

Lecture Notes in Electrical Engineering 1044

Jason C. Hung · Jia-Wei Chang ·  
Yan Pei *Editors*

# Innovative Computing Vol 1 - Emerging Topics in Artificial Intelligence

Proceedings of IC 2023

 Springer

## 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*

Samarjit Chakraborty, *Fakultät für Elektrotechnik und Informationstechnik, TU München, München, Germany*

Jiming Chen, *Zhejiang University, Hangzhou, Zhejiang, China*

Shanben Chen, *School of Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China*

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

Rüdiger Dillmann, *University of Karlsruhe (TH) IAIM, Karlsruhe, Baden-Württemberg, Germany*

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

Gianluigi Ferrari, *Dipartimento di Ingegneria dell'Informazione, Sede Scientifica Università degli Studi di Parma, Parma, Italy*

Manuel Ferre, *Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de 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, *Intelligent Systems Laboratory, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland*

Alaa Khamis, *Department of Mechatronics Engineering, German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt*

Torsten Kroeger, *Intrinsic Innovation, Mountain View, CA, USA*

Yong Li, *College of Electrical and Information Engineering, Hunan University, Changsha, Hunan, China*

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

Ferran Martín, *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*

Subhas Mukhopadhyay, *School of Engineering, Macquarie University, NSW, Australia*

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

Toyoaki Nishida, *Department of Intelligence Science and Technology, Kyoto University, Kyoto, Japan*

Luca Oneto, *Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genova, Genova, Genova, Italy*

Bijaya Ketan Panigrahi, *Department of Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India*

Federica Pascucci, *Department di Ingegneria, Università degli Studi Roma Tre, Roma, Italy*

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

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

Joachim Speidel, *Institute of Telecommunications, University of Stuttgart, Stuttgart, Germany*

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

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

Walter Zamboni, *Department of Computer Engineering, Electrical Engineering and Applied Mathematics, DIEM—Università degli studi di Salerno, Fisciano, Salerno, Italy*

Junjie James Zhang, *Charlotte, NC, USA*

Kay Chen Tan, *Dept. of Computing, Hong Kong Polytechnic University, Kowloon Tong, Hong Kong*

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](mailto:leontina.dicecco@springer.com).

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

### **China**

Jasmine Dou, Editor ([jasmine.dou@springer.com](mailto:jasmine.dou@springer.com))

### **India, Japan, Rest of Asia**

Swati Meherishi, Editorial Director ([Swati.Meherishi@springer.com](mailto:Swati.Meherishi@springer.com))

### **Southeast Asia, Australia, New Zealand**

Ramesh Nath Premnath, Editor ([ramesh.premnath@springernature.com](mailto:ramesh.premnath@springernature.com))

### **USA, Canada**

Michael Luby, Senior Editor ([michael.luby@springer.com](mailto:michael.luby@springer.com))

### **All other Countries**

Leontina Di Cecco, Senior Editor ([leontina.dicecco@springer.com](mailto:leontina.dicecco@springer.com))

**\*\* This series is indexed by EI Compendex and Scopus databases. \*\***

Jason C. Hung · Jia-Wei Chang · Yan Pei  
Editors

# Innovative Computing Vol 1 - Emerging Topics in Artificial Intelligence

Proceedings of IC 2023

 Springer

*Editors*

Jason C. Hung  
Department of Computer Science  
and Information Engineering  
National Taichung University of Science  
and Technology, FIET  
Taichung City, Taiwan

Jia-Wei Chang  
Department of Computer Science  
and Information Engineering  
National Taichung University of Science  
and Technology, FIET  
Taichung City, Taiwan

Yan Pei  
Computer Science and Engineering  
University of Aizu  
Aizuwakamatsu, Fukushima, Japan

ISSN 1876-1100                      ISSN 1876-1119 (electronic)  
Lecture Notes in Electrical Engineering  
ISBN 978-981-99-2091-4              ISBN 978-981-99-2092-1 (eBook)  
<https://doi.org/10.1007/978-981-99-2092-1>

© The Editor(s) (if applicable) and The Author(s), under exclusive license  
to Springer Nature Singapore Pte Ltd. 2023, corrected publication 2023

This work is subject to copyright. All rights are solely and exclusively licensed 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

## **The International Workshop on Technique for Language and Literature Information Modeling (LIM 2023)**

A Study on Data Mining for Type of Korean Painting Poetry .....	3
<i>Haeyoung Park, Younghoon An, and Hwayoung Jeong</i>	
A Big Data Based Learning Model from Student Questionnaire .....	10
<i>Hwa-Young Jeong</i>	
A Phonetic Investigation of Korean Monophthong Vowels by Vietnamese Female Speakers .....	17
<i>Juhee Lee</i>	
Korean Causal Connective Expressions in a Cross-Linguistic and Cultural Perspective .....	26
<i>Sujeong Choi and Sinhye Nam</i>	

## **ICIC2023-vol1**

Design and Implementation of Electromagnetic Field Simulation Platform for AC Testing .....	37
<i>Liang Zhang, Lijun Fu, Yanling Guan, Jingyuan Chi, Han Wang, and Jianquan Liang</i>	
Organizational Layout and Optimization Model of Agricultural Logistics Industry Based on Ant Colony Algorithm .....	44
<i>Jingjun Shu</i>	
Application of Image Recognition in Equipment Monitoring .....	54
<i>Haidong Zou, Shaoqiang Yang, and Wei Wu</i>	
Stability of Marine Physics Detection Sensor Based on Artificial Intelligence Technology .....	60
<i>Xiran Liu</i>	
Optimization System of Microbial Test on Account of Genetic Algorithm ....	68
<i>Mingming Shao</i>	
The Application of Virtual Reality Technology in Ophthalmology .....	76
<i>Jingying Wang and Yi Zhang</i>	

Research on the Application of BIM Technology in the Whole Process Cost Management of Construction Project .....	84
<i>Kangyan Zeng and Zhen Wen</i>	
Modal Parameter Identification of Bridge Structure Based on Hybrid Genetic Algorithm .....	92
<i>Rong Hu</i>	
Overload Damage Detection Method of Motor Car Axle Based on Neural Network Algorithm .....	102
<i>Pin Xia</i>	
Ground Wire Status Object Detection Based on Cloud Application .....	110
<i>Haoran Li, Shengli Xie, and Qian Wang</i>	
Design of Multimedia Assisted Course Based on Immersion Japanese Teaching .....	120
<i>Xiaomin Zhi, Liang Gao, and Guohe Qin</i>	
Analysis and Research of Intelligent Distribution UAV Control System Based on Optical Flow Sensor .....	128
<i>Shuai Sun, Yanqi Han, Yue Li, Wanting Zhao, Qifeng Gao, Yang Yang, and Heng Li</i>	
Method of Improving Logistics Distribution Efficiency by Combining Small Multi-rotor UAV with the Visual Recognition System .....	138
<i>Shuai Sun, Weilun Xia, Qiankun Zhu, and Zeping Han</i>	
Prediction Model of Popular Tourist Attractions Based on Big Data Fusion Algorithm .....	145
<i>Yan Wang</i>	
Application of Genetic Algorithm in Rail Transit Comprehensive Monitoring System .....	152
<i>Hui Lin</i>	
Construction of Inland VHF Communication System Based on Opposite-Reason Model .....	159
<i>Chen Kong, Zhongli Yi, Xiqi Liu, Shengbo Qin, and Shanshan Wang</i>	
Intelligent On-Line Monitoring System of General Electrical Equipment Based on Remote Communication .....	168
<i>Xinhe Cui, Fei Xia, Fanbo Meng, Xiaobo Huang, and Xin An</i>	

Sports Video Moving Target Detection and Tracking Based on Particle Filter and Related Algorithms ..... 176  
*Xiaocheng Wang*

Visual Reconstruction Design Based on Image Technology Emotion ..... 184  
*Meng Hou*

System of Cross-Border E-commerce Network Pattern Evolution on Account of Bayes-BP Algorithm ..... 191  
*Lina Zhang*

Construction of Internet Financial Risk Early Warning Model Based on Data Mining Algorithm ..... 199  
*Jingshuang Liu and He Yu*

Research on an Intelligent Logistics Blockchain Consensus Algorithm ..... 206  
*He Yu and Jingshuang Liu*

Improved ID3 Algorithm in Sports Tourism Service System ..... 215  
*Quanan Gui*

Design of Public Sports Service System in the Era of 5G + Artificial Intelligence ..... 223  
*Xinying Liu*

Automatic Design System of Product Shape Based on Intelligent Algorithm ..... 231  
*Ren Li*

Sports Video Tracking Technology Based on Optimized Decision Tree Algorithm(DTA) ..... 239  
*Zhong Wu*

Entrepreneurship Platform Under Virtual Reality Technology (VRT) ..... 246  
*Lili Wang*

A Model of Diversified Fiscal and Taxation System Based on Computer Network Technology ..... 254  
*Jianwei Zhang*

Research on Smart City Development and Construction Driven by Big Data Technology ..... 262  
*Yongling Chu, Zhiling Luan, and Shaochun Li*



Discipline Inspection and Supervision Big Data Information Query Platform Based on Distributed Computing .....	269
<i>Changyu Liu, Jinhua Chen, Hongchang Wen, Fanjun Zeng, and Sanjiong Wang</i>	
Digital Audit Platform Based on Visual Data Analysis .....	280
<i>Qi Liu, Jinhua Chen, Hongchang Wen, Guodong Qi, and Yujing Li</i>	
The Division of Labor and Matching System of Teaching Reform Model Based on Multi-objective Classification Algorithm .....	291
<i>Min Liu</i>	
Domestic AI Chip Based on Internet of Things Technology and Artificial Intelligence Algorithm in Power Grid Infrastructure Construction .....	299
<i>Weihua Zhong, Jun Liu, Jianxiang Xie, and Ji Zhang</i>	
Data Compression Algorithm of Power System .....	307
<i>Xinyu Zhang, Yanduo Xia, Shuangwu Quan, and Guozhi Zhang</i>	
Bank Credit Risk Prediction Based on MCDM and CNN .....	315
<i>Yuyue Cong</i>	
Application of NC Machining Technology in Aircraft Manufacturing .....	322
<i>Weiwen Ye</i>	
Barriers of Mathematical Language Representation of College Students Based on Data Mining .....	328
<i>Haiyan Wu</i>	
Mathematical Model (MM) and Signal Processing (SP) in Computer Simulation (CS) on Account of Improved Genetic Algorithm (IGA) .....	334
<i>Bo Wang</i>	
English Translation Intelligent Recognition Model Based on Reverse Maximum Matching Segmentation Algorithm .....	342
<i>Ruichao Li</i>	
A New Retraceable Clustering Algorithm .....	350
<i>Yefan Liu, Yijing Liang, and Yina Yin</i>	
Digital Transformation of the Smart Home Industry .....	357
<i>Haijie Shen</i>	

Identity Authentication Method of IoT Equipment Based on Alliance Chain .....	363
<i>Zhongzheng Xiang, Yanjun Hu, Xiaokang Chai, Kai Zheng, and Xuemei Shen</i>	
Financial Accounting Risk and Prevention System Construction Analysis Based on Fuzzy Influence Diagram Evaluation Algorithm .....	372
<i>Sitong Chen</i>	
Development of Rural Labor Resource Optimization Platform Based on Spring Boot Technology .....	381
<i>Jingyu Ye, Wei Huang, Jian Sun, Bin Huang, and Longfei Li</i>	
Research on Performance Enhancement of Parallel Computing in Intelligent Proofreading System .....	389
<i>Zhishuai Guo and Fucheng Zhang</i>	
Ethical Reflections on Artificial Intelligence: Practical Dilemmas and Mitigation Strategies .....	396
<i>Hao Cheng</i>	
Analysis of the Direction of Library Development in the Future Based on Knowledge Graph Analysis .....	402
<i>Jie Dong and Lingxiang Luo</i>	
Integrated Energy Service Business Applicability Based on MRW Model .....	408
<i>Hongcai Dai and Rui Chen</i>	
Application of Internet of Things and Virtual Reality Technology in the Intelligent Reconstruction of Old Residential Quarters .....	417
<i>Ming Lv, Miao Yu, and Wenting Ding</i>	
Construction of English Teaching Multimedia Resource Library Based on Cloud Computing Technology .....	424
<i>Huan Jiang</i>	
Design and Application of Japanese MOOC Teaching Resources System Based on User Collaborative Filtering Recommendation Algorithm .....	431
<i>Xuemei Zhang</i>	
A Study of Teaching Resource Retrieval System Based on Semantic Web .....	439
<i>Weiwei Zeng and Shirui Peng</i>	

Construction of Indoor Space Atmosphere by UVLED Light Source Design .....	448
<i>Yu Gui, Yangyang Fan, Jian Tian, Zonghui Bo, Yuxin Miao, Yuxiao Fan, Shunzi Ren, and Xi Chen</i>	
Seismic Stability Assessment of Civil Building Projects Based on BIM Technology .....	456
<i>Rongguo Zhao</i>	
Design and Implementation of Virtual Simulation Teaching Platform of Sports Action Technology Based on Maya .....	464
<i>Sicong Shan and Shuang Sun</i>	
Industrial Economy and Regional Growth Strategy Based on Particle Swarm Optimization Algorithm .....	471
<i>Hengran Bian and Yi Liu</i>	
Design and Implementation of Online Adaptive Teaching System Based on Data Clustering Algorithm .....	479
<i>Jie Ding</i>	
Design and Implementation of Online Teaching System Based on J2EE .....	486
<i>Qingfeng Guo</i>	
Design and Implementation of Classroom Teaching Line Analysis Platform Based on Pose Recognition Algorithm (MCG-CPR) .....	493
<i>Xiaobo Sun</i>	
Double Attention Mechanism Text Detection and Recognition Based on Neural Network Algorithm .....	500
<i>Yongjun Qi, Hailin Tang, and Li Huang</i>	
A Text Detection and Recognition System Based on Dual-Attention Mechanism with Artificial Intelligence Technology .....	511
<i>Yongjun Qi, Chenggao Li, and Li Huang</i>	
Data Transmission and Isolation Method for Power Bus Communication .....	522
<i>Yuxuan Ye, Rongtao Liao, Lei Li, Shengwei Wang, and Jian Zhang</i>	
Configuration Method of Safety Link of Electric Power Communication Network .....	530
<i>Huanjun Hu, Shengwei Wang, Yixi Wang, Lei Li, and Hao Ning</i>	
Method for Locating Communication Device Faults .....	538
<i>Lei Li, Zheng Zhou, Yixi Wang, Shengwei Wang, and Jian Zhang</i>	

Digital Twin Simulation System for Computer Room Equipment ..... 547  
*Lei Li, Zheng Zhou, Rongtao Liao, Yuxuan Ye, and Huanjun Hu*

Fault Detection Methods for Communication Equipment ..... 555  
*Hongyu Ke, Xi Zhang, Jing Dai, Chenxi Dong, and Lei Li*

Design and Implementation of Teaching Assistant System Based  
on Intelligent Recommendation Algorithm ..... 564  
*Jie Ding*

Construction of User Profile for Teaching Archive Knowledge Service  
Domain ..... 571  
*Wenyao Guo, Liming Du, Fengying Wang, Lei Cui, and Jie Dong*

Artificial Intelligence Technology in Computer Network Security ..... 579  
*Cuijin Lao and Shen Qin*

Dynamic Capacity Expansion System and Control Method for User-Side  
Power ..... 587  
*Rongtao Liao, Huanjun Hu, Yixi Wang, Yuxuan Ye, and Jian Zhang*

Secure and Non-secure Data Processing of Power Bus Communication ..... 597  
*Guowei Zhu, Dangdang Dai, Jing Li, Jie Huang, and Shengwei Wang*

Dynamic Monitoring Method for Power Big Data Quality ..... 606  
*Ting Huang, Li Yuan, Yujia Zhang, Fei Xie, Zhengqin Zuo,  
Yuefeng Peng, Shengjie Yang, and Haige Ren*

Detection of Load Loss of Large Capacity Transformer ..... 614  
*Feifeng Wang, Shengchao Jiang, Haoze Zhuo, Tailin Li, Bin Wang,  
and Yunqing Pei*

Intelligent Power Measurement System ..... 622  
*Chengwu Zou, Dan Ye, Qing Shen, Naiqing Wang, Jiaqi Lu,  
Mingquan Zeng, and Hangkang Li*

Routing Algorithm for Connected Vehicles Based on Bus Route  
Information ..... 633  
*Ning Tong, Man Jin, and Kunrui Wang*

Processing Method of Computer Data ..... 642  
*Xiaolin Jiang*

Storage System of Computer Data ..... 651  
*Peng Wang*

Design of Human-Computer Interaction Experience Platform in Virtual Space Under Scientific Information Technology .....	659
<i>Yi Fu and Yuzhuo Jiang</i>	
Design of Intelligent Color Matching System for Cultural and Creative Products Based on Data Analysis Algorithm .....	667
<i>Xuan Wan and Yantao Liang</i>	
Design of an Intelligent Educational Virtual Classroom Platform Based on the Mobile Terminal APP of the Internet of Things .....	675
<i>Xu Xing, Xilin Zhang, and Qiong Zhang</i>	
Design and Development of English Autonomous Learning Database Framework Based on Artificial Intelligence Algorithm .....	684
<i>Ximeng Wen and Shaoxian Hong</i>	
Devise and Application of Intelligent Guidance System in Exhibition Space .....	692
<i>Heng Shen</i>	
Configuration of Crop Supply Chain Based on Gale-Shapley Algorithm .....	700
<i>Wenting Kan and Bin Xiao</i>	
Application of Particle Swarm Optimization in BIM Building Modeling .....	707
<i>Guang Yang and Xin Guo</i>	
Intelligent Prediction Model of Agricultural Environment Based on Intelligent Algorithm .....	716
<i>Shuxin Zhang and Fangbo Hou</i>	
Mathematical Modeling and Optimal Design Based on Information Technology Algorithms .....	721
<i>Hua Li</i>	
Internet of Things Data Mining Technology in Sports Physical Health Analysis .....	729
<i>Qian Zhang</i>	
Application of Artificial Intelligence Technology in Big Data Nining .....	737
<i>Xueyun Zhou, Yongjun Qi, and Hailin Tang</i>	
Enterprise Credit Rating Prediction Model Based on Data Mining Algorithm .....	745
<i>Yingying Song</i>	

Application of Machine Learning Algorithm in Financial Market Risk Prediction .....	752
<i>Juyi Wang</i>	
Credit Risk Assessment Model of Supply Chain Finance Under the Background of Big Data .....	759
<i>Jun Ma</i>	
Large Screen for 3D Data Visualization Based on RFG-SVM Algorithm .....	767
<i>Guojuan Wang, Lijun Wang, Heqi Zhao, Qi Mu, and Xiangchun Ji</i>	
Construction of Inspection Work Management System Based on Data Mining Technology .....	775
<i>Hongchang Wen, Huiwen Peng, Ying Zhu, Yameng Wan, and Hao Niu</i>	
Constructing and Evaluating Service Quality Evaluation Index System of Airline Using PSM and Entropy Weight .....	784
<i>Liying Wang, Hongtao Li, Jing Bai, Xingming Lu, Pengpeng Zhou, and Zhijie Liu</i>	
Improving Service Quality Measurement with Servqual: Evidence from Airport Industry .....	792
<i>Liying Wang, Hongtao Li, Xingming Lu, Chenxi Zheng, and Licheng Xu</i>	
Research on Visualization of Student Management System Based on Distributed Storage Technology .....	800
<i>Fengxia Zhang, Xiaoguang Chen, and Ningning Tong</i>	
Banking Fintech Capability Evaluation System Based on Data Mining .....	805
<i>Jie Fan</i>	
Technology Financial Service Platform Based on Data Mining Algorithm .....	815
<i>Shuyu Hu</i>	
Overlapping Region Extraction Method of 3D Laser Point Cloud Registration on Account of Artificial Intelligence Algorithm .....	823
<i>Yaguang Wang, Wuzhan Yu, Yongdong Liang, and Bing Liu</i>	
Smart Urban Carbon Emission Management Platform Based on Energy Big Data .....	831
<i>Xue Li, Zheyong Piao, Yushan Zheng, Jieping Han, and Rijie Cong</i>	

A Distribution Network Reconfiguration Method Considering Optimal Distribution of DG Output .....	841
<i>Chao Feng, Yuanhong Hou, Shijun Li, Hui Xue, Yulai Yang, and Shunhai Xue</i>	
A Study of Intersubjectivity in Translation Based on Communicative Behavior Theory and Computer Technology .....	848
<i>Lu Liu, Nan Peng, and Jing Guo</i>	
Analysis and Research on Children’s Cognitive Psychological System Based on Digital Learning Terminal .....	855
<i>Zixin Yang</i>	
Analysis of E-Commerce Live Broadcast Marketing Strategy Based on Big Data Algorithm .....	862
<i>LinLin Sun</i>	
Analysis of Flow Data of E-Commerce Project Incubation Park Based on Computer Algorithm .....	870
<i>Xingrong Zhang</i>	
Analysis of Teaching Management Data Based on Decision Tree Algorithm .....	878
<i>Dan Yu</i>	
Analysis on Reasonable Planning and Design of Port Waters and Channels .....	884
<i>Wenting Liu and Zhiwei Liu</i>	
Ant Colony Algorithm Theory and Its Application in Control Engineering .....	890
<i>Chunxiang Huang, Nenjun Ben, and Guojun Yan</i>	
Application and Optimization of Style Transfer Algorithm in Contemporary Cultural and Creative Products .....	896
<i>Jinxia Chen and Gui Wang</i>	
Application of Ant Colony Algorithm in Chinese Language and Literature Education .....	902
<i>Qianyu Ma</i>	
Application of BP Neural Network Algorithm in Teaching Evaluation and Learning Prediction .....	908
<i>Zhangrong Hu</i>	

Application of Computer Aided Technology in Kindergarten Interior Design Under the “New Crown” Epidemic .....	914
<i>Shaowei Liu</i>	
Application of Data Mining Algorithm in Computer Mathematical Modeling .....	922
<i>Yanqin Yang</i>	
Application of Hybrid Recommendation Algorithm in College Life Safety Education Management System .....	928
<i>Shuchao Liu and Yuanyuan Xu</i>	
Application of Improved Genetic Algorithm in Distribution Network Planning with Distributed Generation .....	937
<i>Yuanhong Hou, Chao Feng, Yougui Aer, Shunhai Xue, Yuqiang Zhou, and Hongchao Wang</i>	
Application of Mathematical Simulation Method in Phosphate Ore Dressing .....	944
<i>Lingpan Du and Yongjie Guo</i>	
Application of New Light and High Strength Materials in Prefabricated Buildings .....	951
<i>Jia Liu, Mingfei Huang, Shanshan Li, and Jinfeng Chen</i>	
Application of Virtual Simulation Technology in Microbial Image Segmentation in Sewage .....	958
<i>Wenguang Xu</i>	
Application of Virtual Simulation Technology in Traditional Handicraft Protection .....	964
<i>Xiong Huang and Yuanjing Zhang</i>	
Big Data Visual Accounting Model for Ecotourism Carrying Capacity .....	971
<i>Yang Ye</i>	
Biomethanation System for Carbon Dioxide Capture in Renewable Energy Hydrogen Production Coupling Power Plant .....	979
<i>Changchun Xu</i>	
Children’s Safety in Smart Home Based on Intelligent Optimization Algorithm .....	986
<i>Wenxin Yang</i>	






Clustering Analysis of English Scores Based on K-means Clustering Algorithm .....	993
<i>Yueying Shen</i>	
College Students' Learning Engagement Based on Engineering Thinking .....	1001
<i>Mingyuan Ma</i>	
Comparison and Research of Roof Waterproof Materials Based on Computer Algorithm .....	1010
<i>Xinjun Wang</i>	
Comprehensive Evaluation System of Cultural Industry Development Based on Computer-Aided .....	1016
<i>Jianghong Yuan</i>	
Correction to: Innovative Computing Vol 1 - Emerging Topics in Artificial Intelligence .....	C1
<i>Jason C. Hung, Jia-Wei Chang, and Yan Pei</i>	
<b>Author Index</b> .....	1023

**The International Workshop  
on Technique for Language  
and Literature Information Modeling  
(LIM 2023)**



# A Study on Data Mining for Type of Korean Painting Poetry

Haeyoung Park<sup>1</sup> , Younghoon An<sup>2</sup> , and Hwayoung Jeong<sup>1</sup> 

<sup>1</sup> Humanitas College, Kyung Hee University, 26, Kyungheedaero, Dongdaemun-gu, Seoul, Republic of Korea

{hy000p, hyjeong}@khu.ac.kr

<sup>2</sup> Department of Korean Language and Literature, Kyung Hee University, 26, Kyungheedaero, Dongdaemun-gu, Seoul, Republic of Korea

yhnahn@khu.ac.kr

**Abstract.** This study introduced the work of analyzing the meaning by using the computer information processing method. We extracted the painting poetry of Mukjukdo (Bamboo Paintings) in the early Joseon Dynasty from the *Hanguk Munjip Chonggan* (Korean Literary Collections in Classical Chinese) of *Hanguk Gojeon Jonghap DB* (Korea Classics DB). Through the data mining method, we divided the types by extracting and cataloging the painting poetry from the Korean classical literature. Then, the scope was narrowed down to painting poetry of Mukjukdo in the early Joseon Dynasty, and the text was analyzed in units of syllable corpus.

