheng, 564

B

Badenko, V., 315
Baidoo, Evans, 759
Bai, Fang, 1188, 1194
Bai, Xue, 482
Bai, Zhen, 870
Ba, Xuejing, 1063
Bi, Guoan, 942
Bi, Jianning, 804
Bo, Zhang, 983, 1015
Bu, Fan, 1031

C

Cai, Dongna, 898
Cai, Qiurun, 1089
Cao, Bei, 1153
Cao, Xiaozhong, 80
Chen, Bingcai, 962
Chen, Dongqiang, 175
Cheng, Jing, 105
Cheng, Yishun, 1073
Chen, Huili, 606
Chen, Jiahua, 911
Chen, Kaibai, 558
Chen, Sihao, 523
Chen, Siyun, 252
Chen, Xiaolong, 303
Chen, Xiu, 175
Chen, Yanlin, 683

Chen, Yidi, 225
Chen, Yu, 962
Chen, Zhengchuan, 950
Chen, Zhuang, 212
Che, Wen, 584
Chi, Yaping, 1063
Chong, Nannan, 881
Cui, Peng, 505, 656
Cui, Shaohua, 154, 753, 870, 875
Cui, Xinxin, 459
Cui, Yang, 128

D

Dai, Fei, 59
Dai, Fusheng, 1100
Dai, Lu, 950
Di, Chong, 25, 34
Ding, Hua, 25
Dong, Hairui, 836
Dong, Heng, 829
Dong, Liyuan, 459
Dong, Shunjie, 233
Dong, X.R., 813, 990
Dou, Xinyu, 422
Duan, Jianli, 933
Du, Jialu, 505, 656

E

Erkang, Fu, 193

F

Fang, Jinhui, 584
Fang, Yong, 606
Fan, Qiang, 128
Fan, Shangan, 1007

Fan, Shangang, 829
 Fan, Xingang, 745
 Fan, Yaping, 829
 Fan, Zhendong, 399
 Fei, Dongliang, 523
 Feng, Shuai, 42, 286
 Ferdinando, Hany, 474
 Fu, Meixia, 121
 Fu, Wei, 813, 990
 Fu, Yusong, 852

G

Gao, Han, 261
 Gao, Jean X., 279
 Gao, Jing, 707
 Gao, Junying, 922
 Gao, Yang, 791, 799
 Garg, R.D., 315
 Ge, Changyun, 499
 Gong, Guanghong, 836
 Gong, Lei, 1055
 Guangzhen, Wang, 538
 Guan, Jian, 303
 Gui, Guan, 50, 59, 87, 829, 1007, 1133
 Gu, Mingliang, 168
 Guo, Haonan, 42, 295
 Guo, Jing, 105, 113
 Guo, Ke, 105, 113
 Guo, Lei, 564
 Guo, Li, 791
 Guo, Liang, 1178
 Guo, Mingda, 34
 Guo, Ning, 707
 Guo, Wei, 999
 Guo, Xiaolei, 804
 Guo, Xiwei, 513
 Guo, Yongan, 852, 861
 Guo, Zijian, 628

H

Hang, Zhiyuan, 1211
 Han, Tingting, 384, 459
 Han, Xiao, 175
 Han, Zhuoran, 244
 He, Lianxing, 523
 He, Liu, 399
 He, Rongxi, 218, 405, 413
 He, Yadi, 270, 647
 Hou, Fen, 933
 Hou, Xueshi, 606
 Hou, Yu, 378
 Huai, Yongjian, 898
 Huang, Hao, 87, 1133
 Huang, Jinchao, 34

Huang, Jing, 468
 Huang, Xiaochen, 621
 Huang, Yali, 429, 436
 Huang, Yinguo, 687
 Huang, Yirui, 967
 Huang, Zhiyong, 687
 Hu, Dou, 891
 Hu, Han, 1178
 Hu, Jurong, 759
 Huo, Zhongxing, 1141, 1168
 Hu, Xiaocheng, 105, 113
 Hu, Yangyang, 1194

J

Jia, Jian, 999
 Jia, Li, 538
 Jiang, Aiping, 261
 Jiang, Bin, 59
 Jiang, Ting, 999
 Jiang, Yinan, 105, 113
 Jiang, Ying, 829
 Jiang, Youhui, 67
 Jiang, Yuan, 933
 Jiang, ZhanQing, 212
 Jianxi, Yang, 96
 Jia, Yunjian, 950
 Jing, Guo, 983, 1015
 Jing, Lei, 983, 1015
 Jing, Meng, 769
 Jin, Qianqian, 1178
 Jin, Shuqi, 753
 Jin, Wenli, 1188
 Jin, Yun, 168
 Ji, Yamin, 628
 Ji, Yuelan, 1141

K

Khoshghalbvash, Fariba, 279
 Kong, Zhen, 128

L

Lan, Tian, 184
 Lei, Fei, 175
 Liang, Jing, 204
 Liang, Liang, 950
 Liang, Wei, 492
 Li, Bing, 628
 Li, Bingyi, 724
 Li, Bo, 922, 975
 Li, Chunyang, 233
 Li, ChunYu, 212
 Li, Dapeng, 852, 861
 Li, Fangqi, 286
 Li, Guojun, 950