**Keywords:** Data mining · Painting poetry · Bamboo paintings · Extracting · Cataloguing · Type Classification

## 1 Introduction

This study introduced the work of analyzing the meaning by using the computer information processing method. We extracted the painting poetry of Mukjukdo (Bamboo Paintings) in the early Joseon Dynasty from the *Hanguk Munjip Chonggan* (Korean Literary Collections in Classical Chinese) of *Hanguk Gojeon Jonghap DB* (Korea Classics DB). This is one of the basic tasks of research on Hansi (poems in the Chinese style), which has subdivided types of Hansi according to their material. This study presents expression techniques and meanings objectively by converting them into numerical values by using a computer information processing method. Traditional research on Hansi has mainly relied on researchers' intuition, however, this digital data analysis, which is a quantitative method, can complement the traditional method.

## 2 Research Methods

### 2.1 Data

The basic data for this study is *Hanguk Munjip Chonggan*; the Database of Hanguk Gojeon Beonyeokwon (Institute for the Translation of Korean Classics). It is called

“Hanguk Gojeon Jonghap DB”. Hanguk Gojeon Beonyeokwon organizes and translates Korean classics into this database and then digitally converts them and discloses them to the public. This data is provided in XML format in the form of openAPI [1]. This data is categorized by author, book name, style, title, original text, year of publication, etc. Based on this, the painting data from the early Joseon Dynasty were extracted and analyzed.

## 2.2 Data Mining Techniques

Data mining is a technique of extracting useful information by analyzing statistical patterns, rules, and relationships in large amounts of data [2]. Today, it is used in various fields such as computerization, statistics, and management. We also intend to apply this technology to the study of Korean classical literature. Currently, 142 anthologies from the early Joseon Dynasty are included in the *Hanguk Munjip Chonggan*. The primary data mining is to select the painting poetry based on this database and to classify them by their type. The secondary mining is to analyze the text by narrowing the scope [3] to painting poetry of Mukjukdo in the early Joseon Dynasty.

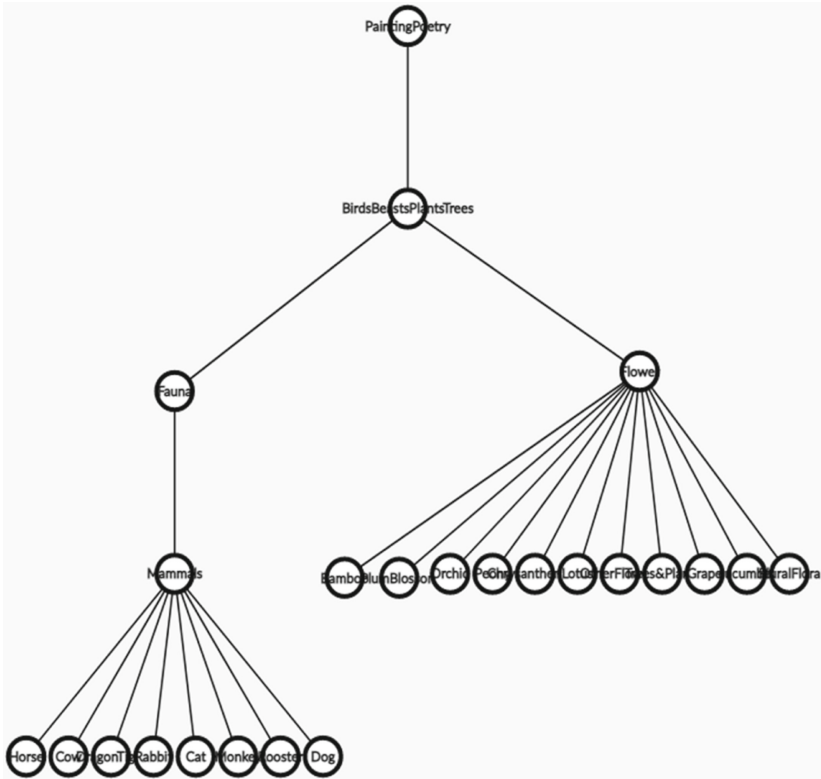
## 3 Result

### 3.1 Extracting and Cataloguing

We chose the poetry from the style category at Hanguk Gojeon Jonghap DB, and selected poems with characters in the title, such as ‘do (picture)’, ‘hwa (painting)’, ‘je (mention)’, ‘muk (ink)’, ‘hoe (drawing)’, ‘sa (drawing)’, ‘byeong (folding screen)’, ‘jok (hanging scroll)’, ‘cheop (album)’, and ‘chuk (scroll)’. After that, we compare the title and the content of each poem to determine whether the poem was painting poetry or not. Through this, it was found that there are a total of 842 poetry poems currently included in *Hanguk Munjip Chonggan*.

### 3.2 Type Classification

Based on the previously extracted lists, the painting poetry of the early Joseon Dynasty was classified by type based on the subject matter of the painting [4]. This classification is possible because the subject matter of the picture is presented as the name of each object before and after ‘do’ or ‘hwa’, which usually means a picture. For example, the titles of most painting poems of Mukjukdo appear in expressions such as “Mukjukdo” or “Jejuk (reciting a bamboo)”. The figure below is the result of categorizing Mukjukdo painting poems using the TOPIC MAP based on the list of painting poems in the early Joseon Dynasty and the types of paintings (Fig. 1).



**Fig. 1.** A Type classification of Mukjukdo painting poetry using TOPIC MAP

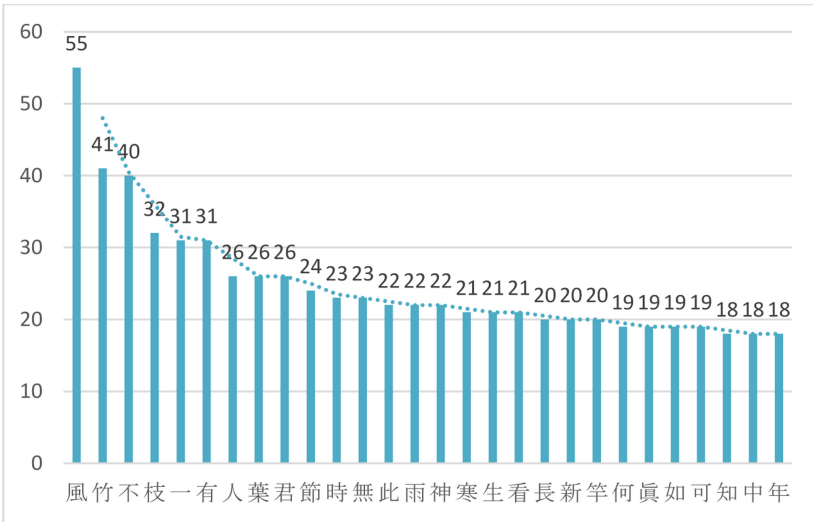
### 3.3 Syllable Unit Segments and Statistics

Next, we analyzed the text data of the painting poetry of Mukjukdo in the early Joseon Dynasty. In order to interpret the poem [5], it is convenient to divide it into morphemes, the smallest unit of meaning. Since Hansi is composed of Chinese characters, it can be divided into morphemes up to one syllable unit. In addition, the units of these syllables are basically combined into five and seven words to form a row. Therefore, we set one row as the basic unit. Among the original text of painting poetry, o'eon (five-character) was divided from 1 to 5, and chil'eon (seven-character) was divided from 1 to 7 syllables. In succession, per line, o'eon generated five corpus of one syllable, two syllables, three syllables, four syllables, and five syllables, and chil'eon generated seven corpus of one syllable, two syllables, three syllables, four syllables, five syllables, six syllables, and seven syllables (Table 1).

**Table 1.** The corpus unit of o'eon and chil'eon poem

If 1 line of o'eon poem is assumed to "1 2 3 4 5", the corpus unit	If 1 line of chil'eon poem is assumed to "1 2 3 4 5 6 7", the corpus unit
1 syllable corpus(5): 1/2/3/4/5	1 syllable corpus (7): 1/2/3/4/5/6/7
2 syllable corpus(4): 12/23/34/45	2 syllable corpus (6): 12/23/34/45/56/67
3 syllable corpus(3): 123/234/345	3 syllable corpus (5): 123/234/345/456/567
4 syllable corpus(2): 1234/2345	4 syllable corpus (4): 1234/2345/3456/4567
5 syllable corpus(1): 12345	5 syllable corpus (3): 12345/23456/34567

The following is the result of dividing the original data of Mukjukdo painting poetry by syllables and calculating the frequency. This is the figure measured using the JAVA program to see how many syllable corpus units are repeated in the actual painting poetry text data. The numerical value of 1–5 syllable corpus is the sum of the data for each syllable of o'eon and chil'eon, and the numerical values of 6–7 syllables is for chil'eon only. In these values, only words from 1 to 3 syllables were actually valid, and all 4–7 syllables were meaningless because the word combination was under twice (Figs. 2, 3 and 4).



**Fig. 2.** The frequency of one syllable corpus in Mukjukdo painting poetry

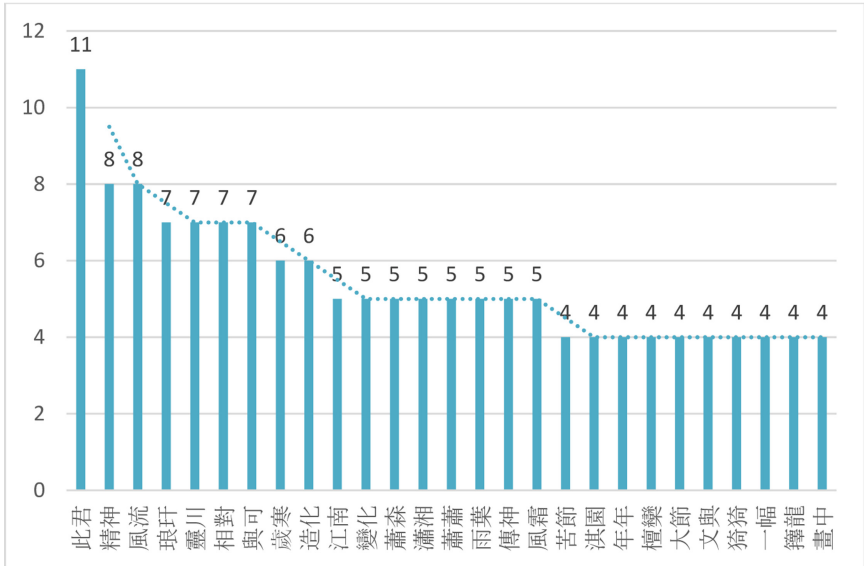


Fig. 3. The frequency of two syllable corpus in Mukjukdo painting poetry

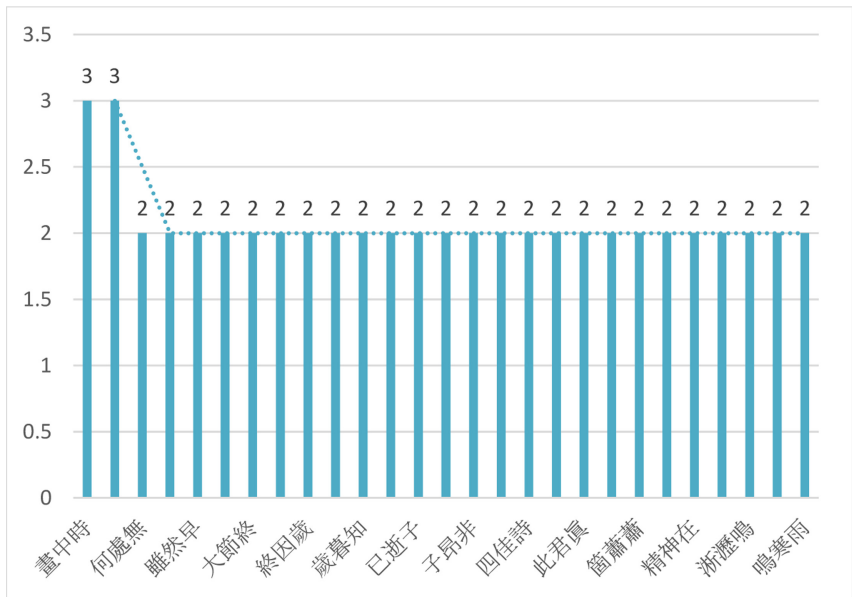


Fig. 4. The frequency of three syllable corpus in Mukjukdo painting poetry

### 3.4 Analysis of the Meaning of Corpus

The most commonly used one-syllable word in painting poetry of Mukjukdo is “poong (wind: 55 times)”. “poong” is a combination of words in the order of “poong ryu (a taste for the arts: 8 times)”, “poong sang (wind and frost: 5 times)”, “Chun poong(the spring breeze: 3 times)” etc. in 2 syllables. Poong ryu is presented when referring to the wonderful scenery of bamboo. Pung sang is wind and frost, which gives bamboo trials and tribulations. Chun poong appears when describing a situation in which bamboo shoots sprout.

Next is “juk (a bamboo: 41 times)”, but “juk” refers to bamboo itself, so it is not a discriminating result. And “bul (no: 40 times)”, which belongs to an adverb that represents negativity in Chinese characters, is not very meaningful in single syllables, but further emphasizes its meaning when combined with other words. And “ji (Branch: 32 times)”, which is connected in two syllables: “poong ji (a branch swaying in the wind: 2 times)”, “so ji(a slender branch: 2 times). It excludes the numerical value that does not form a special meaning, such as juk, bul, ji.

The next significant figure is “goon (a man of virtue: 26 times)”. “goon” is a combination of words in the order of “cha goon (this man of virtue: 11 times)”, “goon jin (a man of virtue is truly ~: 3 times)”, “goon dok (a man of virtue alone ~: 2 times)” etc.in 2 syllables. And “jeol” is 24 times, it means principles. “jeol” is a combination of words in the order of “go jeol(a distressed principles: 4 times)”, “no jeol (mature principles: 2 times)”, “jik jeol (a straight principles: 2 times)”, etc., in 2 syllables. The jeol represents the principles of bamboo, and the attitude of keeping faith firmly is matched with “go” and “no”, to reinforce the unchanging properties.

## 4 Conclusion

The data figures of corpus by syllable in Mukjukdo painting poetry are in line the symbolic meanings of bamboo in the early Joseon Dynasty. The symbolic meaning of Mukjukdo [6] is as follows. As an object, a bamboo was personified as the ideal existence of a god dragon and a Junzi. As an event, the virtual space of bamboo about the bamboo of Two queens and Qu Yuan was realized as the sorrowful and unworldly image.

Through this, the measurement of the frequency by dividing the original text of Mukjukdo painting poetry by syllables and setting it as a corpus composition unit is meaningful in revealing the symbolic meaning of bamboo in painting poetry. Conversely, the symbols of objects in various types of painting poetry can be demonstrated and shown with an objective indicator of language.

## References

1. Lee, B., Min, K.: A Study on Visualization of the Analysis between the Collections of Korean Literature in Korea Classic DB: Hanguk Gojeon Beonyeokwon. *Natl. Cult.* **57**, 5–32 (2021)
2. ko.wikipedia.org/wiki/Data Mining. Accessed 1 Dec 2022
3. Kim, D., Jeong, H.: A study of computational literature analysis based classification for a pairwise comparison by contents similarity in a section of Tokkijeon, ‘Fish Tribe Conference’: The Korea Contents Association. *J. Korea Contents Assoc.* **22**, 15–25 (2022)



4. Haeyoung, P.: A study on the painting poetry of the former part of the Joseon Dynasty. Kyunghee University Graduate school, pp. 19–34 (2021)
5. Lee, B.: A Study on the Construction and utilization of Korean Chinese poetry corpus: geunyeoghanmunhaghoe. *hanmunhaknonjib* **53**, 153–177 (2019)
6. Haeyoung, P.: The symbol of the painting poetry about a bamboo painting in the former part of Joseon Dynasty: Society of Yol-Sang Academ. *Yeol-sang J. Class. Stud.* **73**, 123-151 (2021)



# A Big Data Based Learning Model from Student Questionnaire

Hwa-Young Jeong<sup>(✉)</sup>

Humanitas College, Kyung Hee University, 1 Hoegi-dong, Dongdaemun-gu,  
Seoul 130-701, Republic of Korea  
hyjeong@khu.ac.kr

**Abstract.** The results of analyzing students' requirements become very important data for teachers. Because it is a very important criterion for improving learning satisfaction while increasing students' learning effects. In this paper, students' requirements are analyzed and what subjects or contents students want are investigated. The purpose of this paper is to construct a learning model that reflects the needs of students.

**Keywords:** Big Data · Word Cloud · Student's Needs · Learning Model

## 1 Introduction

Until the late 21st century, general teaching methods were mainly based on frontal teaching in classrooms. The use of digital tools has the effect of increasing learning efficiency, but the method of education insisted on traditional classroom lectures [3]. After COVID-19 pandemic, university education was changed and operated from offline to online, and many students became familiar with online education [1]. Teleconferencing platforms such as Zoom have become hugely popular with the spread of online education due to COVID-19 pandemic. This has become a catalyst for many institutions to seek traditional education methods as various educational methods that can conduct synchronous or asynchronous remote education [3]. In online education, learners' activities and learning data were analyzed and processed using a LMS (Learning Management System). The large amount of learning data extracted from the LMS platform provided basic information for both teachers and students that could help improve learning satisfaction and educational goals [1]. In addition, students' learning satisfaction is generally obtained through a post-learning survey, and most of them analyze and use it when students answer multiple-choice questions written by teachers. However, the multiple-choice questions given by the teacher are not sufficient data for the student satisfaction survey, and the results of the student's data analysis may vary depending on the teacher's intention. For this purpose, it is necessary to analyze students' learning satisfaction and their needs through big data analysis when presented as a subjective question, not a multiple-choice question, and students freely describe their opinions and submit answers to the questionnaire.

NLP (Natural Language Processing) is natural human language and communication variables such as voice, text, audio, and video, interpretations and applications [2] We call NLP as non-formal data. Text mining that is one of technique in Big data is well known a process to extract meaning words from the data and analyze them.

This paper aims to make a learning model using analysis of students' needs. For this process, students get a subjective questionnaire and they write their opinion freely without any form and select items. Their answers is used to analyze what they want to study or learn. Consequently, this research shows a frequency of their opinions that can be the students' needs from their study.

## 2 Related Works

### 2.1 Big Data

Big data is a technology that can handle unstructured data, unlike databases that only dealt with structured data. Of course, this can perform both processing of collecting and analyzing unstructured and structured data.

It also guarantees both volatility, speed, volume, diversity and integrity. In big data, data is collected in various ways, such as a web browser and a mobile web with various data formats. In the existing analysis method, different formats of unstructured and structured data could not be managed, but both big data are possible. In big data, Hadoop is cost-efficient, scalable, and enables fast and flexible parallelism. It also uses the Hadoop framework for big data analysis because it provides availability, resilient properties, security, and authentication. It is also well known for its open-source software architecture, which includes processing and storage. The part to be stored is HDFS (Hadoop Distributed File System), and the part to be processed is MapReduce. D.K. Jain et al. [4] depicted the Hadoop architecture as shown in Fig. 1.

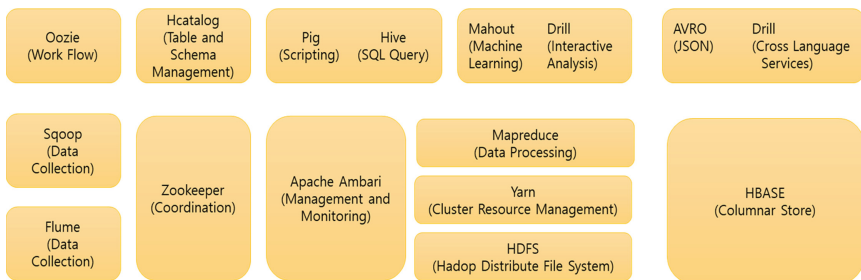


Fig. 1. Hadoop architecture by D.K. Jain et al.

Hadoop, first run by Doug Cutting and Mike Caparella in 2005, is a Java-based open-source software. Providing a distributed framework for processing and managing big data is known as Hadoop's advantage. Therefore, using Hadoop can manipulate a large amount of data. In addition, the Hadoop Distributed File System (HDFS) can store MapReduce for the purpose of the process. MapReduce is used to analyze and generate

a wide range of datasets using distributed and parallel processing methods in clusters [4].

Big data is being used in various fields. In particular, techniques for analyzing the frequency of words in unstructured data to show importance and visualize them are widely known. Figure 2 shows an example of the results analyzed by word cloud of big data for CC (Cognitive Computing) [5].



Fig. 2. A sample visualization by analyze word cloud of Big Data.

## 2.2 Learning Model

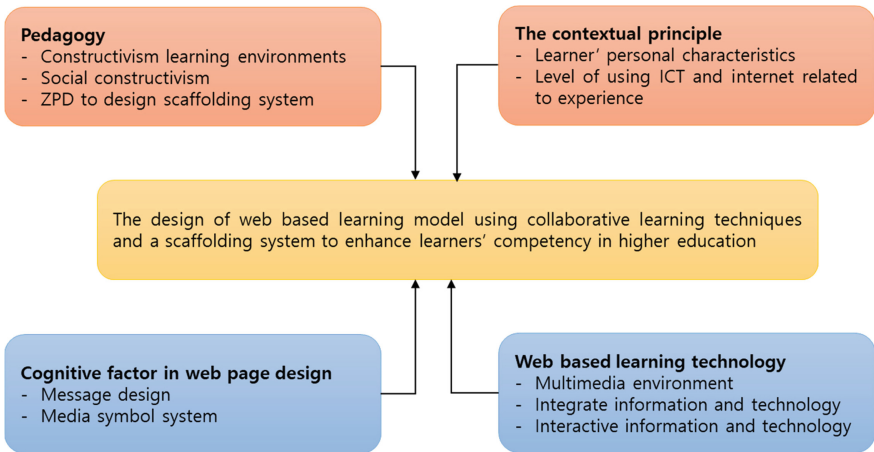


Fig. 3. The web based learning model by K. Deejrjing

The e-learning system promotes learning as teachers and students interact in a non-face-to-face manner, and learning materials and evaluations are also conducted online.

This approach can be a good solution to accommodating an exponentially growing number of students and their curriculum, but only popular large-scale online curriculums can cause problems such as student crowding, student performance, failure and departure [6]. K. DeeJring depicted e-learning model as shown in Fig. 3 [7].

### 3 Data Analysis for Students' Needs

#### 3.1 Survey of Students' Requirements for Their Class

For this study, a survey was conducted on 40 general students who did not major in IT in the liberal arts subject of K University in Seoul. The survey items are as follows.

**Question:** *Feel free to describe any topic or content you want to hear in the lecture. Or if there is a subject that you want to make a lecture in general culture, please write it with the reason.*

The comments as below represent one of the student's answers among the results of the survey.

**Answer:** *In my opinion, the topic we need in the Fourth Industrial Revolution is 'human ennui among the negative aspects of liberal utopia'. In the 4th Industrial Revolution, as each of us could access information or technology so easily, it destroyed the feudal system of the past and transformed it into a nation by the public, giving everyone a chance to politics by democracy and capitalism. As educational opportunities diversified, it began to turn into a world where everyone creates opportunities through education. But we must think about whether we will be happy. I get everything I want and come across, but I feel despondent. The reason is that in the era of the Fourth Industrial Revolution, we have it whenever we want, so we work hard to achieve something and do not feel the joy, satisfaction, and fulfillment that comes from it. Therefore, we maximize convenience and efficiency, but we may not be able to enjoy our mental satisfaction and happiness.*

#### 3.2 Frequency from the Survey

Students' survey results are unstructured data written freely. If you analyze the frequency of words using the word cloud of big data, it is shown as shown in the Fig. 4. The results of this analysis were all processed students' opinions, and they were not able to extract actually important words from the students' answers, but only the frequency was processed from the entire content.

Figure 5 shows the results except for meaningless words. What these results show is that students want subjects related to technology and are paying a lot of attention to preparing for the future. In particular, if you look at topics such as education, development, paradigm, artificial, computer, big data, metabus, and program, it can be seen that students want lectures related to this.



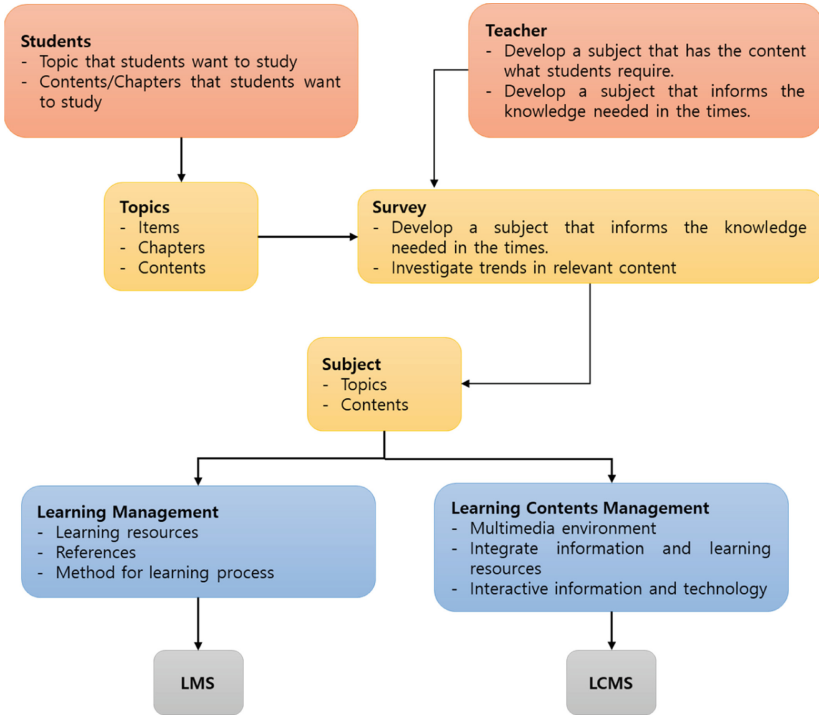
**Fig. 4.** The result of word cloud



**Fig. 5.** The result of word cloud filtering meaningless word

### 3.3 Learning Model

A learning model as shown in the Fig. 6 is constructed based on the results of student requirement analysis.



**Fig. 6.** A learning model that reflects student needs

## 4 Conclusion

This paper analyzed what students wanted to learn in class and applied it to the process of developing learning subjects. In order to know the topics what students want to study, the research was applied survey with unstructured data. In the survey result, we can see the topics that students' needs such as education, development, paradigm, artificial, computer, big data, metabus, and program. The results showed that students wanted to be able to develop more technical subjects and take them. Therefore, this paper shows the learning model that applied their requirement. Each course indicated how the teacher should reflect it and develop a learning subject after analyzing the students' requirements.

**Acknowledgement.** This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2020S1A5A2A01042497).

## References

1. Cantabella, M., et al.: Analysis of student behavior in learning management systems through a Big Data framework. *Futur. Gener. Comput. Syst.* **90**, 262–272 (2019)
2. Bhattacharjee, S., et al.: Business and government applications of text mining & Natural Language Processing (NLP) for societal benefit: Introduction to the special issue on text mining & NLP. *Decis. Support Syst.* **162**, 113867 (2022)
3. Yorkovsky, Y., Levenberg, I.: Distance learning in science and mathematics - advantages and disadvantages based on pre-service teachers' experience. *Teach. Teach. Educ.* **120**, 103883 (2022)
4. Jain, D.K., et al.: An intelligent cognitive-inspired computing with big data analytics framework for sentiment analysis and classification. *Inf. Process. Manage.* **59**, 102758 (2022)
5. Sreedevi, A.G., et al.: Application of cognitive computing in healthcare, cybersecurity, big data and IoT: a literature review. *Inf. Process. Manage.* **59**, 102888 (2022)
6. Neves, F., et al.: Chapter 8. Assisted education: Using predictive model to avoid school dropout in e-learning system. *Intelligent Systems and Learning Data Analytics in Online Education*, pp. 153–178 (2021)
7. Deejring, K.: The design of web-based learning model using collaborative learning techniques and a scaffolding system to enhance learners' competency in higher education. *Procedia. Soc. Behav. Sci.* **116**, 436–441 (2014)





# A Phonetic Investigation of Korean Monophthong Vowels by Vietnamese Female Speakers

Juhee Lee<sup>(✉)</sup>

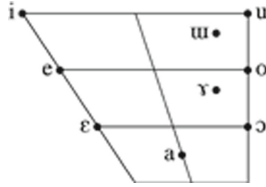
Kyung Hee University, Seoul, Republic of Korea  
juhee@khu.ac.kr

**Abstract.** In this study, we discuss the phonetic characteristics of Korean monophthong vowels produced by Vietnamese female speakers at the beginner's level. Unlike Korean female speakers, the quantitative results for /e/ and /ɛ/ show that there was a statistically significant difference in F1 values ( $p < 0.05$ ) as well as F2 values ( $p < 0.05$ ) in the production of vowels by Vietnamese learners of Korean. Therefore, the difference in height between the two vowels was contrastive by Vietnamese learners of Korean differently from native Korean female speakers. On the other hand, the back vowels /u/ and /o/ produced by female Vietnamese learners of Korean had a statistically significant difference in both F1 and F2 ( $p < 0.05$ ) values. Therefore, female Vietnamese learners' pronunciation was contrastive by the difference in tongue height (F1) as well as the position of the tongue (F2) in the production of Korean rounded back vowels /u/ and /o/, whereas these two vowels were contrastive only by the difference in tongue position (F2) in native Korean female speakers. Therefore, all pairs of monophthong vowels produced by Vietnamese female speakers formed in opposition to each other, while Korean native speakers formed a seven-vowel system due to the merger of /e/ and /ɛ/. Moreover, Vietnamese female speaker's Euclidean distance between /u/ and /o/ was shorter (88.4) than Korean female speaker's (146.5). Thus, it can be also argued that the pronunciation of the Korean vowels /u/ and /o/ produced by Vietnamese female learners of Korean are considerably different to the production of Korean rounded back vowels.

**Keywords:** formants · Korean vowels · Vietnamese · beginner level · phonetics

## 1 Introduction

In this paper, we discuss the acoustic and phonetic characteristics of Korean monophthong vowels produced by northern Vietnamese female learners of Korean at the beginner's level, and we identify the differences compared to those of Korean female speakers. Since Vietnam has a long topography from north to south, the pronunciation of vowels varies across the north, central, and south, so the experiment was conducted based on the northern dialect. The monophthong vowels of Vietnamese are nine vowels in the northern Hanoi dialect, which is considered as standard (Han 1966, Kirby 2011, Đào and Nguyễn 2018). Let us consider the following in Fig. 1:



**Fig. 1.** Vietnam monophthong vowels (Kirby 2011: 384)

In Fig. 1, Vietnamese vowels between /e/ and /ɛ/ are distinguishable, while Korean forms a seven-vowel system (/i, e, ʊ, ʌ, a, u, o/) that is based on the current spoken language due to the merger of /e/ and /ɛ/ (Yoon and Kang 2014, Lee et al. 2016, Kang and Kong 2016). Thus, we discuss the phonetic characteristics of Korean monophthong vowels produced by Vietnamese female speakers at the early stage of learning.

## 2 Methods

### 2.1 Participants

A total of 22 subjects participated in the experiment: 12 Vietnamese female learners of Korean (beginner level: less than 6 months of learning experience; average age: 21.4) from the northern part of Vietnam, including Hanoi. All of them were attending a Korean language school at the universities in Seoul. 10 Korean female speakers (average age: 21.1) who indicated their birthplaces and residences were in Seoul and Gyeonggi. All of them are students attending universities in Seoul. A predetermined honorarium was paid to all subjects.

### 2.2 Procedure

The words used in the experiment were 24 nonce words in Korean with two syllables composed of eight monophthong vowels. Due to the merging of /e/ and /ɛ/, the monophthong vowels of Seoul Korean form a seven-vowel system based on the production level (Yoon and Kang 2014, Lee et al. 2016, Kang and Kong 2016). However, we included nonce words with eight vowels to examine how Vietnamese speakers pronounced /e/ and /ɛ/, respectively, and how they differed from subjects who speak Korean as their mother tongue. The nonce words in Table 1 are in the form of “V<sub>1</sub> + CV<sub>2</sub>” so that eight monophthong vowels (V<sub>1</sub>) are the word-initial position, while the following consonants are Korean plosives /p, t, k/.

**Table 1.** Nonce words.

	_ka	_ta	_pa
i	ika	ita	ipa
e	eka	eta	epa
ɛ	ɛka	ɛta	ɛpa
uɯ	uɯka	uɯta	uɯpa
ʌ	ʌka	ʌta	ʌpa
a	aka	ata	apa
u	uka	uta	upa
o	oka	ota	opa

The experiment was conducted in a manner that the subjects repeatedly read the 24 nonce words presented in Table 1 three times, and the words were read with the carrier sentence ‘This is \_\_\_\_\_’. In this way, 72 pieces of data were collected per subject, and the total number of data obtained through each group of Korean (10 people  $\times$  24 words  $\times$  3 repetitions) and Vietnamese (12 people  $\times$  24 words  $\times$  3 repetitions) was 1,584. The recording was performed using an LG Gram laptop, which is included in the public software Praat version 6.2.23. A SONY ECM-LV1 pin microphone was attached to the upper body of the subject and connected to the laptop. The recording took place in a quiet space on the campus without noise.

For the data analysis, using Praat version 6.2.23, we checked the spectrogram of all vowels and set the section with the least variation in formant as the stable section. Since all the subjects were women, the maximum formant value was set to 5,500 (Hz), the number of formants was set to five, and the window was set to 25 (ms). The measurement of the formant value was calculated by finding the mid points of the stable section found in the collection. Through this process, an analysis was conducted on a total of 1,584 tokens. Formant values were summarized using Excel, and statistical analysis was performed using IBM’s statistical analysis program SPSS 26 to understand the statistical significance.

### 3 Results

#### 3.1 Production

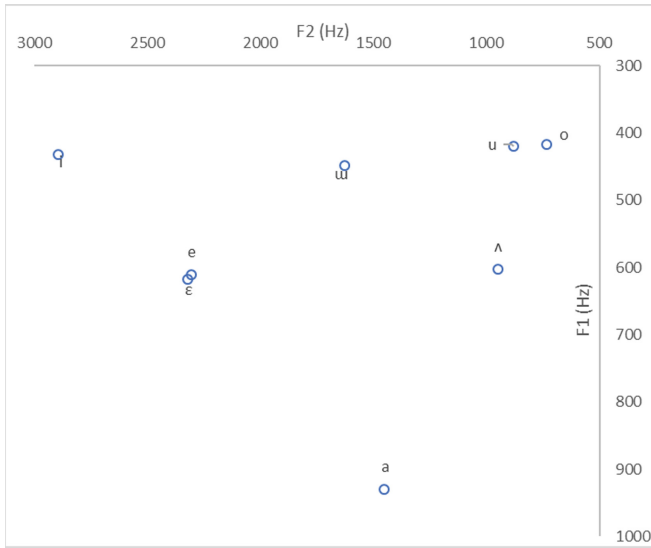
To discuss the acoustic and phonetic characteristics of Korean monophthong vowels produced by Korean female speakers, we present the formant values in Table 2. To judge the significant differences among Korean vowels, the results of repeated-measures ANOVA followed by Bonferroni post-hoc analysis are reflected.

**Table 2.** Korean female speakers estimated mean in F1/F2 (Hz)

Vowel (F1)	Estimated mean	Standard error	95 % Confidence interval	
			minimum	maximum
i	433.5	7.1	419.5	447.5
e	611.1	4.0	603.3	619.0
ɛ	618.8	4.1	610.6	627.0
u	449.2	5.5	438.3	460.1
ʌ	603.1	5.8	591.5	614.7
a	931.0	6.8	917.6	944.4
o	420.0	4.8	410.5	429.5
o	418.0	4.7	408.6	427.4
Vowel (F2)	Estimated mean	Standard errors	95 % confidence interval	
			minimum	maximum
i	2759.7	61.9	2636.8	2882.6
e	2210.9	44.0	2123.7	2298.2
ɛ	2229.6	43.1	2144.0	2315.2
u	1565.5	29.7	1506.5	1624.4
ʌ	932.3	11.1	910.3	954.3
a	1424.0	14.6	1395.0	1453.0
o	854.2	15.7	822.9	885.4
o	715.2	11.8	691.8	738.6

In Table 2, in F1 values, there is no significant difference between /e/ and /ɛ/ ( $p > 0.05$ ) and /u/ and /o/ ( $p > 0.05$ ). On the other hand, in F2 values, there is no significant difference between /e/ and /ɛ/ ( $p > 0.05$ ), but contrast is maintained between /u/ and /o/ ( $p < 0.05$ ) in the position of the tongue in F2 values. Based on the estimated mean in Table 2, we present the Korean female speakers' vowel production in Fig. 2.

In Fig. 2, the vowel diagram shows that the traditional Korean phonological category (the so-called eight-vowel system) does not match the production of Korean female speakers. However, unlike in Korean, the Vietnamese vowel system distinguishes front vowels between /e/ and /ɛ/, as shown in Fig. 1. It is unclear whether Vietnamese female speakers pronounce these two vowels differently, or simply merge them in the same way as Korean speakers. Thus, we investigated the acoustic and phonetic characteristics of Korean monophthong vowels produced by Vietnamese female learners of Korean at the beginner's level. The results of repeated-measures ANOVA followed by Bonferroni post-hoc analysis are also reflected in this analysis.



**Fig. 2.** Korean female speakers Korean vowel production (Hz)

**Table 3.** Vietnam female speakers (beginner) F1/F2 values (Hz)

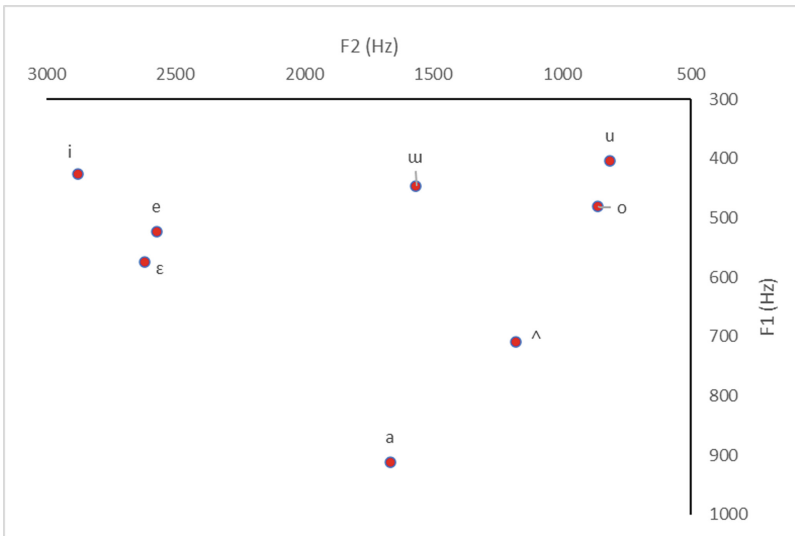
Vowel (F1)	Estimated mean	Standard errors	95% confidence interval	
			minimum	maximum
i	425.6	5.8	414.0	437.1
e	524.4	7.9	508.7	540.0
ɛ	573.8	9.1	555.8	591.8
ʊ	446.5	6.2	434.3	458.7
ʌ	708.7	11.6	685.7	731.7
a	912.7	9.5	893.9	931.5
u	404.5	5.8	393.0	416.0
o	480.4	8.5	463.6	497.2
Vowel (F2)	Estimated mean	Standard errors	95% confidence interval	
			minimum	maximum
i	2875.0	14.5	2846.1	2903.8
e	2569.7	18.6	2532.7	2606.6
ɛ	2618.3	18.4	2581.8	2654.8

(continued)

**Table 3.** (continued)

Vowel (F2)	Estimated mean	Standard errors	95% confidence interval	
			minimum	maximum
u	1565.6	15.6	1534.6	1596.6
ʌ	1179.0	9.501	1160.2	1197.9
a	1665.3	12.0	1641.4	1689.1
u	814.1	9.7	794.9	833.3
o	859.5	8.4	842.8	876.2

Unlike Korean female vowel production, the quantitative analysis of F1 and F2 values in Table 3 shows that all eight vowels are categorically significant in F1 and F2 values ( $p < 0.05$ ). Based on the estimated mean in Table 3, we present the diagram for Vietnamese female learners' vowel production.



**Fig. 3.** Korean vowel production of Vietnamese female learners' (beginner) (Hz)

In Fig. 3, the distance between each vowel is statistically contrastive among Vietnamese female speakers. Therefore, we argue that the native (Vietnamese) phonological category, as shown in Fig. 1, is reflected in the production of Korean vowels.