- Li, Haibo, 1007
 Li, Hongbo, 599
 Li, Huibo, 105
 Li, Jiaqi, 499
 Li, Jing, 716
 Li, Manhong, 621
 Li, Mei, 80
 Lin, Bin, 218, 933
 Lingfei, Zhang, 193
 Ling, Xiaofeng, 665
 Ling, Zhiting, 1063
 Lin, Ziwei, 413
 Li, Qi, 1117, 1126
 Li, Sen, 218
 Li, Shao, 564
 Li, Shenghong, 34, 42, 244, 286, 295
 Li, Shidang, 168
 Li, Shiyong, 159
 Li, Shuhao, 1063
 Liu, Deliang, 513
 Liu, Dun, 549
 Liu, Fengzhi, 218
 Liu, Huan, 683
 Liu, Jinlong, 1109
 Liu, Mingliang, 628
 Liu, Rui, 73
 Liu, Shuang, 80
 Liu, Sile, 429, 436
 Liu, Songyan, 683
 Liu, Tongtong, 413
 Liu, Wenlong, 891
 Liu, Wenyuan, 575
 Liu, Xiaoming, 638
 Liu, Xiaoning, 724
 Liu, Ying, 1178
 Liu, Yuxiang, 933
 Li, Wantong, 933
 Li, Weihua, 967
 Li, Xiang, 1046
 Li, Xinwei, 261
 Li, Xujie, 759
 Li, Yan, 429, 436
 Li, Yating, 745
 Li, Ying, 340
 Li, Yuliang, 348, 445
 Li, Yuning, 898
 Li, Yunyi, 87
 Li, Zhi, 898
 Li, Zhihuai, 911, 922
 Luan, Xiuzhen, 492
 Lu, Liguó, 1007
 Luo, Jin, 468
 Luo, Wang, 128
 Luo, Xiaolong, 687
 Lu, Shaoyun, 529
 Lyu, Fengcheng, 1031
- M**
- Ma, Chuang, 891
 Man, Chaoyan, 791, 799
 Mao, Weiyong, 592
 Ma, Yong, 168
 Meng, Huimin, 1073
 Mou, Xiaoqian, 303
 Mo, Yangan, 1022
- N**
- Ning, Bo, 1038, 1081
 Niu, Yifei, 683
 Ni, Xiaoyong, 759
- P**
- Pan, Jinqiu, 829
 Peng, Cunchao, 861
 Peng, Qiwei, 128
- Q**
- Qiao, Libo, 348, 445
 Qin, Danyang, 1153
 Qingqing, Lin, 769
 Qiu, Xiaoying, 999
 Qu, Fang, 549
- R**
- Ren, Jie, 204
 Ren, Qiang, 59
- S**
- Seppänen, Tapio, 474
 Sheng, Haoxuan, 606
 Shenghong, Li, 25
 Shen, Yabo, 168
 Shen, Yingli, 168
 Shen, Yongliang, 3
 Shi, Cheng, 1081
 Shi, Jifu, 474
 Shi, Jun, 1100
 Shi, Mingzhu, 80
 Shi, Zhenfeng, 1117, 1126
 Shi, Zhiping, 1031
 Shuai, Li, 769
 Shuai, Liu, 769
 Shuai, Weiyi, 813, 990
 Song, Lizhong, 592
 Song, Qi, 184
 Song, Qingyan, 606
 Song, Yang, 599
 Su, Debin, 740, 745

Sun, Honghao, 392
 Sun, Hongxia, 1031
 Sun, Houjun, 159
 Su, Ningyuan, 303
 Sun, Lingyu, 621
 Sun, Shang, 1117, 1126
 Sun, Songlin, 121
 Sun, Tianyuan, 724
 Suo, Jidong, 638

T

Tan, Kejun, 492
 Tao, Wen, 983, 1015
 Tian, Feng, 852, 861
 Tian, Xuanxuan, 184
 Tian, Yu, 1055
 Tu, Ya, 499

U

Ullah, Yasir, 121

W

Wang, Bowen, 459
 Wang, Guilin, 145
 Wang, Haixin, 482
 Wang, Haorong, 776
 Wang, Hongjin, 1073
 Wang, Jie, 87
 Wang, Jinpeng, 330
 Wang, Jun, 813, 990
 Wang, Linwei, 776
 Wang, Meng, 716
 Wang, Mengqi, 1022
 Wang, Nan, 665
 Wang, Niannian, 252
 Wang, Qian, 673
 Wang, Ronghe, 804
 Wang, Shimin, 612
 Wang, Shuguang, 159
 Wang, Wei, 348, 445
 Wang, Wenrui, 1089
 Wang, Wenyuan, 3
 Wang, Xiaowen, 683
 Wang, Xiao-yi, 740
 Wang, Xinyi, 891
 Wang, Yanwei, 699
 Wang, Ye, 252
 Wang, Yicheng, 687
 Wang, Ying, 784
 Wang, Yinshan, 529
 Wang, Yu, 87
 Wang, Yuanyuan, 270, 647
 Wang, Zhifang, 1211
 Wan, Liang, 967