### 3.2 Euclidean Distance

Based on the analysis of the F1 and F2 values, the distance between the two points corresponding to each vowel was obtained by applying the following formula to mathematically determine the distance between the vowels. Euclidean distance is a technique for calculating the distance between two points, and when two points have coordinates of  $(P_1, P_2, P_3 \dots P_n)$  and  $(Q_1, Q_2, Q_3 \dots Q_n)$ , the distance between two points is expressed by the Euclidean distance formula as follows.

$$d(p, q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2}$$

The distance between two points (two paired vowels) can be calculated using the formula to measure the two distances. The Euclidean distance between each vowel is shown as in Table 4.

**Table 4.** KF/ VF (beginner) speakers' Euclidean distance for Korean vowels

KF	$\Delta F1$	$\Delta F2$	ED	VF	$\Delta F1$	$\Delta F2$	ED
i-e	170.5	589.7	613.9	i-e	98.4	305.0	320.5
<b>e-ε</b>	<b>7.7</b>	<b>18.2</b>	<b>19.8</b>	<b>e-ε</b>	<b>49.8</b>	<b>48.3</b>	<b>69.4</b>
ε-a	338.9	872.1	926.0	ε-a	338.9	953.0	1011.5
i-ɯ	15.7	1267.7	1267.8	i-ɯ	20.9	1309.4	1309.6
ɯ-u	29.2	746.7	747.3	ɯ-u	42.0	751.5	752.7
ɯ-o	31.2	893.2	893.7	ɯ-o	33.9	706.1	999.2
ɯ-ʌ	153.9	676.4	693.7	ɯ-ʌ	262.2	386.6	467.1
ɯ-a	175.9	481.8	512.9	ɯ-a	466.2	99.7	457.2
<b>u-o</b>	<b>2.0</b>	<b>146.5</b>	<b>146.5</b>	<b>u-o</b>	<b>75.9</b>	<b>45.4</b>	<b>88.4</b>
o-ʌ	185.1	216.8	285.1	o-ʌ	228.3	319.5	392.7
ʌ-a	327.9	500.5	598.3	ʌ-a	204.0	486.3	527.4

In Korean speaker's production, the Euclidean distance between the vowels /e/ and /ε/ was shorter (/e: ε/, 19.8) than Vietnamese speaker's production. The difference in tongue height (F1) between these vowels was not significant for Korean female speakers ( $p > 0.05$ ), while these two vowel categories were statistically contrastive in Vietnamese learners' production ( $p < 0.05$ ). For the case of the rounded back vowel /u/ and /o/, Korean production of these vowels' distance is short in F1 (/u: o/, 2.0) but the distance was maintained in F2 (/u: o/, 146.5). Thus, the Euclidean distance between /u/ and /o/ is 146.5. However, the Euclidean distance for the /u: o/ in Vietnamese learners' production was shorter than the Korean realization of /u: o/.

## 4 Discussion

The results of quantitative formants for /e/ and /ɛ/ vowels show that there was a statistically significant difference in F1 as well as in F2 values ( $p < 0.05$ ) in the production of Vietnamese learners of Korean. Therefore, the difference in height between the two vowels was contrastive by Vietnamese speakers differently from native Korean female speakers. The back vowels /u/ and /o/ produced by female Vietnamese speakers had a statistically significant difference in F1 value ( $p < 0.05$ ), but there was no statistically significant difference in F2 value ( $p > 0.05$ ). Therefore, female Vietnamese speakers' production was contrastive by the difference in tongue height (F1), whereas the distance was maintained only by the difference in tongue position (F2) in native Korean speakers. Therefore, all pairs of monophthong vowels produced by Vietnamese female speakers formed an opposition to each other. Interestingly, Vietnamese female speakers' Euclidean distance between /u/ and /o/ was shorter (88.4) than Korean female speakers' (146.5). Thus, statistical formant analysis concluded that the production of the Korean vowels /u/ and /o/ produced by Vietnamese female learners have considerable phonetic differences compared Korean speakers, as shown in Fig. 4.

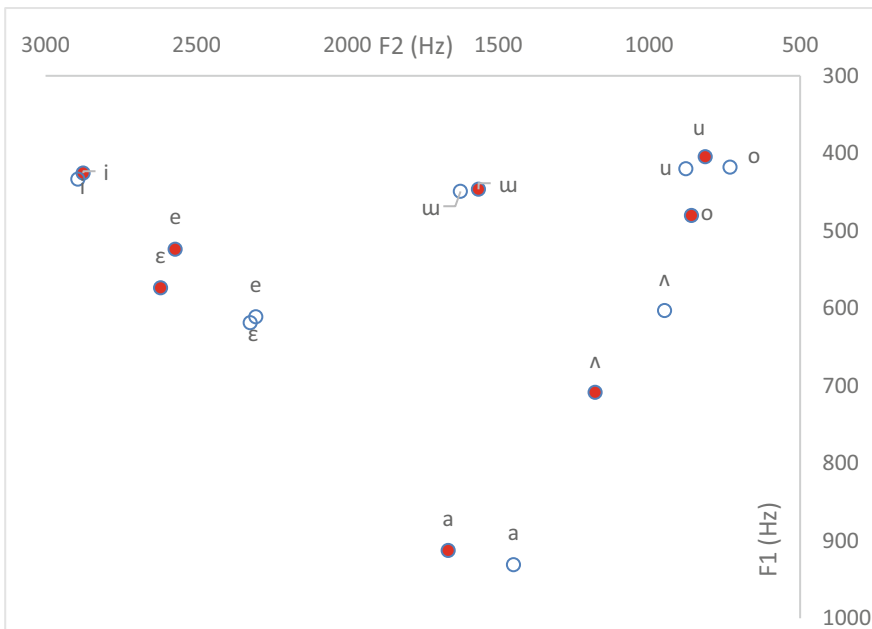


Fig. 4. Korean vowels produced by KF (white dot) and VF (beginner)

## References

1. Đào, Đích M.ục, Nguyễn, Anh-Thư T.: L1 Korean vocalic transfer in adult L2 Korean learners' production of Vietnamese monophthong vowels. *Asian-Pacific J. Second Foreign Lang. Educ.* 3(1), 1–20 (2018). <https://doi.org/10.1186/s40862-018-0055-1>



2. Han, M.S.: Vietnamese Vowels (Studies in the Phonology of Asian Languages 4). University of Southern California, Acoustic Phonetics Research Laboratory, Los Angeles, CA (1966)
3. Kang, J., Kong, E.J.: Static and dynamic spectral properties of the monophthong vowels in Seoul Korean: Implication on sound change. *Phonet. Speech Sci.* **8**(4), 39–47 (2016). <https://doi.org/10.13064/KSSS.2016.8.4.039>
4. James, P.K.: Illustration of the IPA: Vietnamese (Hanoi Vietnamese). *J. Int. Phonetic Assoc.* **41**(3), 381–392 (2011)
5. Lee, J., Yoon, K., Byun, K.: A study of vowel shifts in Seoul Korean. *J. Stud. Lang.* **31**(4), 979–998 (2016)
6. Park, S.-G., Kim, J.-Y.: A study on the analysis of the L1–L2 similarity between Korean and Vietnamese Monophthongs. *Korean J. Linguist.* **42**(4), 691–716 (2017)
7. Thompson, L.C.: *A Vietnamese Grammar*. University of Washington, Seattle, WA (1965)
8. Yoon, T.-J., Kang, Y.: Monophthong analysis on a large-scale speech corpus of read-style Korean. *Phonet. Speech Sci.* **6**(3), 139–145 (2014)



# Korean Causal Connective Expressions in a Cross-Linguistic and Cultural Perspective

Sujeong Choi<sup>1</sup> and Sinhye Nam<sup>2</sup>

<sup>1</sup> KDI School of Public Policy and Management, 263 Namsejong-ro, Sejong-si 30149,  
Republic of Korea

schoi@kdis.ac.kr

<sup>2</sup> Kyung Hee University, 26, Kyungheedaero, Dongdaemun-gu, Seoul 02447,  
Republic of Korea

namsh@khu.ac.kr

**Abstract.** This study aims to explore the reason why there are various causal connective expressions in Korean in comparison with English from a cultural perspective. In this study, Korean causal connective expressions are linguistically analyzed from the cross-linguistic perspective with English expressions in the Korean-English Parallel Corpora. Then, the differences between Korean causal connective expressions and the corresponding English expressions are interpreted from a cultural point of view. The findings of this study are as follows. First, there are various causal connective expressions implying a negative meaning in Korean, and negative nuances can be indirectly delivered through grammatical expressions. Second, in Korean the causal connective expression that includes the conjecture meaning is much more frequently used compared to English. The phenomenon of using a lot of guessing expressions in Korean can be interpreted as Korean speakers exhibiting a tendency to express their thoughts or opinions mildly and indirectly rather than strongly and clearly, and this is also related to showing politeness toward the listener. Third, there is the causal connective expression implying that it is one of various reasons in Korean. It implies that Korean speakers intend to avoid conclusive expressions by emphasizing that it is one of several reasons rather than concluding a single reason. This can also be viewed as a way to keep from expressing one's intentions too strongly and avoid causing the other person to lose face.

**Keywords:** Cross-linguistics · Causal Connective Expressions · Interpretive Ethno-grammar

## 1 Introduction

Communication competence encompasses not only linguistic and grammatical competence, but also cultural competence. In order to communicate properly in a particular language society, speakers must not only understand basic linguistic structures such as vocabulary, phonemes, and grammar, but also know how to speak appropriately in a given social and cultural context (Saville Troike 2003). This is because a language is

shaped by the culture. Depending on the culture, there may be differences in the way of expressing the same subject in each language.

For example, the Hanunoo language of the Philippines has dozens of words for different kinds of *rice* (Conklin 1957) and Russian has several different everyday words for different kinds of *friends* rather than one basic everyday word like the term *friend* used in English (Wierzbicka 1997), and this phenomenon is not limited to words. Compared to other languages, in Korean, there are a lot of connective expressions that express a cause or reason. This phenomenon can be analyzed along the same lines as the phenomenon in which *rice* in the Hanunoo and *friends* in Russian appear significantly more than in other languages. In other words, it can be said that the expressions that indicate the cause or reason appear very diverse due to the cultural characteristics of Koreans' tendencies or communication.

This study aims to explore the reason why there are various causal connective expressions in Korean from a cultural point of view. For this, first, Korean causal connective expressions are linguistically analyzed from the cross-linguistic perspective with English expressions. Then, the differences between Korean causal connective expressions and the corresponding English expressions are interpreted from a cultural point of view. If speakers are aware of the reasons for the variety of expressions representing the cause or reason in Korean from a linguistic cultural perspective, it will be possible to help foreign learners avoid misuse of language in a cultural context and achieve successful cross-cultural communication.

## 2 Korean Causal Connective Expressions

Choi (2022) analyzed the grammar items from 10 classes of Korean textbooks and 3 classes of grammar books to categorize them based on their meanings. As a result, a total of 68 semantic categories of grammar items were presented and a list of synonymous grammar items was organized for each category. It was found that the meaning category which has the greatest number of synonymous grammar items was the [cause] category. According to Choi (2022), there are 24 synonymous grammar items expressing the cause or reason in Korean: *-a/eoseo*, *-(eu)nikka*, *-(eu)meuro*, *-gi ttaemune*, *-neurago*, *-neun barame*, *-gilrae*, *-gie*, *-deoni*, *-at/eotdeoni*, *-(eu)n/neun mankeum*, *-a/eo gajigo*, and so on. These all have a common meaning, which is the cause or reason for the following clause in the sentence. However, they differ in terms of syntactic conditions, contextual formality, or semantic features.

For example, in Korean textbooks, the following causal grammar items are provided with descriptions as they have the additional semantic features compared to the basic/neutral causal grammar items '*-a/eoseo*, *-(eu)nikka*, *-(eu)meuro*, *-gi ttaemune*'.

- (1) *Ast-(eun) tase*, *Vst-neun tase*: It is a negative expression of cause or reason.
- (2) *Ast/Vst-(eun) nameoji*: It is used when some action or situation in the first clause becomes worse, and it leads to a negative result in the following clause.
- (3) *Vst-neurago*: It indicates a cause, reason, or purpose getting a negative result.
- (4) *Vst-neun barame*: It indicates a cause or reason. The preceding situation negatively affects the following action.

- (5) *Vst-neun tonge*: It indicates a cause or reason that caused the negative situation or result of the following clause.
- (6) *Ast/Vst-a/eoseo geureonji*: It is used when the previous action or situation seems to be the cause and reason for what follows but cannot be determined for sure.
- (7) *Ast/Vst-go haeseo*: This is used to show that the preceding clause is one of many reasons for the content in the following clause.

As underlined above, the causal connective expressions in (1) to (7) have additional semantic features: the grammar items in (1) to (5) contain [+negative] features, the grammar item in (6) has [+uncertainty] features, and the grammar item in (7) implies [+plurality]. Like this, Korean causal connective expressions are not only able to express the cause or reason for the following clause, but they can also imply additional semantic features or nuances.

As such, it seems there are numerous grammar items expressing cause or reason in Korean. At this point, the question of whether this phenomenon exists in other languages as well, or whether this is a characteristic unique to Korean can be raised.

In order to explore the above question in this study, seven causal connective expressions that contain additional semantic features are compared with the corresponding English expressions. In addition, through these analysis results, we attempt to interpret the reason why there are many causal connective expressions in Korean from a cultural perspective.

### 3 Data and Methodology

The data of this study was collected by the Korean-English parallel corpora developed by *AI-Hub*. *AI Hub* is an AI integration platform operated by Korea's Ministry of Science and ICT and the National Information Society Agency. The Korean-English parallel corpora were released as a part of the data construction project for artificial intelligence learning. The Korean-English parallel corpora consist of three styles of corpus: literary, colloquial, and colloquial conversation, covering 1,600,000 sentences.

In this study, 200,000 sentences (2,658,545 words) from the news article corpus in the literary style and 100,000 sentences (779,541 words) from the conversation corpus in the colloquial style were analyzed as follows (Table 1).

**Table 1.** Information of Korean-English Parallel Corpora

No.	Style	Context	The Number of Sentences/Words	
1	Literary	News Articles	200,000 sentences (2,658,545 words)	300,000 sentences (3,438,086 words)
2	Colloquial (Conversation)	Conversation in a meeting, shopping, school, restaurant, etc	100,000 sentences (779,541 words)	

A total of 300,000 sentences (3,438,086 words) in the Korean-English parallel corpora were examined in order to explore the corresponding English expressions with Korean reason/causal connective expressions.

	A	B	C	D	E
	Korean	English	Korean Sentences	English Sentences	
1	<i>Vst-neun barame</i>	and	반이 인도네시아인 부랑기에 날 '안도'지만, '기세' 앞날로 후기를 포함하는 비행기의 이륙이 지난 2일 아침에야 완전히 도착했다.	when the trip from Korean Embassy in Indonesia, due to issues with the aircraft, he had to go back and she was able to arrive at the site in the morning of 2nd after 2 days.	국제, 아시아어
309	<i>Vst-neun barame</i>	because	북한 선수들이 백니스와 골키퍼가 골을 제대로 차단하지 못하는 데다가 골키퍼가 있던 전방에서 코스타안테 페르네스에게 패스를 내주고, 후반 33분엔 루카 모드리치(33세)의 슛으로 골문을 열어주었다.	At the 8th minute of the second half, they gave out the first goal to Croatia's Ante Rebic (25, Frankfurt) because the goalkeeper could not handle the ball properly during the back pass, and they allowed the mid-range goal to the Luka Modric (33, Real Madrid) at the 33th minute of the second half.	스포츠, 축구, 한국프로축구
310	<i>Vst-neun barame</i>	and	말씀하신 대로 41분을 살피다 못한 방수 골문 후반 41분 상대 수석 골키퍼인 공이 들어온 데다가 골문으로 향하는 비행이 골을 넣어주고 있었다.	41 minutes into the second half of the game, Korea, which was not able to make use of the several opportunities to flip the game, had the ball that I'm Seon-joo was trying to pull away from the opponent's shooting hit his head and head into the goalposts, and gave away the final goal.	스포츠, 축구, 해외축구
311	<i>Vst-neun barame</i>	so	홍콩 민주화운동 투쟁당 '민주'의 임원들은 오후 6시에 회의를 취소하고, 회의를 다시 취소할 예정이지만 이번에도 한 차례의 회의가 단적으로 불려와 합의하는 비행이 안전 채 마무리되었다.	As the morning meeting was canceled, Members of the Democratic Party of Korea, including, Hong Joo, the chairman of the subcommittee, tried to resume the meeting again in the afternoon. But, this time again, 6 members of the Liberty Korea Party (a group and protest) for the passing of the bill was founded.	정치, 국회, 정당
312	<i>Vst-neun barame</i>	due to	김장보교 측은 그로부 집행위원장/홍남철위원장/전남지회장이 '연세사건은 (연세사건이 아니라) 김대중 군인이 일으킨 반한 '연세교' 발원하는 비행이 소인이 들어갔다.	During the report, Executive Committee Head Go Hyeo-yu/Vietnam Veterans Association Jeonnam Branch Head, stated that the uproar occurred due to the statement: "The 'yeonun incident was a revolt caused by leftist soldiers (and not the citizens of Yeosu)".	지역, 전남
313	<i>Ast-(eun) tase, Vst-neun tase</i>	due to	제품이 제대로 검수할 하지 않은 탓에 하자 제품이 출고되었는데 고객님이 원하시므로 환불 처리하도록 하겠습니다.	The defected good was delivered due to our lack of inspection, so if you want, we will refund it.	여행/쇼핑
314	<i>Ast-(eun) tase, Vst-neun tase</i>	because	이산 품의약품이 금지된 탓에 해외로 수출 불허다.	Because domestic medicine is banned, it has turned to foreign countries.	사회, 의료, 건강
317	<i>Ast-(eun) tase, Vst-neun tase</i>	because	이 공업 지원이 끊긴 탓에 수출이 어렵게 되었다.	Labour assistance has intensified because such support has been cut off.	사회, 노동, 복지
318	<i>Ast-(eun) tase, Vst-neun tase</i>	because	이러한 특성을 수용한 탓에 수출도 어려움이 크다.	Because it has already accepted the quota system, exportation is also very difficult.	경제, 산업, 기업
319	<i>Ast-(eun) tase, Vst-neun tase</i>	because of	제품부터 크게 지출 탓에 공사가 곳곳에 있었다.	The vacant lots were everywhere because of the large construction from the beginning.	경제, 산업, 기업
320	<i>Ast-(eun) tase, Vst-neun tase</i>	due to	스타트업이 도입된 탓에 높은 회전율을 보인다.는 점도 특징이다.	Another characteristic is that it has high turnover due to dense startups.	경제, 무역, 방법
320	<i>Ast-(eun) tase, Vst-neun tase</i>		가시불이 현금으로 받거나 현금 계좌 차액이 확인이 되어		

Fig. 1. An Example of Data Analysis

As shown in Fig. 1, 479 sentences using the seven Korean causal connective expressions were searched from the Korean-English parallel corpora and the corresponding English expressions for each Korean causal connective expression were extracted from the English sentences.

After examining all the corresponding English expressions, they were categorized into two groups: reason/causal expressions and non-reason/causal expressions. If the English expressions are used to indicate the reason or cause, they were included in the reason/causal expressions. In contrast, if the English expressions are not used to indicate the reason or cause, they were included in the non-reason/causal expressions.

### 4 Korean Causal Connective Expressions in a Cross-linguistic Perspective

Table 2 shows the frequency and rate of Korean causal connective expressions and corresponding English expressions in the Korean-English parallel corpora.

Table 2. Korean Causal Connective Expressions and Corresponding English Expressions

No	Korean	N	English	N	%
1	<i>Ast-(eun) tase, Vst-neun tase</i>	145	Reason/causal expressions	because (of)	39.31
				due to	22.07
				as	17.24

(continued)

**Table 2.** (continued)

No	Korean	N	English	N	%	
				since	14	9.66
				so	2	1.38
				by	1	0.69
				for	1	0.69
				therefore	1	0.69
			Non-reason/causal expressions	cause	4	2.76
				after	2	1.38
				and	1	0.69
				as a result	1	0.69
				result in	1	0.69
				lead	1	0.69
				so ... that	1	0.69
				while	1	0.69
2	<i>Ast/Vst-(eun) nameoji</i>	24	Reason/causal expressions	because	5	20.83
				as	3	12.50
				for	2	8.33
			Non-reason/causal expressions	and	5	20.83
				so... That	3	12.50
				∅	5	20.83
by ~ ing	1	4.17				
3	<i>Vst-neurago</i>	13	Reason/causal expressions	because	4	30.77
				due to	1	7.69
				from	1	7.69
				as	1	7.69
				with	1	7.69
				~ ing	2	15.38
			Non-reason/causal expressions	and	1	7.69
				bring	1	7.69
				on	1	7.69
4	<i>Vst-neun baram</i>	153	Reason/causal expressions	because	61	39.87
				so	24	15.69

(continued)

**Table 2.** (continued)

No	Korean	N	English	N	%	
				as	14	9.15
				since	11	7.19
				due to	6	3.92
				from	1	0.65
			Non-reason/causal expressions	and	12	7.84
				∅	9	5.88
				cause	6	3.92
				, which	2	1.31
				after	2	1.31
				when	2	1.31
				drive	1	0.65
				disrupt	1	0.65
				get	1	0.65
5	<i>Vst-neun tonge</i>	4	Reason/causal expressions	since	1	25
				due to	1	25
			Non-reason/causal expressions	and	1	25
				confuse	1	25
6	<i>Ast/Vst-a/eoseo geureonji</i>	124	Reason/causal expressions	because	31	24.80
				maybe (it's) because	28	22.40
				so	13	10.40
				probably because	11	8.80
				perhaps because	7	5.60
				since	7	5.60
				due to	3	2.40
				maybe that's why	2	1.60
				maybe the reason	2	1.60
				not sure if it is because	2	1.60
				with	2	1.60
				guess it's because	1	0.80
				perhaps... so	1	0.80
perhaps due to	1	0.80				

(continued)

**Table 2.** (continued)

No	Korean	N	English	N	%	
				whether it's from	1	0.80
			Non-reason/causal expressions	∅	8	6.40
				and	2	1.60
				as to whether	1	0.80
				no wonder	1	0.80
				so... That	1	0.80
7	Ast/Vst-go haeseo	9	Reason/causal expressions	because	2	22.22
				so	2	22.22
				since	1	11.11
				due to	1	11.11
				as	1	11.11
			Non-reason/causal expressions	∅	2	22.22

As described in Sect. 2, the Korean causal connective expressions (1) to (5) connote [+negative] meaning, the expression in (6) has [+uncertainty] meaning, and the expression in (7) contains [+plurality] meaning additionally. We analyzed whether these additional semantic qualities appear in the corresponding expressions in English, and the results are examined by each additional semantic quality below.

#### 4.1 [+Negative] Feature in Causal Connective Expressions

The English expressions that correspond with the Korean causal connective expressions *Ast-(eun) tase*, *Vst-neun tase*, *Ast/Vst-(eun) nameoji*, *Vst-neurago*, *Vst-neun baram*, *Vst-neun tonge*, which imply the [+negative] feature, were analyzed. As a result, in English there were no causal expressions that have the [+negative] feature noticed in Korean expressions, and only basic/neutral causal expressions such as *because*, *due to*, *as*, *since* were used as the corresponding expressions.

In English, it was found that there was a large tendency to use direct negative vocabulary when trying to connote a negative meaning rather than implying such negative meaning through grammar items. In that, in English, words with negative meanings are used directly when expressing negative intentions, whereas in Korean, grammatical expressions implying the negative meaning are used somewhat indirectly. It means in Korean, even if the speaker does not use direct vocabulary to express negative intentions, negative nuances can be indirectly delivered through grammatical expressions. The phenomenon of speakers indirectly expressing their intentions in Korean can also be found when they express their thoughts or opinions with conjecture expressions such as *-(eu)n/neun/(eu)l geot gat-* or when they express their plans with *-(eu)lkka ha-*, *-(eu)lkka sip-*, and so on.



## 4.2 [+Uncertainty] Feature in Causal Connective Expression

The English expressions that correspond with the Korean causal connective expression *Ast/Vst-a/eoseo geureonji*, which implies the [+uncertainty] feature, were analyzed. In English, conjecture expressions such as *maybe*, *probably*, and *perhaps* were not included in as many as 50% of the corresponding English expressions with *Ast/Vst-a/eoseo geureonji* in the reason/causal expression category. In that, cases involving speculation in the causal expressions appeared much more in Korean.

As mentioned above, this coincides with the phenomenon of using a lot of guessing expressions in Korean. In Korean, when expressing a cause or reason, there is a strong intention to express it mildly rather than strongly and clearly, and this is also related to showing politeness toward the listener.

## 4.3 [+Plurality] Feature in Causal Connective Expression

The English expressions that correspond with the Korean causal connective expression *Ast/Vst-go haeseo*, which implies the [+plurality] of the reason, were analyzed. In Korean, this expression shows that the preceding clause is one of many reasons for the content in the following clause. However, in English, this connotative meaning was not expressed. Unlike English, the causal expression implying that it is one of various reasons frequently appears in Korean because it implies that the speaker intends to avoid conclusive expressions by emphasizing that it is one of several reasons rather than concluding a single reason. This can also be said to be a way to keep from expressing one's intentions too strongly. This is used in a context where the speaker wants to avoid definitive reasons: when the speaker has to refuse the other's request or suggestion, the speaker does not want to cause the other person to lose face by giving the connotative meaning that there are many reasons for being unable to accept the request or suggestion, or when the speaker does not want to explicitly express their actual reasons.

## 4.4 Additional Findings

Through the analysis, a few findings can be discussed further.

First, the corresponding English expressions with Korean causal expressions were mostly one of four expressions such as *because*, *since*, *due to*, and *as*, and it shows that in English various grammar items expressing the cause or reason are not used as much as in Korean.

Second, even though there are some different causal expressions in English, their difference is the degree of formality rather than semantic features. However, in Korean, there are various causal expressions, and their difference can be explained in many aspects: formality, syntactic condition, and semantic features.

Third, there are cases in which Korean causal expressions do not correspond with the non-reason/causal expressions in English. This indicates that causal expressions are used more frequently in Korean, and sometimes the causal expression is used for other intentions such as supporting their thoughts rather than expressing the actual cause or reason in a logical context.

## 5 Conclusion

It is generally recognized that languages differ in the amount – and kind – of attention given to different aspects of reality through their lexical systems: Arabic has numerous words for *sand*, Hanunoo for *rice*, and so on (Wierzbicka 2002). This study considered that it also applies to grammatical expressions and not only to the lexical system. Thus, this study focused on the phenomenon that there are various causal connective expressions in Korean in comparison with English from a cultural perspective.

As a result of the linguistic analysis from the cross-linguistic perspective between Korean causal connective expressions and corresponding English expressions, it was found that the following three aspects were characteristics unique to Korean and these were interpreted in a cultural perspective.

First, there are various causal connective expressions implying a negative meaning in Korean, whereas there are no such expressions in English. In Korean, the negative nuance can be indirectly delivered through grammatical expressions.

Second, in Korean the causal connective expression that includes the conjecture meaning is much more frequently used compared to English. As mentioned above, the phenomenon of using a lot of guessing expressions in Korean can be interpreted as Korean speakers exhibiting a tendency to express their thoughts or opinions mildly and indirectly rather than strongly and clearly, and this is also related to showing politeness toward the listener.

Third, there is the causal connective expression implying that it is one of various reasons in Korean, whereas no such expression is used in English. It implies that Korean speakers intend to avoid conclusive expressions by emphasizing that it is one of several reasons rather than concluding a single reason. This can also be said to be a way to keep from expressing one's intentions too strongly and avoid causing another person to lose face.

The results of this study from the linguistic-cultural perspective will be able to help foreign learners choose the proper causal connective expressions so that they can express their intentions with a cultural understanding and avoid misuse of language in a cultural context.

## References

- Choi, S.: A Study of the Meaning-Based Categorization of Grammar Items for Synonymous Grammar Education of Korean Language, Doctoral dissertation. Yonsei University, Seoul (2022)
- Conklin, H.: Hanunoo Agriculture. Food and Agriculture Organization of the United Nations, Rome (1957)
- Muriel, S.T.: The Ethnography of Communication: An Introduction, 3rd edn. Blackwell, Oxford (2003)
- Wierzbicka, A.: Understanding Cultures through Their Key Words: English, Russian, Polish, German, Japanese. Oxford University Press, New York (1997)
- Wierzbicka, A.: English causative constructions, In: Enfield, N.J. (ed.) Ethnosyntax. Oxford University Press, New York (2002)

**ICIC2023**



# Design and Implementation of Electromagnetic Field Simulation Platform for AC Testing

Liang Zhang<sup>(✉)</sup>, Lijun Fu, Yanling Guan, Jingyuan Chi, Han Wang,  
and Jianquan Liang

China State Grid Heilongjiang Electric Power Company Limited Electric Power Research  
Institute, Harbin 150030, Heilongjiang, China  
sgcc00240231@126.com

**Abstract.** The solution of electromagnetic field boundary value problem is the difficulty and key point in the teaching of electromagnetic theory. Taking the simple static two-dimensional electric field boundary value problem as an example, the research group adopts the analytical method and numerical method to solve it at the same time. It is programmed based on the MATLAB simulation platform, and designs the experiment level by level from the point of view that the conclusions of the analytical method and the numerical method echo each other, so that the students can have a perceptual understanding of the solution method of the electromagnetic field boundary value problem, the abstract and complex mathematical conclusion and the uniqueness theorem.

**Keywords:** Electromagnetic Field Boundary Value Problem · Uniqueness Theorem · Analytical Method · Numerical Method

## 1 Introduction

“Electromagnetic field theory” or “electromagnetic field and electromagnetic wave” is an important basic course for electronic information and radio technology majors in engineering colleges [1]. It involves many vector calculus formulas and abstract concepts. It requires students to have a solid foundation in mathematics and physics. It is difficult for students to generally reflect. Although the teaching of “electromagnetic field theory” should focus on the basic theory of electromagnetic field and electromagnetic wave, it should also focus on the combination with engineering practice, so as to provide more direct theoretical guidance for students in engineering colleges in related courses and directions. Among them, the solution of electromagnetic field boundary value problem is a bridge connecting basic theories such as electromagnetic field Maxwell equation with various complex engineering practices, such as antenna design, electromagnetic interference and electromagnetic compatibility, radar cross-sectional area and other related applications. However, because it involves complex mathematical theories such as mathematical equations, it has become a difficulty in undergraduate teaching [2].

The boundary value problem of electromagnetic field refers to the mathematical equation that satisfies the specific partial differential equation and boundary value condition. The solution of the boundary value problem of static electricity and magnetic field refers to the solution of the mathematical equation that satisfies the Poisson equation or Laplace equation and the specified boundary value condition. The solution methods of electromagnetic field boundary value problems are mainly divided into two categories: analytical method and numerical method. Analytical method is a method that can directly obtain the exact expression of the problem from the electromagnetic theory through formula derivation. This kind of method is usually only suitable for solving some special boundary value problems with simple boundary shape, such as the boundary shape is regular planar, spherical or cylindrical. Numerical method is a series of methods that divide the solution area into discrete grids and discretize the continuous variables to be solved into a series of discrete points in the solution area [3]. The common finite difference method, finite element method and moment method. Because this kind of method can deal with various complex boundary problems flexibly, it is widely used in practical engineering calculation. At this stage, undergraduate teaching generally only emphasizes the analytical solution of some special boundary value problems, such as the separation of variables method and the mirror image method. Due to the limitations of teaching hours, the more general numerical methods are not explained thoroughly, or even passed by, so it is difficult to achieve the real purpose of serving the engineering practice.

In this paper, taking the solution of two-dimensional electrostatic field boundary value problem as an example, a self validating experiment combining analytical method and finite difference numerical method is designed to improve students' understanding of the solution method of boundary value problem and the uniqueness theorem of the solution of boundary value problem.

## 2 Boundary Value Problems to Be Solved and Their Analytical Solutions

The two-dimensional electrostatic field boundary value problem to be solved is shown in Eq. (1), which satisfies the first kind of boundary value condition, and the surface charge density in the solution region is  $\rho$ , The lengths of x-axis and y-axis are  $a$  and  $b$  respectively, and  $U_0$  is a constant. The partial differential equation and boundary value condition satisfied by solving the potential in the region are written as follows:

$$\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} = -\frac{\rho(x, y)}{\epsilon} \quad (1)$$

$$\varphi(0, y) = 0, \varphi(a, y) = 0 (0 \leq y \leq b) \quad (2)$$

$$\varphi(x, 0) = 0, \varphi(x, b) = U_0 (0 \leq x \leq a) \quad (3)$$

According to the uniqueness theorem of static electromagnetic field, as long as the potential function satisfies the corresponding Poisson equation or Laplace equation and

the given boundary conditions, the potential function is the only correct solution. The boundary value problem can be decomposed into two boundary value problems by superposition. As shown in Eqs. (2) and (3), there are  $\varphi = \varphi_1 + \varphi_2$ .

Under the unified framework of the uniqueness theorem, the above solutions are jointly solved by the separation of variables method, superposition method and series expansion method. The orthogonality principle of trigonometric function is used in the coefficient solution. The derivation process is complex and there are many formulas, which makes it difficult for students to accept. Therefore, the solution of electromagnetic field boundary value problems has always been a difficulty in undergraduate teaching [4].

### 3 Principle of Numerical Method

This experiment uses the finite difference method to solve the problem. The code implementation of this method is relatively simple, which is more suitable for undergraduate students to write their own code [5]. They can not only understand the essence of the numerical calculation method, but also do not occupy too much energy in the implementation. In addition, the different solutions for solving simultaneous equations can also reflect the demand for accelerating the convergence speed while ensuring the calculation accuracy in numerical calculation, so that students can have a more comprehensive understanding of numerical solution methods.

Firstly, the solution area is divided into square grids. The subscripts  $i$  and  $j$  represent the sequence numbers of the nodes in the  $X$  and  $Y$  directions respectively. The grid length is  $h$ . then the difference equation of Poisson equation is shown in Eq. (4).

$$\varphi_{ij} = \frac{1}{4} \left( \varphi_{i-1j} + \varphi_{ij-1} + \varphi_{i+1j} + \varphi_{ij+1} + h^2 \frac{\rho}{\varphi} \right) \quad (4)$$

If the number of nodes in the domain division is  $n$ , then  $n$  simultaneous equations are obtained. Further considering the known boundary conditions, the main solution methods include matrix inversion, Gauss elimination and iterative method. The iterative method is to assign an initial value to each node to be solved, and gradually update the potential of each node according to a certain sequence and formula (4) until the error of the updated value before and after each node is within the given accuracy range. Because the iterative method is more suitable for solving large-scale boundary value problems, more versatile, and simple to implement, the iterative method is used in this experiment. The iteration method is divided into simple iteration method and over relaxation iteration method. If the update order is from left to right and from bottom to top, the nodes on the left and bottom have been updated each time in the simple iteration method, so the over relaxation iteration method is introduced to improve the convergence speed.

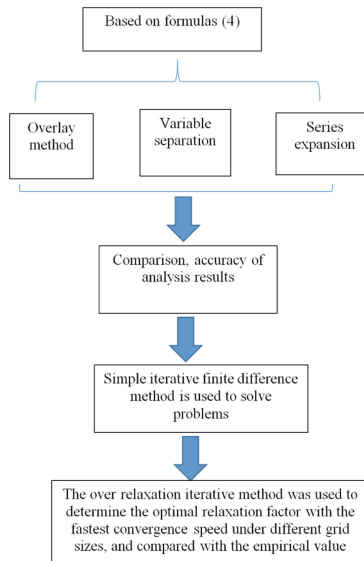
### 4 Experimental Design

This experiment requires students to write their own code to achieve analytical conclusion analysis, numerical method solution, etc. MATALB provides an efficient simulation

environment based on matrix operation, with rich function library and drawing function, simple code writing, and is widely used in theoretical research. This experiment is based on the MATLAB platform to compile the finite difference code and draw and analyze the later data. The experimental design is from the perspective that students can self verify the calculation conclusion. The experimental process and requirements are designed as follows, As shown in Fig. 1 below [6]:

- 1) Based on formula (4), analyze the analytical solution based on superposition method, separation of variables method and series expansion method;
- 2) The simple iterative finite difference method is used to solve the problem. Different grid sizes are used. The numerical solution is compared with the analytical method to analyze the accuracy of the solution results;
- 3) Use the over relaxation iteration method to determine the optimal relaxation factor with the fastest convergence speed under different mesh sizes, and compare it with the empirical value.

This experiment takes the solution of boundary value problem (1) as an example, and the boundary value problem to be solved can be adjusted according to the actual needs. By adjusting the grid size, students can understand the influence of grid dispersion on the accuracy of calculation results and the requirements for computing resources in numerical calculation; The solution of the over relaxation factor also makes students realize the practical requirements of ensuring the solution accuracy and improving the convergence speed in numerical calculation. At the same time, the experiment follows the uniqueness theorem of solving the boundary value problem of electromagnetic field, that is, the same boundary value problem can be solved by different methods, but the obtained



**Fig. 1.** The experimental process and requirements are designed as follows

solution must meet the specified partial differential equation and boundary value conditions. The uniqueness theorem usually passes by in undergraduate classroom teaching, and students are difficult to realize its important position in solving electromagnetic field boundary value problems. Through this experiment, students can deeply understand the significance and importance of the uniqueness theorem, which really plays the purpose of experiment assisted teaching.

## 5 Analysis of Experimental Results

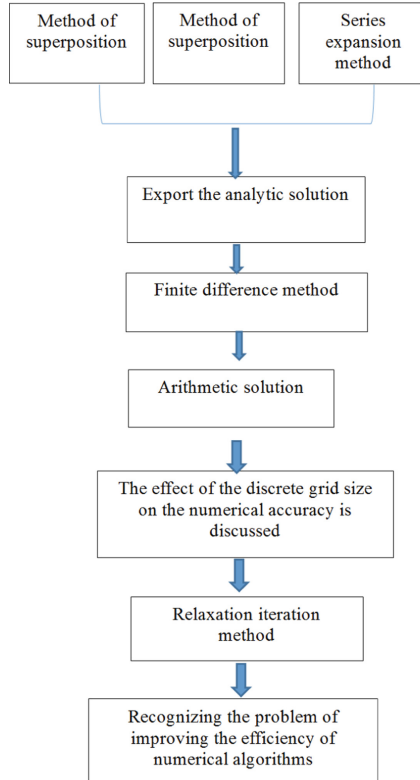
Let the size of the solution area  $a = 10$  m,  $b = 8$  m,  $u_0 = 100$  V, and the charge area density  $\rho = 10 \times (y-1)/\epsilon_0$  pC/m<sup>2</sup>; Three grid sizes  $h = 0.1$  m,  $h = 0.5$  m and  $h = 1$  m are used in the numerical calculation. The maximum difference between the two iterations is required to be  $10^{-10}$  v. the following is the corresponding analysis conclusion. (a) (b) (c) and (d) are the color map of potential distribution in the whole calculation area obtained by the analytical method and the finite difference method with grid sizes of  $h = 1$  m,  $h = 0.5$  m and  $h = 0.1$  m. The finite difference method adopts the simple iterative method. If (a) and (b) are the relative errors of the numerical solution relative to the analytical solution under three grid sizes along the two lines  $y = 1$  m and  $y = 5$  m, respectively. It can be seen that with the decrease of the grid size, the numerical solution gradually tends to the analytical solution, which indicates that the finer the grid is, the higher the calculation accuracy is. At the same time, it also proves that under strictly consistent boundary conditions and partial differential equations, the conclusions of multiple solutions to the same boundary value problem tend to be consistent, that is, the unique correctness of the solution [7].

**Table 1.** Comparison of Iteration Times of Over relaxation Iteration Method and Simple Iteration Method

H(m)	Simple iteration (times)	Optimization relaxation iteration (times)	Ratio%	$\alpha_{opt}$
1	393	45	11.45	1.5
0.5	1508	90	5.97	1.7
1	32771	478	1.45	1.94

The comparison of iteration times between the final simple iteration method and the optimal over relaxation iteration is shown in Table 1, where the ratio represents the ratio of the minimum iteration times of the over relaxation iteration method to the iteration times of the simple iteration method [8]. It can be seen that by adjusting the over relaxation factor, the calculation efficiency can be greatly improved on the premise of ensuring the calculation accuracy, and the more the number of nodes, the more obvious the improvement effect is, which is of great significance for the numerical solution of large-scale boundary value problems.





**Fig. 2.** The main content of the article study

## 6 Conclusions

The solution of electromagnetic field boundary value problem has always been a difficult point in the undergraduate teaching of electromagnetic field and electromagnetic wave because of the complexity of mathematical formulas and the diversity of methods involved. In this paper, the solution of a two-dimensional electrostatic field boundary value problem is taken as an example. Firstly, the analytical solution is derived based on the superposition method, the separation of variables method and the series expansion method; Then, the finite difference method is further used for numerical solution, and the influence of the size of discrete grid on the accuracy of numerical calculation is discussed [9]. The over relaxation iteration method is used to make students realize the problem of improving the efficiency of numerical algorithm. As shown in Fig. 2 below. Each step of the conclusion of this experiment has a reference conclusion, which students can self verify; the algorithm code based on MATLAB is relatively simple. Students have deepened their understanding of the algorithm by writing the code themselves [10]. The experimental design deepens the students' understanding of the abstract uniqueness theorem and its important position in solving electromagnetic field boundary value problems from

the perspectives of analytical method and numerical method, helps the students understand the numerical calculation method of electromagnetic field, and further promotes the students' flexible learning and application of basic equations, boundary conditions and various second-order partial differential equations in electromagnetic field theory.

**Acknowledgements.** Project supported by State Grid Corporation of China: Research on the technology of electric field control at residential platform near transmission lines (contract number: 8100-202055155A-0-0-00).

## References

1. Kim, M., Park, S.W., Jung, H.-K.: An advanced numerical technique for a quasi-static electromagnetic field simulation based on the finite-difference time-domain method - ScienceDirect. *J. Comput. Phys.* **373**, 917–923 (2018)
2. Kalis, H., Marinaki, M., et al.: On numerical simulation of electromagnetic field effects in the combustion process. *Math. Model. Anal.: Matematicinis modeliavimas ir analize* **23**(2), 327–343 (2018)
3. Pevneva, N.A., Kopshai, A.A., Gurskii, A.L.: Simulation of electromagnetic field distribution in the measuring cell for determining the dielectric permittivity of materials at microwave frequencies. *Doklady BGUIR* **18**(6), 75–80 (2020)
4. Orozco, E.A., Estupiñán, A., Dugar-Zhabon, V.D.: Simulation of the electromagnetic field in a rectangular multimode chamber of a 14GHz ion source. *J. Phys. Conf. Ser.* **1159** (2019)
5. Tong, S.Y., Feng, G., Lian, Z.M., et al.: Simulation on temperature field of initial segment of electromagnetic rail launch. *IOP Conf. Ser. Mater. Sci. Eng.* **657**, 012067 (2019)
6. Caiquanwen, Y.Z.: Electromagnetic field simulation of large hollow reactor. *IOP Conf. Ser. Earth Environ. Sci.* **512**, 012112 (2020)
7. Maurya, A., Kumar, R., Jha, P.K.: Simulation of electromagnetic field and its effect during electromagnetic stirring in continuous casting mold. *J. Manuf. Process.* **60**, 596–607 (2020)
8. Yarymbash, D.S., Yarymbash, S.T., Kotsur, M.I., et al.: Computer simulation of electromagnetic field with application the frequency adaptation method. *Radio Electron. Comput. Sci. Control* **1**, 65–74 (2018)
9. Zhang, S.S., Lei, C.J., Liu, Y.H., et al.: On simulation of 3D-FDTD for transient electromagnetic field. *Dianzi Keji Daxue Xuebao/J. Univ. Electron. Sci. Technol. China* **48**(1), 13–21 (2019)
10. Osaci, M.: Numerical simulation methods of electromagnetic field in higher education: didactic application with graphical interface for FDTD method. *Int. J. Mod. Educ. Comput. Sci.* **10**(8), 1–10 (2018)



# Organizational Layout and Optimization Model of Agricultural Logistics Industry Based on Ant Colony Algorithm

Jingjun Shu<sup>(✉)</sup>

Wuhan Business University, Wuhan, Hubei, China  
hubeiwuhansjj@163.com

**Abstract.** With the rapid rise of rural e-commerce and the steady increase in the level of rural distribution demand, due to the weak rural logistics infrastructure and imperfect distribution infrastructure and distribution system, the traditional conditions of rural logistics and distribution include uneven resource allocation, information asymmetry, and enterprises. Inability to communicate with each other, lack of knowledge sharing network and other issues. The lack of system integration and distribution has led to problems such as low efficiency, high cost, and poor performance at the logistics end, which have not been effectively solved in rural areas for a long time. Based on the ant colony algorithm, this paper studies the organizational layout and optimization model of the agricultural logistics industry. This paper analyzes the cost of rural express logistics distribution, and analyzes the design goals and principles of the optimization model. The experimental results show that the scale of rural netizens and the rural Internet penetration rate have continued to increase in the past five years, reaching 293 million and 44.71% of the population and popularization respectively. With the rapid popularization of the Internet in rural areas, rural logistics products have high expectations and great potential.