Wei, Xiaotian, 911
 Wu, Bin, 422
 Wu, Pengfei, 1153
 Wu, Qiongzhi, 482
 Wu, Wei, 942
 Wu, Xuanli, 942
 Wu, Yanwen, 315
 Wu, Zhilu, 1109

X

Xiang, Yu, 18
 Xia, Yuanyi, 1178
 Xia, Yun, 673
 Xie, Xiaobo, 1007
 Xie, Yizhuang, 724
 Xie, Yongping, 18
 Xi, Leiping, 836, 845
 Xing, Hao, 638
 Xinyun, Li, 193
 Xiong, Jian, 50, 59, 1007
 Xiu, Yue, 3
 Xue, Lei, 145, 225
 Xue, Xinpeng, 330
 Xu, Hongguang, 1022
 Xu, Kun, 821
 Xu, Manyu, 784
 Xu, Wenhai, 340
 Xu, Zhengwu, 42, 286

Y

Yan, Di, 990
 Yang, Chaochao, 513
 Yang, Deshan, 340
 Yang, Hao, 354, 363
 Yang, Hongjuan, 975
 Yang, Hua, 468
 Yang, Jichao, 42, 295
 Yang, Jie, 50, 59, 87, 1007
 Yang, Jingping, 384
 Yang, Manrou, 962
 Yang, Mingda, 1038, 1081
 Yang, Mingqi, 813
 Yang, Sen, 836, 845
 Yang, Shuaijun, 529
 Yang, Sujuan, 50
 Yang, Yongjie, 612, 1141, 1168
 Yang, Yunxiang, 105, 113
 Yang, Zhutian, 1109
 Yang, Zi, 1038
 Yan, Zhicong, 286, 295
 Yao, Guo-wei, 673
 Yao, Nianmin, 1073
 Yaping, Chi, 96
 Ya, Tu, 505, 656

Ye, Liang, 474
 Yin, Hongxi, 422
 Yin, Yue, 87
 Yu, Changjun, 599, 776
 Yu, Cunqian, 405
 Yue, Dianwu, 716
 Yu, Hou, 538
 Yu, Jia, 252
 Yujia, Zhang, 193
 Yu, Ling, 799
 Yu, Muchen, 942
 Yunxiang, Yang, 983, 1015
 Yuqing, Zhang, 1015
 Yu, Shaopeng, 459
 Yu, Ying, 330
 Yuzhou, Yu, 96

Z

Zhang, Baoju, 145, 225
 Zhang, Bo, 105, 113, 804
 Zhang, Dizhou, 845
 Zhang, Feng, 673
 Zhang, Haijun, 87
 Zhang, Hong, 233
 Zhang, Jian, 204
 Zhang, Jie, 1133
 Zhang, Junjiao, 1168
 Zhang, Lei, 315
 Zhang, Li, 405
 Zhang, Lidong, 270, 647
 Zhang, Lin, 303
 Zhang, Linghui, 564
 Zhang, Minglu, 621
 Zhang, Qian, 1031
 Zhang, Qinyu, 175
 Zhang, Quan, 399
 Zhang, Ruijian, 875
 Zhang, Shuangshuang, 270
 Zhang, Shujun, 1031
 Zhang, Tai, 128
 Zhang, Tingting, 184, 1022
 Zhang, Wang, 340
 Zhang, Wei, 399
 Zhang, Xiaojun, 621

Zhang, Xin, 392
 Zhang, Xinhai, 113, 804
 Zhang, Xinqiang, 499
 Zhang, Xiu, 392
 Zhang, Xueying, 113
 Zhang, Yang, 10
 Zhang, Yu, 759, 791, 799, 967
 Zhang, Yueyong, 1046
 Zhang, Zhijian, 967
 Zhang, Zhong, 80
 Zhang, Zhongpei, 3
 Zhao, Chenglin, 154, 753, 870, 875
 Zhao, Dongmei, 168
 Zhao, Feichao, 270, 647
 Zhao, Gaofeng, 128
 Zhao, Jian, 575
 Zhao, Liang, 733
 Zhao, Nan, 87
 Zhao, Peng, 175
 Zhao, Qiuming, 975
 Zhao, Ruigang, 1194
 Zhao, Ruina, 564
 Zhao, Xiujie, 499
 Zhao, Yanan, 1100
 Zhao, Yongqiang, 212
 Zhen, Dan, 429, 436
 Zheng, Dewei, 154
 Zheng, ShiHui, 733
 Zhen, Jiaqi, 699
 Zhong, Xiaoling, 606
 Zhong, Xingjian, 399
 Zhou, Huaji, 1202
 Zhou, Tian, 50
 Zhou, Yiran, 1089
 Zhou, Yiyang, 1089
 Zhu, Liang, 128
 Zhu, Lin, 575
 Zhu, Pengli, 270, 647
 Zhu, Xiaomei, 218
 Zhu, Ying, 136
 Zou, Jie, 707
 Zou, Nianyu, 330
 Zuo, Jin, 384