**Keywords:** Ant Colony Algorithm · Agricultural Logistics · Industrial Organization Layout · Optimization Model

## 1 Introduction

In the process of implementing the rural revitalization strategy, rural express logistics undertakes the important historical mission of activating the rural economy. With the rapid rise of rural e-commerce and the steady increase in the level of rural distribution demand, it has become a key link in the “connecting the past” in rural economic and social development. An important indicator to measure the quality of online shopping is the speed of logistics distribution, and the speed of logistics distribution is determined by the pros and cons of logistics route selection. A fast logistics route selection scheme can give customers a good shopping experience and reduce The overall cost of an e-commerce platform or my country’s logistics and distribution [1, 2].

In the relevant research, Tadi mentioned that logistics is the main means to effectively realize the flow of people, goods and information in the rural tourism supply chain (RTSC) and improve the competitiveness of tourism products [3]. Logistics provides material and non-material basis for rural tourism services. The author analyzes the key issues and structures of RTSC, constructs the logistics structure of agritourism, and analyzes specific fields from the perspective of logistics processes, processes and activities. Vakhidov et al. proposed a model for calculating the braking parameters of transport and technical agricultural machinery equipped with ultra-low pressure wheels [4]. The difference between this model and the previous model is that its output parameter is not the braking efficiency, but the time difference between the front and rear axles locked. The results show that satisfying the advance locking condition of the front axle ensures the stability of the tractor movement during emergency braking, which has a positive impact on road traffic safety.

The main purpose of this paper is to study the organizational layout and optimization model of the agricultural logistics industry based on the ant colony algorithm. This paper analyzes the cost of rural express logistics distribution, and analyzes the design goals and principles of the optimization model. In this paper, the ant colony algorithm is used to solve the VRP problem, and the parameters are set based on the current status of the system. The research is of great significance for improving the process of rural logistics. In practice, the rural logistics distribution model developed in this paper provides a model for the development of rural logistics distribution; the control measures obtained from the study have important reference value for promoting rural logistics and policy making.

## 2 Design Research

### 2.1 Analysis of the Cost of Rural Express Logistics Distribution

#### (1) High transportation cost

The rural express logistics infrastructure is weak, its transportation organization is unreasonable, the network layout and rural residents are widely distributed, and the number of express parcels in the same area is unstable, which leads to roundabout transportation, repeated transportation and empty vehicle transportation of delivery vehicles to the countryside. At the same time, the unloaded rate of delivery vehicles to the countryside is high, resulting in high transportation costs for express delivery to the countryside [5, 6].

#### (2) High transit costs

The rural distribution area is wide, but the rural express logistics distribution system is imperfect compared to the city, and a professional and stable distribution system has not been formed. Usually, it needs to go through multiple transfers to reach the terminal distribution network or agent. The operation leads to an increase in the transfer cost and the storage and transportation cost during the transfer process; and the problems such as the prolonged distribution time caused by this increase the time cost of distribution.

#### (3) The delivery cost is high

Due to the small and scattered demand for rural end distribution, the establishment of distribution outlets will increase the cost of outlet construction. In practice, most

companies are unwilling and unable to set up rural distribution outlets, resulting in reduced “last mile” distribution efficiency in rural distribution; Or affected by other factors, it is often difficult to complete a one-time pickup, which leads to an increase in the cost of secondary distribution.

(4) High cost of operation and implementation

Restricted by the objective environment in rural areas, the resources of rural express logistics and distribution are limited, the hardware foundation is poor, and the “poor, narrow, and weak” rural roads in remote areas lead to low road accessibility and high vehicle depreciation costs; To achieve interconnection, the phenomenon of waste of distribution resources is serious.

## 2.2 Design Goals and Principles

The design goal of this system is to improve the parts that do not conform to the business process based on the existing system functions of the distribution center. At the same time, an intelligent dispatching module is added to realize the system automatically dispatching vehicles and planning the distribution path, so as to provide the distribution center with the business process and operation. Simple information management, vehicle scheduling services [7, 8].

In order to achieve the system goals, the system should follow the following principles when designing:

(1) Security principle

There is a large amount of enterprise internal information stored in the system, so information security is very important. The vehicle dispatching platform of the distribution center is used in the intranet of the enterprise, and the data transmission is also connected with the internal system of the enterprise, and the security is relatively strong. When the user logs in to the system and performs operations, the user’s identity should be verified, and the user’s password should be irreversibly encrypted, and the setting of multiple input wrong passwords to lock access should be activated to prevent software attacks that crack the password. The system administrator should clean up the user information of the resigned staff in a timely manner to prevent the information from being stolen.

(2) The principle of reliability

The system needs to work 24 h a day to process orders from various e-commerce platforms at any time. If there is a failure, it can be recovered within 12 h, and the backup data can be used to ensure the normal operation of the system [9, 10].

(3) The principle of scalability

With the continuous improvement of the business scale, the business scale of the distribution company will also continue to expand. To meet business requirements at the same time, the design of the system should enable integration between small and functional units.

(4) The principle of portability

There are many urban distribution centers in logistics enterprises, and the system should be designed to be as universal as possible, so that the system can be used in local distribution centers after briefly modifying some parameters or adding or subtracting business modules.

(5) The principle of ease of use

The purpose of using the system is to improve work efficiency, and the vehicle dispatchers and managers faced by the system are usually not computer professionals. Therefore, the system needs to provide a user-friendly operation interface, minimize manual operations, and facilitate learning and use [11, 12].

### 2.3 Algorithm Operation Process

The steps to solve the VRP problem using the ant colony algorithm are as follows:

(1) Parameter initialization

$m$  is the number of insects,  $\alpha$  is the main pheromone factor,  $\beta$  is the heuristic activity,  $\rho$  is the vaporized pheromone,  $Q$  is the total amount of pheromone released, and  $n$  is the maximum number.

(2) Constructing the solution space

All insects are placed in a distribution center, and each insect selects the next distribution point for distribution based on the concentration of “pheromone”. The calculation process is:

$$p_{ij}^k = \begin{cases} \frac{[\tau_{ij}(t)]^\alpha \cdot [\eta_{ij}(t)]^\beta}{\sum_{s \in I_k} [\tau_{ij}(t)]^\alpha \cdot [\eta_{ij}(t)]^\beta}, & s \in I_k \\ 0, & s \notin I_k \end{cases} \quad (1)$$

where  $\tau_{ij}(t)$  is the pheromone concentration. At the beginning of the analysis, the pheromone concentration is the same between receptor sites, so assuming  $\tau_{ij}(0) = 0$ , is a set of receptors that do not transmit  $k$ , a heuristic function calculated by Eq. (2).

$$\eta_{ij}(t) = \frac{1}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}} \quad (2)$$

The larger the number of heuristic activities, the higher the probability of insect selection until all insects have completed the delivery of all collection points and returned to the distribution center.

(3) Update pheromone

After the search is completed, the path length of each insect is calculated, the shortest path in the current iteration is recorded, and the pheromone concentration  $T_{Ary}(i,j)$  between each receiving point is updated according to formula (3):

$$\begin{cases} \tau_{ij}(t+1) = (1 - RHO)\tau_{ij}(t) + \Delta\tau_{ij} \\ \Delta\tau_{ij} = \sum_{k=1}^n \Delta\tau_{ij}^k \end{cases} \quad (3)$$

$\Delta\tau_{ij}^k$  represents the pheromone concentration,  $k$  is the number of animals, and  $i, j$  are the receiving points.

The calculation formula (4) of the antcyclesystem model is as follows:

$$\Delta\tau_{ij}^k \begin{cases} Q/\text{Length}_{ij}, & \text{The } k\text{th ant visits } j \text{ from the receiving point } i \\ 0, & \text{other} \end{cases} \quad (4)$$

#### (4) Judging termination conditions

When the number of iterations reaches the preset maximum number of iterations, stop the work and obtain the optimal solution; otherwise, delete the insect path record and resume the second step.

### 3 Experimental Study

#### 3.1 System Status Analysis and Parameter Setting

Through investigation, it is found that the distribution center has a complete logistics information management system, and the informatization level is higher than the industry average. First, the existing system of the distribution center will be analyzed. If the system calls the algorithm proposed above for vehicle allocation, the following parameters are required:

- (1) Address number
- (2) The latitude and longitude coordinates of the delivery address
- (3) The total volume of goods to be delivered at the receiving point
- (4) Available fleets, number of available franchise vehicles and corresponding models

However, during the investigation, it was found that the data management functions of the existing system have the following defects:

##### (1) The address library is not fully utilized

According to the survey, the system is embedded with a GIS system, which can automatically obtain the information of provinces, cities, districts, counties, and latitude and longitude of the receiving address and mark it on the map. However, in the actual car distribution process, the address database information is only used to divide the order area according to the administrative district and county where the delivery address is located, and the rest of the data is not fully utilized.

##### (2) Special orders cannot be marked through the system

In the process of allocating vehicles, the dispatcher is sometimes required to handle orders with special needs alone, but such orders cannot be identified by the system, and can only be obtained by the dispatcher recording the order number through manual inquiry, and corresponding processing. Therefore, the system cannot judge the order entering the automatic vehicle distribution process (normal order), and thus cannot calculate the total volume of the goods to be delivered at the corresponding receiving point.

##### (3) The vehicle management is chaotic and the information is not updated in time

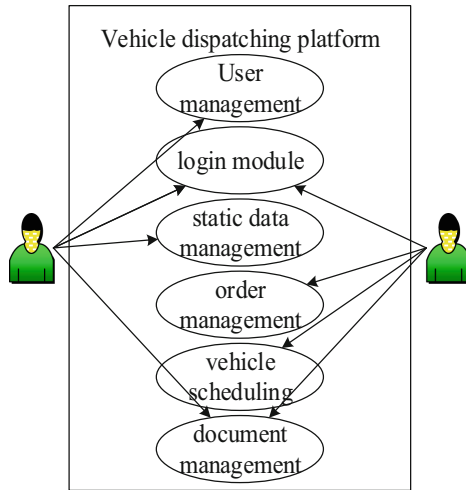
The distribution center carrier is divided into two categories: fleet and franchised vehicles, but the system has loopholes in the management of the two. First of all, the basic information is incomplete, and the system cannot reflect the actual docking situation of the fleet and the franchised vehicles. Secondly, the system cannot reflect the real-time status of the vehicle, and cannot know through the system whether the participating vehicle is currently on the way or idle, and whether the fleet can undertake the delivery task. The above information is obtained by the dispatcher through offline

inquiry. Therefore, the system cannot provide the available fleet, the number of available franchise vehicles and the corresponding model.

The above problems all lead to the inability of the system to provide corresponding parameters for automatic vehicle allocation, and are also the key problems to be solved in the following system analysis and design.

### 3.2 Functional Module Requirements

According to the business requirements and role analysis of the distribution center, the vehicle scheduling platform should include the following six functional modules, namely: login module, user management, static data management, order management, vehicle scheduling, and document management. The specific functional requirements are shown in the following Fig. 1:



**Fig. 1.** The overall functional module requirements of the vehicle dispatching system

#### (1) User management

User management is mainly used to record and verify the identity of the system user, and the user needs to log in with the user name and password. The account number and initial password are uniformly assigned by the system administrator, and the user can change the password after logging in for the first time.

#### (2) Static data management

Static data management is mainly used to record data with high frequency and low update frequency in the system, including fleet information, franchise vehicle information and address database (for B2B business, customers are relatively fixed, so addresses are regarded as static data). The administrator can add, modify, delete and query the data of the joined car, fleet and address database. In order to ensure the stability of the system operation, the vehicle dispatcher can only query static data as required, and cannot perform other operations.



### (3) Order management

The orders of the distribution center come from major e-commerce websites, which are uniformly processed by the order processing system and converted into standard formats and then directly imported into the vehicle dispatching platform. Therefore, there is no need to do anything with the generated order unless there are special circumstances. If there are special circumstances (such as expedited delivery, etc.), the system administrator and the vehicle dispatcher can mark the order specially, and the dispatcher will arrange it separately when arranging the vehicle delivery.

### (4) Vehicle scheduling

The vehicle dispatching module is the core of the whole platform and is only operated by the vehicle dispatcher. The vehicle dispatcher needs to manually allocate vehicles for the specially marked orders according to the original operation process. For ordinary orders, no redundant operations are required, and the platform automatically allocates the delivery vehicles according to the algorithm and issues the rush orders. After the order grab is over, the system will automatically generate a dispatch order, a delivery order and a delivery order, and the dispatcher can confirm and print it after confirming that it is correct.

### (5) Document management

Document management is mainly used to record the execution of the order. After the driver completes the delivery work and returns the delivery order, the system will confirm it, forming a closed-loop operation. The dispatcher can query the carrier and delivery status of the order according to the document management.

## 3.3 Non-functional Requirements

### (1) Response speed requirements

Since the distribution center has strict requirements on delivery timeliness, the response time of the system should be fast, and it should not take more than 10 min from order input to completion of vehicle distribution.

### (2) Input and output requirements

The system input mainly comes from two aspects, one is the order information entered by the order management system and the order grabbing result returned by the order grabbing system, and the other is the administrator and operator, which requires a system interface and a friendly input interface.

The system output mainly includes printing documents, sending orders for grabbing orders and sending order distribution results to the WMS system, so the system needs to have a system interface and connect to a printing device.

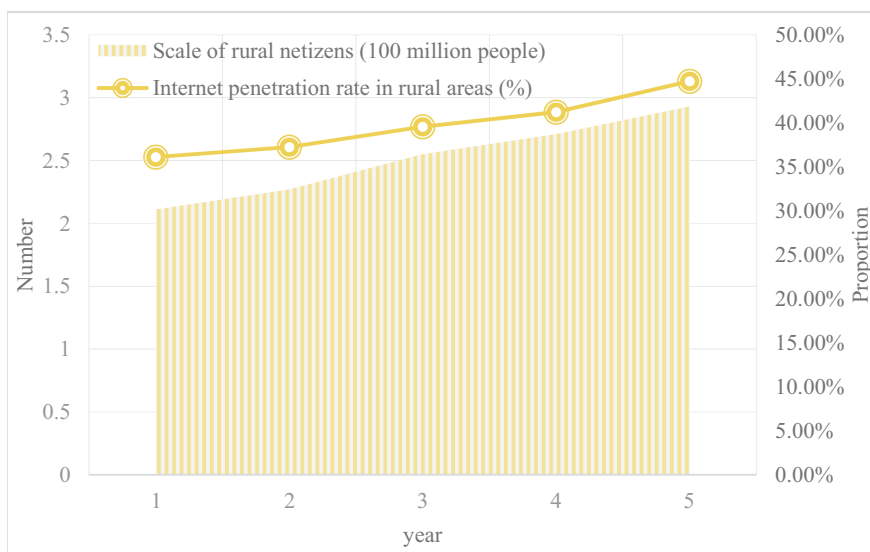
## 4 Experiment Analysis

### 4.1 Characteristics of Rural Express Logistics and Distribution

It is mainly reflected in the huge market potential and the booming of rural e-commerce. The following are the statistics on the scale and popularization of rural netizens in the past five years as shown in Table 1:

**Table 1.** Statistics on the scale of rural netizens and the Internet penetration rate in the past five years

years	1	2	3	4	5
Scale of rural netizens (100 million people)	2.11	2.27	2.55	2.71	2.93
Internet penetration rate in rural areas (%)	36.10%	37.23%	39.54%	41.21%	44.71%

**Fig. 2.** Analysis of the scale of rural netizens and the Internet penetration rate in the past five years

Analysis of Fig. 2 shows that the scale of rural netizens and the rural Internet penetration rate have continued to increase in the past five years, reaching 293 million and 44.71% of the population and popularization respectively; The use of retail, whether it is rural online retail sales and market share, is also increasing year by year, and rural logistics products have high expectations and great potential.

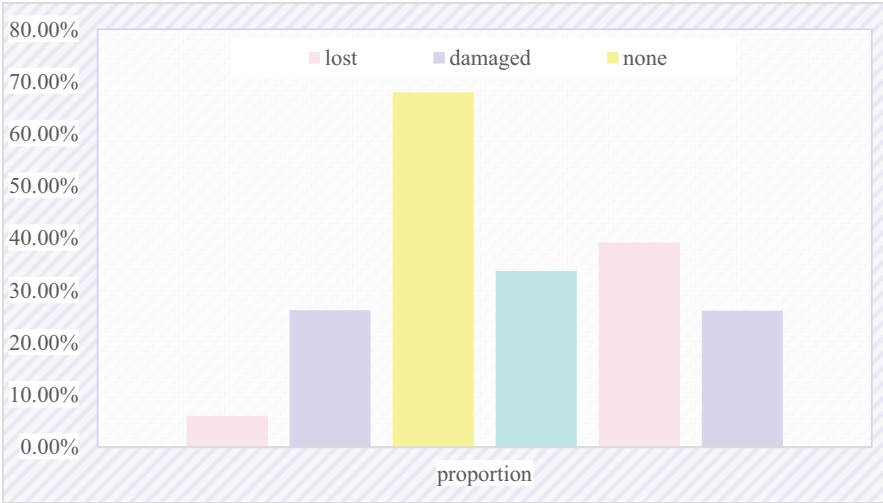
## 4.2 Delivery Service Issues

### (1) Problems occur from time to time

Due to the low level of specialization in the logistics distribution of rural express terminals and the uneven quality of human distribution, component problems often occur in the distribution of rural logistics terminals. The parcels of express parcels are damaged, the quantity of goods is in short supply or even lost. The percentage of problem pieces in the survey is as follows Table 2 and Fig. 3.

**Table 2.** Proportion of problem parts of express shipments and statistics of complaint handling of problem parts

problems and attitudes	express problem			attitude towards results		
	lost	damaged	none	dissatisfied	generally	satisfy
proportion	5.9%	26.2%	67.9%	33.7%	39.2%	26.1%



**Fig. 3.** Analysis of the proportion of problem pieces of express shipments and the results of complaint handling of problem pieces

Due to the large number of delivery links at the end of rural express logistics, ineffective transit links and long delivery time, it is difficult to divide the boundaries of responsibilities in the delivery process. Complaint handling issues with low satisfaction.

## 5 Conclusions

Rural express logistics is a concept of reduced regional logistics. Rural express logistics is to serve the vast number of rural residents in rural areas, including cargo handling, packaging, storage, sorting, distribution, delivery, distribution processing and information services. And a series of logistics activities to meet the fast demand of rural residents for the growing production and living materials. With the popularity of online shopping, more and more people choose to buy the goods they need online, but the speed of receiving the goods will directly affect the user's choice direction. The quality of the logistics path selection is an important factor in determining the delivery speed of the goods. How to establish an optimized logistics distribution route selection scheme is an important issue.

**Acknowledgements.** The National Social Science Fund of China: Research on the coordinated development mode of agricultural industry system based on agricultural logistics park (2018BJY138).

## References

1. Olkhova, M., Roslavtsev, D., Matviichuk, O., et al.: City delivery routes planning based on the ant colony algorithm. *Sci. Tech.* **19**(4), 356–362 (2020)
2. Bhutta, M., Ahmad, M.: Secure identification, traceability and real-time tracking of agricultural food supply during transportation using Internet of Things. *IEEE Access* **9**, 65660–65675 (2021). <https://doi.org/10.1109/ACCESS.2021.3076373>
3. Tadi, S., Veljovi, M.: Logistics of rural tourism. *Int. J. Traffic Transp. Eng.* **10**(3), 323–350 (2020)
4. Vakhidov, U.S., Kurkin, A.A., Levshunov, L.S., et al.: Ensuring the stability of agricultural transport and technological machines equipped with ultra-low pressure tires during braking. *Eng. Technol. Syst.* **30**(4), 609–623 (2020)
5. Radovi, S.: Innovative solution in the modeling of food distribution channels as a factor of successful organization of agricultural production. *Poljoprivredna Tehnika* **45**(3), 38–43 (2020)
6. Antara, M., Sumarniasih, M.S.: Featured food commodities for food security support in Bali Province, Indonesia. *Agricul. Soc. Econ. J.* **20**(2), 147–158 (2020)
7. Agholor, A.I., Nkosi, M.: Sustainable water conservation practices and challenges among smallholder farmers in Enyibe Ermelo Mpumalanga province South Africa. *J. Agricul. Exten.* **24**(2), 112–123 (2020)
8. Barakat, A., Saad, N.A., Hammad, M.A.: key performance indicators of cold supply chain practices in agriculture sector empirical study on the egyptian export companies. *Open Indust. Manufac. Eng. J.* **14**(10), 1002–1006 (2020)
9. Kupalova, H.I., Goncharenko, N.V.: State stimulation of the development of organic crop production in Ukraine. *Probl. Econ.* **2**(44), 144–152 (2020)
10. Medeiros, D.L., Kiperstok, A.C., Nascimento, F., et al.: Human urine management in resource-based sanitation: water-energy-nutrient nexus, energy demand and economic performance. *Sustain. Product. Consumpt.* **26**(-), 988–998 (2021)
11. Akn, Y., Elen, B., Elen, M.F., et al.: Agriculture and pandemic: how should turkish agriculture change after COVID-19? *EJONS Int. J. Math. Eng. Natl. Sci.* **4**(16), 904–914 (2020)
12. Nguyen, T.D., Nguyenquang, T., Venkatadri, U., et al.: Mathematical programming models for fresh fruit supply chain optimization: a review of the literature and emerging trends. *AgriEng.* **3**(3), 519–541 (2021)



# Application of Image Recognition in Equipment Monitoring

Haidong Zou<sup>(✉)</sup>, Shaoqiang Yang, and Wei Wu

China Satellite Marine Tracking and Control Department, Jiangyin, Jiangsu, China  
abigkun@163.com

**Abstract.** To solve the incomplete remote monitoring status of equipments, video streaming is used to monitor the indicators of equipment in the front panel. With the help of image recognition technology, the working status of equipments can be automatically get. By using methods such as image binarization, grayscale processing, positioning and calculation, the status of the equipment indicators are analyzed. The corresponding working status of equipments is automatically determined to further improve the equipment status monitoring. The effectiveness of image recognition technology in equipment status monitoring is verified through practical tests in this paper.

**Keywords:** Image Recognition · Statement Monitoring · Automatic · Video

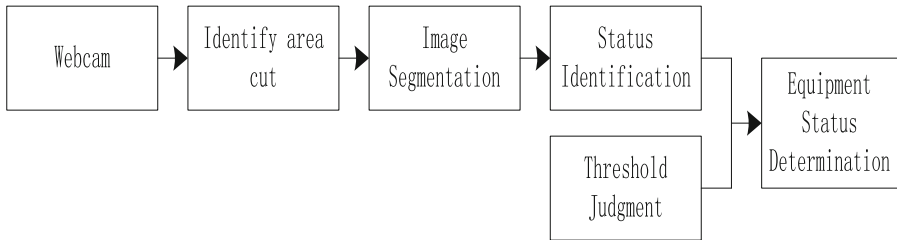
## 1 Introduction

Due to the inadequate remote monitoring data of some devices, it is difficult for the managers to judge the working status of this device, which in turn affects the verdict on the working status of the whole system. Such problems occur from time to time, which is not conducive to the unattended demand of the machine room. Take a certain type of inverter as an example, its remote monitoring data does not show the working status of the equipment, but only provides the working parameter settings of the equipment. In case of equipment failure, it requires management personnel to dispose of it on site, which seriously affects the efficiency of problem disposal. However, the content of the panel indicator of the device is relatively rich, including information such as power indication, this vibration alarm, alarm storage, etc. From the field working status of the indicator of the device panel, you can basically judge the working status of the device quickly. If the panel indicator field work status, in the form of data communication to remote management personnel, can greatly enhance the system fault disposal efficiency.

Given the richness of current video surveillance means and the development of target recognition technology based on video surveillance, its application to practical problems can improve the processing efficiency [1]. In this paper, we intend to use video surveillance to monitor the status of equipment panel instructions in real time, and then realize the remote monitoring function of the working status of some equipment through automated technologies such as image recognition and data processing. This helps to improve the efficiency of equipment management and to provide technical support for the realization of unmanned server rooms.

## 2 System Architecture

The system is implemented by a combination of hardware and software. The status indicator information of the device panel is collected through a webcam and transmitted to the computer at the remote end. Then, through software programming, the recognition area of the captured image is cut out and processed as much as possible without missing key information. Then the status indicators in the cut image are segmented by edge recognition and other technologies, and the indicator working status library is established and the indicator status judgment threshold is calculated. Finally, the current working status of the device is determined by combining the status information of all indicators. The specific implementation architecture is shown in Fig. 1.



**Fig. 1.** Architecture of the system

## 3 System Key Technology

Based on image recognition technology, the equipment working panel status monitoring system needs to solve the technical problems of equipment panel area recognition and panel indicator area segmentation, indicator status recognition and equipment working status determination, and so on.

### 3.1 Identification Area Determination

By installing identification tags on the equipment identification area, it is easy to quickly determine the identification area [2, 3]. For fixed video surveillance, the corresponding equipment panel is relatively fixed, so the recorded equipment working panel image is also fixed. That is, the relative position of the equipment panel in a monitoring picture is fixed. After the precise measurement of the acquired image, we can get the relatively accurate equipment working panel area, and remove the irrelevant area to improve the efficiency of subsequent image processing.

### 3.2 Image Segmentation

There are many indicators in the working panel of the equipment, so it is necessary to segment them according to the work requirements to improve the image calculation

speed. Fortunately, the shape of the indicators of the equipment panel in this system is relatively regular, usually mainly round indicators, and the boundary of the area between indicators is relatively obvious. According to the rules and characteristics of image processing, this paper adopts the image segmentation method based on edge detection to realize the image segmentation of panel indicators [3–5].

### 3.3 Image State Recognition

The state of the device panel indicator is only bright and off, but its color distribution is richer, commonly used are red, green, yellow, white, orange, etc.. To accurately determine the state of the indicator, it is necessary to first determine the value of each color indicator in the current lighting situation in the off state as the reference value 1, and then take the value of each color indicator in the on state as the reference value 2, with the difference between the two reference values to determine the threshold value [6]. The subsequent judgment of the working status of all indicators is based on the judgment threshold determined by their respective positions. The basic algorithm is as follows:

- a. Determine the reference value of the indicator off state 1. For the N images acquired, calculate the mean values  $r_i$ ,  $g_i$ , and  $b_i$  for each indicator in the area R, G, and B channels that are off, and then calculate the mean values  $r_m$ ,  $g_m$ , and  $b_m$  for these three channels as the reference value 1.

$$r_i = \text{mean}(\text{mean}(\alpha_r(\text{pos}_{on}))) \quad (1)$$

$$g_i = \text{mean}(\text{mean}(\alpha_g(\text{pos}_{on}))) \quad (2)$$

$$b_i = \text{mean}(\text{mean}(\alpha_b(\text{pos}_{on}))) \quad (3)$$

$$(r_m, g_m, b_m) = \frac{1}{N} \sum_{i=0}^{N-1} (r_i, g_i, b_i) \quad (4)$$

- b. Determine the reference value for the indicator on state 2. For the N images acquired, calculate the mean values  $R_i$ ,  $G_i$ ,  $B_i$  for each indicator in the area R, G, B channels that are off, and then calculate the mean values  $R_m$ ,  $G_m$ ,  $B_m$  for these three channels as the reference value 2. Since the calculation method is the same, this step can continue to use the four calculation formulas in the above step, only the distinction needs to be made for the off state indicator light value.
- c. Determine the judgment thresholds ( $R_t$ ,  $G_t$ ,  $B_t$ ). The difficulty of this step is that each indicator due to color differences and the impact of various possible relationships at the scene, resulting in their respective judgment thresholds have a large difference, so in determining the threshold value, the need to have tolerance considerations. The respective calculated base value 1 and base value 2 can be used as the final qualitative judgment indicator by taking the median value of both for the judgment threshold.

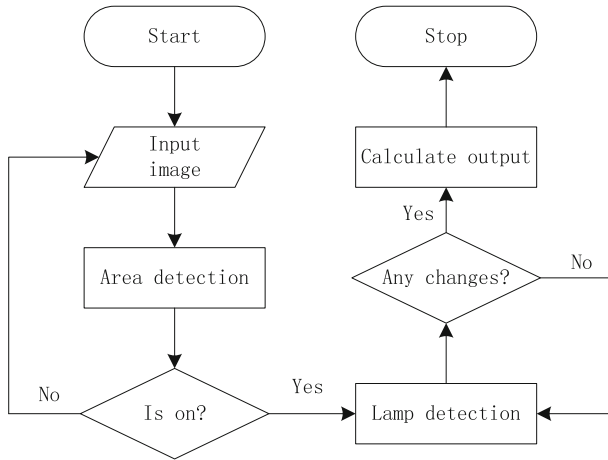
$$(R_t, G_t, B_t) = \text{mean}((r_m, g_m, b_m), (R_m, G_m, B_m)) \quad (5)$$

- d. The result is obtained by comparing the current state value with the judgment threshold. If the current value is greater than the judgment threshold, the indicator will be turned on, otherwise it will be turned off.

$$state = ((R, G, B) > (Rt, Gt, Bt)) ? 1 : 0 \quad (6)$$

### 3.4 Equipment Status Determination

The determination of the operating status of the device requires a comprehensive judgment combined with the operating status values of each indicator as described above. The basic working logic of the program is shown in Fig. 2.



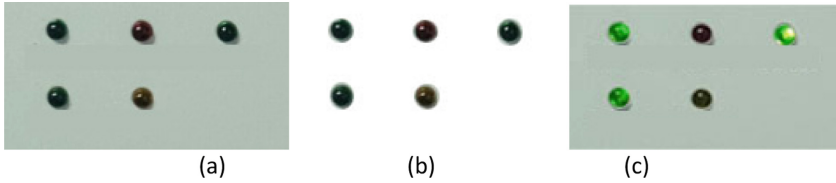
**Fig. 2.** Flow chart for determining the working status of the equipment

After the program is started, it will ask for the input of the picture to be processed, and after the detection of the positioning of the relevant working area in the picture is completed, the recognition program is started to judge whether the equipment is already in working state [7–9]. If the result is judged to be true, it enters the monitoring stage of the equipment status indicator, focusing on whether there is any abnormality in each key status indicator, and then outputs the status determination result of the equipment in this way [10].

## 4 System Test

Taking an inverter as an example, the site uses a webcam from Beiqingshitong with an effective pixel of 3 million. By calculating the working indicator area of the device and then cutting the picture, the working area of the device panel indicator is obtained, as shown in Fig. 3(a). As can be seen from the figure, the power indication, remote control and internal reference source indicators are green [9, 10], while the local vibration alarm





**Fig. 3.** (a) Working area map (b) Splitting effect map (c) Test map

and alarm storage indicators are red, and the color of each indicator differs slightly due to its location, so it needs to be segmented. After using the edge detection method, the segmentation effect is shown in Fig. 3(b).

As Fig. 3(a) is the device shutdown status indicator display, after taking the value of each indicator separately, we can get its respective base value 1. Similarly, take the value of each indicator after it is lit, and then calculate the respective base value 2, and then use it to calculate the judgment threshold. For example, if the “Benzene Alarm” indicator has a reference value of 55 and a reference value of 125, then the threshold value for determining its operating status is 90.

When the test chart is shown in Fig. 3(c), after processing and calculation, the system comes up with the current device status: the equipment is on and working normally. It is in line with the actual situation.

## 5 Conclusions

By using of image recognition technology, with the status recognition of the indicator in front of equipment panel, the status of the equipment can be checked. It can help the management to effectively dispose of the equipment problems in a timely manner and improve the system operation and maintenance efficiency. The system employs frame-splitting processing of the video images captured by the webcam to cut the effective positions in the images in order to improve the efficiency of subsequent image processing. An edge detection algorithm is used to extract the position of each indicator, and then calculate the respective state value separately, compare it with the corresponding judgment threshold, and finally determine the working status of the device. The method has good generality and has a certain degree of generalization. Subsequently, the research on state monitoring in different environments and under different lighting influences should be enhanced to improve the applicability of the system.

## References

1. Fagundes, L.G., Santos, R.: Development of computer graphics and digital image processing applications on the iPhone. In: 2010 23rd SIBGRAPI Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T). IEEE Computer Society (2010)
2. Goyal, A., Meenpal, T.: Patch-based dual-tree complex wavelet transform for kinship recognition. *IEEE Trans. Image Process.* **30**, 191–206 (2021)
3. Chakraborty, S., Chatterjee, A., Goswami, S.K.: A dual-tree complex wavelet transform-based approach for recognition of power system transients. *Expert Syst.* **32**(1), 132–140 (2014)

4. Park, J., Low, C.Y., Andrew, B.: Divergent angular representation for open set image recognition. *IEEE Trans. Image Process.* **31** (2022)
5. Chakraborty, et al.: A dual-tree complex wavelet transform-based approach for recognition of power system transients. *Expert Syst. Int. J. Knowl. Eng.* **32**(1), 132–140 (2015)
6. Dubnov, Y.A.: The feature selection method based on a probabilistic approach and a cross-entropy metric for the image recognition problem. *Sci. Tech. Inf. Process.* **48**(6), 430–435 (2021). <https://doi.org/10.3103/S0147688221060022>
7. Keys, R.G.: Cubic convolution interpolation for digital image processing. *IEEE Trans. Acoust. Speech Signal Process.* **29** (2003)
8. Solomon, C., Breckon, T.: *Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab*. Wiley, Hoboken (2011)
9. Mahmood, A., Khan, A.Q., Mustafa, G., et al.: Remote fault-tolerant control for industrial smart surveillance system. *Math. Probl. Eng.* (2021)
10. Varma, N., Epstein, A.E., Irimpen, A., et al.: Efficacy and safety of automatic remote monitoring for implantable cardioverter-defibrillator follow-up: the Lumos-T safely reduces routine office device follow-up (TRUST) trial. *Circulation* **122**(4), 325–332 (2010)



# Stability of Marine Physics Detection Sensor Based on Artificial Intelligence Technology

Xiran Liu<sup>(✉)</sup>

Changchun College of Electronic Technology, Changchun 130114, Jilin, China  
liuxiran1122@163.com

**Abstract.** In order to maintain maritime safety, marine physical exploration plays an important role in it. For the specific application of marine physical detection, marine monitoring sensors are deployed in the extremely complex and variable marine environment to realize real-time monitoring of the ocean. Therefore, the combination of artificial intelligence technology is of great significance to the design and stability research of marine physical detection sensors. The purpose of this article is to study the stability of marine physical detection sensors based on artificial intelligence technology. This article introduces the functional modules and software system of the sensor, and analyzes the stored data of each sensor. This article tests the stability of the entire marine physical detection sensor, simulates the actual environment to measure temperature and electrode information, compares and analyzes the experimentally measured data, and draws a response conclusion. Experimental test results show that during 1–30 min, the temperature measured by the sensor fluctuates between 4.70031 °C–4.69890 °C, and the resolution of the temperature detection module can reach at least five decimal places. It can be seen that the performance of the sensor is stable, and the measurement accuracy basically meets the requirements of use.

**Keywords:** Ocean Exploration · Seabed Observation Network · Sensor Stability · Artificial Intelligence

## 1 Introduction

Ocean development is inseparable from the development of ocean exploration technology [1, 2]. The marine industry is rising day by day, and the research on marine physical exploration is also receiving more and more attention [3, 4]. Among them, the ocean magnetotelluric method reflects the distribution of subsea media through electrical parameters, which can provide more valuable information [5, 6]. However, the marine environment is complex and changeable, and the seawater has great interference to electromagnetic signals, making it more difficult to obtain valuable signals [7, 8]. Therefore, it is of great significance to develop a high-performance marine physical detection sensor and complete its performance stability test.

Regarding the research of ocean exploration, many scholars at home and abroad have conducted multi-directional and in-depth discussions on it. For example, Liu CH uses

optical video image processing technology to intelligently identify and classify weak targets on the sea and non-ocean waves [9]; Bell K proposed a detection scheme that combines coarse and fine detection of ship targets [10]; Wagner combines radar digital signal processing with machine learning to realize an efficient algorithm for exploring marine targets [11]. It can be seen that since the development of ocean exploration technology, the scientific development of its related technologies has been concerned by the majority of researchers. Therefore, this article combined with artificial intelligence technology is of great significance to the subject of marine physical detection sensor stability research.

This article aims to study the stability of marine physical detection sensors based on artificial intelligence technology. This article first introduces the functional modules and software of the sensor, including the processing module, sensor module, positioning module and other hardware. Then the stability of the sensor is tested. The experimental test results verify that the sensor has stable performance, and the measurement accuracy basically meets the requirements of use.

## **2 Stability of Marine Physical Detection Sensors Based on Artificial Intelligence Technology**

### **2.1 Marine Physical Detection Sensor Hardware and Functional Modules Based on Artificial Intelligence Technology**

#### **(1) Processing module**

The processing module includes functions such as power management, distributed processing and storage. The power management optimizes the sensor node in design, extends its life cycle, and manages the power supply. The main controller of the processing module adopts the CC2531 processing chip, which has the characteristics of low power consumption. Distributed processing is the processing of data. When the collected data suddenly becomes larger or smaller, the data will be collected many times by itself, and then the average value will be taken. The function of storage is to save certain parameters and key data in the sensor node.

#### **(2) Sensor module**

The sensors in this study, whether they are analog sensors or digital sensors, can be connected to the node through a common interface. The design of the universal port is to design the analog interface, serial digital interface and parallel digital interface into a universal module in the interface [12].

#### **1) Positioning module**

In marine physical exploration, the sensor will drift due to the interference of weather and natural factors. Therefore, the position of the sensor node must be known in advance during deployment, that is, the relative position of the sensor node is entered into the node or the ID number record of the node is translated into a valid position. This can effectively reduce the power consumption generated by the sensor node's own positioning, reduce the cost, and make the positioning accuracy more accurate. Therefore, in this module, the GPS/mobile base station is omitted.

## 2) Power management

The power supply mode of the sensor node is: a rechargeable lithium battery is combined with a solar battery. When the solar energy is sufficient, it will supply power to the sensor node and charge the lithium battery at the same time. When solar power is insufficient, it is powered by a rechargeable battery. The power management in the processor reads the energy of the power module at intervals to determine the power status. This design method can effectively solve the problem of energy limitation, greatly extend the life cycle, reduce the frequency of manual maintenance, and save labor.

## 3) Serial communication circuit

Using RS485 transmission technology, the data is output from the serial port and transmitted to the deck via the ready-made RS232 to RS485 module, and then sent to the client software via the RS485 to RS232 module.

## (3) Scheme design of temperature detection module

This design uses two-wire molybdenum resistors to form a Wheatstone bridge circuit. The voltage reference chip AD780 of this system is used as the power supply of the bridge, so that the current through the uranium resistor is not more than 1mA. The ADS1256 analog-to-digital converter is selected, which has the advantages of low noise, high resolution, high performance, high precision bits with built-in gain, and perfect self-correction and system correction functions.

## (4) Scheme design of in-depth inspection module

The rated power supply current of the selected pressure sensor is 0.5–2 mA, and the constant current of the designed constant current source is a typical value of 1.5 mA. Because the pressure sensor has a built-in Wheatstone bridge, no additional circuit is required. The analog-to-digital converter is also used. It is a high-precision 24-bit A/D analog-to-digital converter ADS1256.

## (5) Magnetic strength detection module

The magnetic strength detection module mainly includes a constant voltage source circuit, a magnetoresistive sensor, and an A/D mode converter with built-in gain. The constant voltage source module and the analog-to-digital conversion circuit all use the same device model in the temperature and depth detection module. The resistance sensor is a three-axis magnetoresistive sensor, which can measure the parameters of geomagnetic intensity.

## 2.2 Software Design

## (1) Data storage

When data is stored, the experiment-related information and sensor voltage signal data transmitted from the acquisition module are respectively saved. The file format of the data stored in the SD card is TXT, and each file is named by time. Each time the file is opened for writing, the address pointer stays at the end of the last written file. Each write operation sequentially writes a 150-bit array, storing 60-bit GPS information, 48-bit attitude sensor data, 36-bit AD conversion data and 6 newline escape characters.

- (2) **Data browsing**  
 The main function of this module is to read the files saved by the data storage module. Through the channel selection button, the data of different channels can be displayed in the waveform graph. At the same time, two cursors are set in the waveform graph. By dragging the two cursors to select the head value and the tail value as the selected interval, the waveform in the specified interval can be cut out to enlarge the display, and the relevant information of the waveform can be obtained, such as maximum, minimum, mean and variance.
- (3) **Magnetic sensor data processing**  
 The magnetic sensor has a zero adjustment resistance in each measurement direction. After power on, the AD differential channel is connected to the direct differential output of the magnetic sensor to set the zero adjustment, that is, it is considered that the magnetic field strength is zero in the current state, and the magnetic substance appears the change in the magnetic field caused by time is the measured value. Due to the working environment of the buoy, the energy detector is used to determine the threshold change for the Z axis and XY axis, and the change is recorded as 1, and the corresponding azimuth angle is calculated. At the same time, the vector sum is calculated to eliminate the energy change caused by its own rotation. The magnetic sensor data processing process occurs after the data is transmitted via the wireless transceiver module.
- (4) **Vector hydrophone data processing**  
 The working principle of the vector hydrophone is that when a sound wave reaches the sensitive structure, the plastic cilia cylinder will vibrate and resonate with the acoustic signal, and the cilia cylinder will slightly swing in four directions, which in turn drives the cantilever beam below to deform. Due to the piezoelectric effect, the resistance value on the cantilever beam will also change, so that the underwater acoustic signal is converted into an electrical signal.

This study selects the power spectrum analysis method to process the hydrophone data. The process is as follows:

First perform Fourier transform on the signal, then square the modulus of the obtained amplitude spectrum, and then divide by the duration to estimate the power spectrum of the signal, as shown in formulas (1) and (2):

$$X(e^{j\omega}) = \sum_{n=0}^{N-1} x(n)e^{-j\omega n} \quad (1)$$

$$P(\omega) = \frac{1}{N} |X(e^{j\omega})|^2 \quad (2)$$

In the formula,  $x(n)$  represents the time domain vector of the signal,  $X(e^{j\omega})$  represents the frequency domain of the signal, and  $P(\omega)$  represents the power spectrum.

### 3 Experimental Research Design

#### 3.1 Experimental Equipment and Environment

Signal acquisition instrument: DAQ2010 multifunctional data acquisition card; Experimental environment: refrigerator.

### 3.2 Experimental Project

(1) Experiment 1: Electrode sensitivity test

In a normal temperature environment, keep the distance between the two electrodes constant, and pass signals with amplitudes of 5 mV and 30 mv into the water tank with a frequency of 1 Hz. Use Ag/AgCl electrodes to detect this signal and display the output result on an oscilloscope. This experiment uses two eDAQ potentiostats, one as a signal generator and the other as an oscilloscope.

(2) Experiment 2: Temperature measurement and debugging experiment

In the refrigerator constant temperature experiment environment, place the container full of water in the refrigerator for a whole day. After the water temperature is consistent with the temperature in the refrigerator, put the hardware circuit in the refrigerator, and then put the sensor in the water for temperature measurement experiment. After 30 min, take out the hardware circuit and record the collected data.

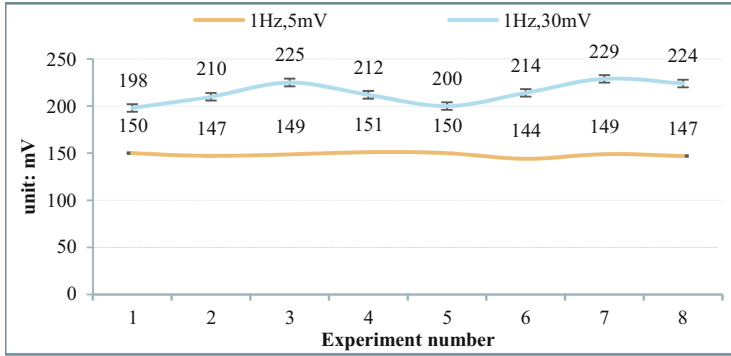
## 4 Analysis of Experimental Results

### 4.1 Sensor Electrode Stability

Experiment 1 was performed 8 times. Table 1 shows the results of the signals detected by the electrodes in different signal amplitudes. It can be seen that the signals detected by the electrodes are relatively stable in the environment of 1 Hz and 5 mV; in the environment of 1 Hz, 30 mV, the signal detected by the electrode fluctuates up and down.

**Table 1.** Electrode detection signal

Experiment number	1 Hz, 5 mV	1 Hz, 30 mV
1	150	198
2	147	210
3	149	225
4	151	212
5	150	200
6	144	214
7	149	229
8	147	224



**Fig. 1.** Electrode detection signal

It can be found from Fig. 1 that due to the limitation of the test instrument, the signal generated by the signal generator is interfered by the power frequency signal. Under the environment of 1 Hz and 30 mV, the signal detected by the electrode has a DC drift of 200 mV. However, no matter the amplitude or phase, the detection signal has no distortion. Therefore, the sensor electrode proposed in this study can sensitively detect the strength of the signal, and can be used to detect the abundant electric field signals existing in seawater.

#### 4.2 Data Analysis of Temperature Measurement and Debugging

In the second experiment, the unreasonable data generated by manual operation was eliminated, and the results are shown in Table 2: during 1–30 min, the temperature measured by the sensor fluctuates between 4.70031 °C and 4.69890 °C. The temperature data in Table 2 shows that the resolution of the temperature detection module can reach at least five decimal places, which meets the design requirements of the system.

**Table 2.** Temperature data measured within 30 min (°C)

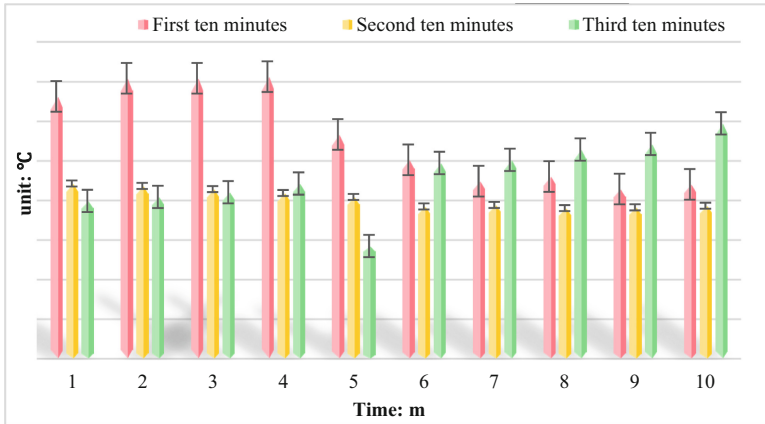
time	temperature	time	temperature	time	temperature
1	4.70031	11	4.69921	21	4.69899
2	4.70054	12	4.69918	22	4.69904
3	4.70054	13	4.69914	23	4.69910
4	4.70056	14	4.69909	24	4.69921
5	4.69983	15	4.69904	25	4.69842
6	4.69951	16	4.69892	26	4.69947
7	4.69924	17	4.69894	27	4.69951

(continued)



**Table 2.** (continued)

time	temperature	time	temperature	time	temperature
8	4.69930	18	4.69890	28	4.69964
9	4.69914	19	4.69891	29	4.69971
10	4.69920	20	4.69893	30	4.69997



**Fig. 2.** Temperature data measured within 30 min (°C)

It can be seen from Fig. 2 that in the first ten minutes, the temperature data collected changes less than the curve, and the curve changes in a wave shape. This is affected by the working mechanism of the refrigerator, and its work is intermittent. When the temperature reaches the set value, the refrigerator stops cooling. After that, the temperature in the refrigerator rises until the refrigeration work restarts. At the same time, due to the larger specific heat capacity of water, the temperature change of the water body is smaller.

## 5 Conclusion

With the development of artificial intelligence technology, sensor monitoring technology combined with artificial intelligence technology is a new driving force for the development of this field. Due to the unique characteristics of the marine environment, it is necessary to study marine physical detection sensors based on artificial intelligence technology suitable for the marine environment. Through research, this paper has completed the following tasks: introduced the functional modules and software system of the sensor, analyzed the stored data of each sensor; tested the stability of the entire ocean physical detection sensor, and verified the stable performance of the sensor. The characteristics and measurement accuracy also meet the requirements of use.

## References

1. Glover, A.G.: Abyssal fauna of the UK-1 polymetallic nodule exploration area, Clarion-Clipperton Zone, central Pacific Ocean: Cnidaria. *Biodivers. Data J.* **4**(4), e9277 (2016)
2. German, C.R., Petersen, S., Hannington, M.D.: Hydrothermal exploration of mid-ocean ridges: where might the largest sulfide deposits be forming? *Chem. Geol.* **420**(1), 114–126 (2016)
3. Wilson, W.H., Gilg, I.C., Moniruzzaman, M., et al.: Genomic exploration of individual giant ocean viruses. *ISME J.* **11**(8), 1736 (2017)
4. Lunine, J.I.: Ocean worlds exploration. *Acta Astronautica* **131**, 123–130 (2016)
5. Ballard, R.D., Leonardi, A.P.: New frontiers in ocean exploration the E/V nautilus and NOAA ship Okeanos explorer 2015 field season epilogue. *Oceanography* **29**(Suppl. 1), 76–77 (2016)
6. Pawlenko, N.: Technology and ocean exploration. *Oceanography* **32**(Suppl. 1), 100–100 (2019)
7. Valette-Silver, N., Cantelas, F., Beaverson, C., et al.: Sponsored projects: NOAA's office of ocean exploration and research introduction. *Oceanography* **32**(Suppl. 1), 118–118 (2019)
8. Gaffney, P.: Ocean exploration: a supply-demand mismatch. *Mar. Technol. Soc. J.* **50**(6), 8–9 (2016)
9. Liu, C.H., Huang, X., Xie, T.N., et al.: Exploration of cultivable fungal communities in deep coal-bearing sediments from 1.3 to 2.5 km below the ocean floor. *Environ. Microbiol.* **19**(2), 803–818 (2017)
10. Bell, K., Copeland, A., Chow, J.S., et al.: All hands on deck: the 2018 national ocean exploration forum. *Oceanography* **32**(Suppl. 1), 4–5 (2019)
11. Wagner, K.: Ocean exploration celebration. *Oceanography* **31**(Suppl. 1), 52 (2018)
12. Voss, J., Pomponi, S.: Cooperative institute for ocean exploration, research, and technology. *Oceanography* **32**(Suppl. 1), 119–119 (2019)



# Optimization System of Microbial Test on Account of Genetic Algorithm

Mingming Shao<sup>(✉)</sup>

Xi'an Medical University, Xi'an, Shaanxi, China  
smmxayxy@yeah.net

**Abstract.** Microorganism is an important part of geochemical cycle and plays an irreplaceable role in ecosystem. Optimization of microbial assay is very important. In this paper, genetic algorithm is used to optimize the microbial test. Using the operating mechanism of genetic algorithm, that is, imitating the basic laws of nature, carrying out natural selection and survival of the fittest, using this principle to treat the detection of microbial detection optimization. Through the natural selection and survival of the fittest, genetic algorithm weight adjustment, so as to achieve more accuracy of the test. By referring to the mathematical formulas (1) and (2) in Part 3 of this paper, the requirements of determining the definition of microbial detection can be achieved by initializing the population of microorganisms and analyzing the global convergence of the samples that meet the standards and do not meet the standards. This paper studies the knowledge of microbial test optimization system based on genetic algorithm, and describes the methods and principles of microbial test. The results show that the optimization effect of microbial test is improved significantly by the optimization system based on genetic algorithm.

**Keywords:** Genetic Algorithm · Microbial Test · Test Optimization · Optimization System

## 1 Introduction

Because microorganisms are ubiquitous in the air, land and water, the inspection results of microorganisms not only represent the quality of the product itself, but also reflect the sanitary conditions of the product processing environment, the health of the processing personnel, the safety of the product transportation and the reasonable conditions of storage. In addition, microorganisms are highly adaptable and easily mutated, and sometimes the mutated individual will have biochemical reaction characteristics completely different from the original individual, so the qualitative test results are also very important. The optimization system of microbial test based on genetic algorithm is beneficial to the optimization treatment of microbial test.

As for the research of genetic algorithm, many scholars at home and abroad have studied it. In foreign studies, Ortiz S proposed a genetic algorithm. Compared with existing path planning methods, the proposed path planning method has many advantages,

combining sliding mode control with classical simultaneous localization and mapping (SLAM) method. This combination can overcome the bounded uncertainty problem in SLAM [1]. Et. Proposed a new fine-grained sentiment analysis model combining convolutional neural network and random forest classifier. The continuous Word bag (CBOW) model is used for vectorizing text input. The most important features are extracted by convolutional neural network (CNN). The extracted features are used for emotion classification by random forest (RF) classifier [2]. Al-obaidi MA proposed an optimization framework based on species conservation genetic algorithm (SCGA) to optimize process design and operational parameters. In order to enable readers to have a deeper understanding of the process, the effects of membrane design parameters on xylenol retention rate, water recovery rate and specific energy consumption level under two different process conditions were studied [3].

In today's society, computer technology continues to develop rapidly, and computers have become necessary equipment for every scientific research institution. Many product inspection items are more or less began to introduce computer software to assist or replace manual operation [4, 5]. However, manual operation is still used in the microbiological testing using the microbiological testing methods formulated by the Ministry of Health. The identification results are obtained through multi-step operation and the judgment results are compared with the manual standards. Of so result issue often time is long, return easy occurrence error, because this artificial judgement result begins to be challenged greatly. The optimization system of microbial test based on genetic algorithm promotes the efficiency and accuracy of microbial test.

## 2 Design and Exploration of Microbial Test Optimization System on Account of Genetic Algorithm

### 2.1 Genetic Algorithm

Genetic algorithm is one of the important algorithms, whose basic principle is to imitate the basic laws of nature, natural selection and survival of the fittest, and optimize the algorithm through these two laws [6, 7]. Genetic algorithm has good application in many fields, but the application requirement is very low. It is a very effective global optimization algorithm with good adaptability when solving fuzziness of data.

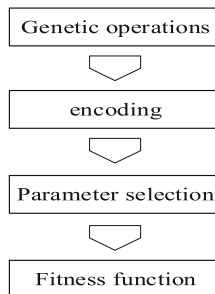


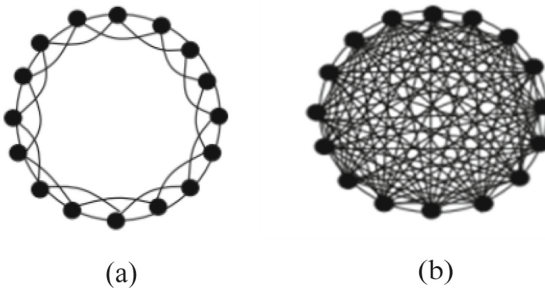
Fig. 1. The basic building blocks of genetic algorithms

The basic elements include genetic operation, coding mode, parameter selection and fitness function [8, 9]. The basic solution process of this algorithm is as follows, refer to Fig. 1:

- 1) Genetic algorithm firstly transcodes data and performs binary codes, which are like gene fragments and constitute elements of genetic algorithm [10, 11]. These elements are grouped into algorithmic populations by certain rules, and the process is like survival of the fittest.
- 2) Perform genetic manipulation on these gene fragments
- 3) To deal with the individual, using the selection strategy;
- 4) After algorithm iteration, a population will be initialized, and the global optimal solution is formed at this time.

When genetic algorithm performs iterative optimization, it must be modeled, which is a complex network model. In this paper, multiple analysis strategies are applied to the complex network model so that the inhomogeneity can be defined more accurately.

Two classical network topology models were widely used in the early stage of complex network research, as follows:



**Fig. 2.** Classic network topology model

- 1) As shown in Fig. 2(a), in the network model, each node only establishes edge relation with its adjacent nodes, and each node has the same number of edge.
- 2) As shown in Fig. 2(b), in the network model, any two nodes have established edge relations, so this model is conducive to information exchange between nodes.

## 2.2 Optimization System of Microbial Test Based on Genetic Algorithm

Microbiological test optimization system, first of all, to conduct microbial test results analysis [12, 13]. The management system needs to meet the laboratory requirements based on microbial testing methods.

(1) User demand analysis

The analysis and management system of microbial test results is mainly managed by the sample receiver or adoption personnel, the inspection personnel manage the sample test results, and other authorized personnel manage the test conclusions [14, 15].

(2) Functional requirement analysis

Microbial inspection conclusion need to analyze the test results can be, usually a sample need many steps of operation, and each step will get a result step by step, finally according to these results comprehensive analysis to determine the final conclusion step by step, so the microbial inspection conclusion analysis of the workload is bigger, the staff to come to the conclusion that often requires repeated comparison standard Therefore, it is very important to make an analysis system of microbial test results to liberate labor force.

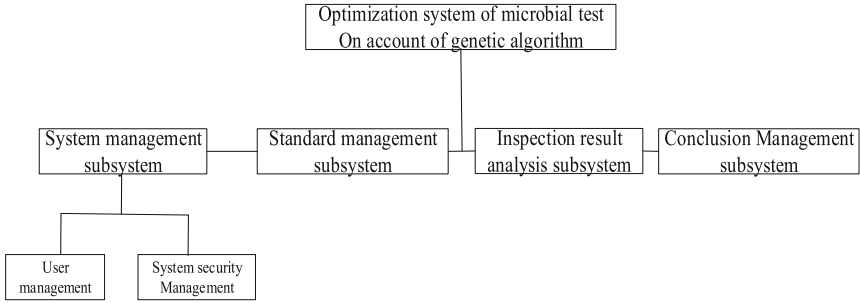
(3) Feature requirement analysis

A system with dual functions of analysis and management of inspection results is required. Especially for arbitration inspection institutions, it is very important to issue inspection conclusions quickly and accurately [16]. In addition, the data confidentiality of the inspection conclusion is very high, and the arbitration inspection generally requires more than 3 years to keep files. Once the data is leaked, it sometimes not only damages the interests of the prosecution, but even causes social chaos. For enterprises, microbial test data not only reflect the quality of products themselves, but also reflect the environmental quality of factories and warehouses, so it belongs to the category of trade secrets. Therefore, the data security of microbial test results analysis and management system is better.

Microorganisms need to be tested before the test results can be obtained, and only after the test results are analyzed and judged can conclusions be obtained. The whole inspection process, result determination and data management must be carried out in accordance with laboratory regulations. According to the business process, the central laboratory adopts the management mode of sampling and separation. After the sample is registered and processed in the sample room, the sample taker takes it to the laboratory. After the samples are tested in the laboratory, the inspection personnel shall draw the inspection conclusion and form the inspection report, which shall be approved by the technical director, the laboratory director and the center leaders in turn and then form an official document to be sent to the person/institution being tested.

### **3 Exploring the Effect of Optimization System for Microbial Testing on Account of Genetic Algorithm**

According to the analysis of functional requirements, the system is divided into five subsystems: system management, standard management, inspection results analysis and inspection conclusion management, as shown in Fig. 3.



**Fig. 3.** Microbial test optimization system based on genetic algorithm

(1) System management subsystem

The system management subsystem mainly manages the user information and the database of the system to ensure the security of data. The system management subsystem mainly includes two functional modules: user management and system security management.

① User Management

User information is designed to add, delete, modify, permission Settings, in order to unified management of users. The user information is unified input by the system administrator, who takes the real name, sets the login password, sets the permissions according to the department, and the permissions are set according to the four subsystems.

② System security management

The backup and recovery function of the existing database is designed to prevent the loss and damage of the database caused by human and non-human factors.

(2) Standard management subsystem

The standard management subsystem is mainly used to input product standards and provide judgment basis for analyzing test results. To ensure the authenticity and effectiveness of the input standard, it is operated by the inspection personnel. Standard management has designed the input, modification and deletion of standard information. Standard information mainly includes: standard name, standard number, internal control code, product category, standard value, release time, implementation time, status.

(3) Test result analysis subsystem

The inspection result analysis subsystem is mainly to analyze and judge the results obtained from the inspection of the sample input by the sample management system. Process: according to the management requirements of our center, the inspection personnel input the inspection results, and other personnel have no right to carry out this operation.

(4) Test conclusion Management subsystem

The inspection conclusion management subsystem mainly realizes the issue of inspection reports to the inspection results after analysis, and queries and prints

related reports according to different requirements. At present, according to the common inquiry methods of our center, it can be divided into: inquiry by product category, inquiry by task source, inquiry by inspection item and inquiry by inspection conclusion. The report is printed in a fixed format according to the query mode.

$z \in x$  is the population with a size of  $N$ , and the population fitness can be calculated according to Eq. (1):

$$f(x) = \max_{z \in x} \{f(z)\} \quad (1)$$

For any initial population  $B(0)$ , if

$$\lim_{t \rightarrow \infty} P\{f(B(t)) = y_h\} = 1 \quad (2)$$

It indicates that the algorithm has global convergence, where  $P(*)$  represents the probability of occurrence of event  $*$ ,  $t$  represents the  $t$ -generation population, and  $y_h$  means that the set is divided into  $H$  subsets.

#### 4 Investigation and Analysis of Optimization System for Microbial Testing on Account of Genetic Algorithm

This software chooses Windows XP as the design platform, uses Visual Basic 6.0 as the design language, and uses Microsoft SQL Server 2000 as the database.

Test method:

Database arbitrary CRUD operations and execution of their respective SQL queries. The database CRUD operation refers to:

C: Create: Creates a user.

R: Retrieve – Performs the retrieve view operation.

U: Update – Updates database information.

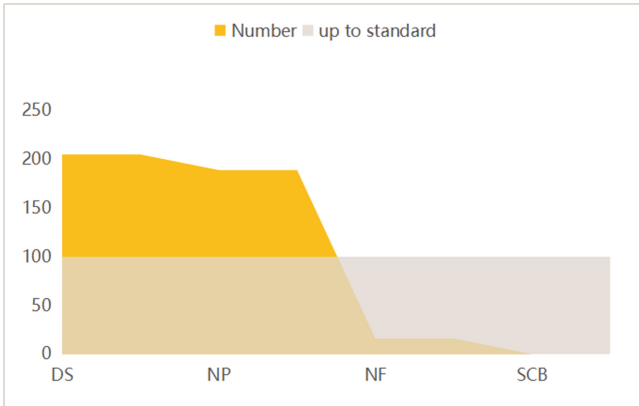
D: Delete: Deletes the database.

Test results:

Database main code is not empty; The outer code is equal to the corresponding main code or is null; The construction of data type, length and index is reasonable to meet the requirements of data and database integrity. All access methods and processes can operate as designed without data damage.

As shown in Fig. 4, the system test data is displayed in the system test checklist. The first row of The table contains Number and up to Standard, and The first column contains The test cases. The test case contains four kinds of data, namely, Design Total number of test case sets (DS), The Number of test case sets passed completely (NP), Number of failed test case sets (NF) and Set of test cases to be tested (SCB). The table means that the Design total number of test case sets (DS) is 205 times, and the number of test case sets passed (NP) is 189 times. The Number of failed test case sets (NF) is 16, and the Set of test cases to be tested (SCB) is 0.





**Fig. 4.** Test quantity chart

As shown in Fig. 4, In the figure, the Design Total number of test case sets (DS) and the number of fully passed test case sets (NP) were 205 times and 189 times, respectively, far exceeding the up to standard line. The test results show that the optimization system of microbial test based on genetic algorithm is very effective.

The data show that the optimization system of microbial test based on genetic algorithm has high performance in the optimization of microbial test.

## 5 Conclusions

This article through to the microbiological determine trival, fees, only a few expensive analysis instrument science problems were discussed, think development suitable for microbial detection methods formulated by the ministry of health of microbial test results analysis and management system to manage digital analysis and test results is necessary. This can not only reduce the working pressure of inspectors, but also can systematically manage the sample information and test results, but also can quickly, accurately and selectively query the test conclusion, can accelerate the pace of office automation center to a certain extent. The optimization system of microbial test based on genetic algorithm is beneficial to improve the efficiency and quality of microbial test.

**Acknowledgments.** 2021 Shaanxi Undergraduate and Higher Continuing Education Teaching Reform Research Project (Project No. 21BY137); 2021 University-Industry Collaborative Education Program of Ministry of Education (Project No. 202102650012); 2021 Annual General Project of Shaanxi Provincial Education Science “14th Five-Year Plan” (Project Approval No. SGH21Y0262).

## References

1. Ortiz, S., Yu, W.: Autonomous navigation in unknown environment using sliding mode SLAM and genetic algorithm. *Intell. Robot.* **1**(2), 131–150 (2021)

2. Siji George, C.G.: Genetic algorithm based hybrid model of convolutional neural network and random forest classifier for sentiment classification. *Turk. J. Comput. Math. Educ. (TURCOMAT)* **12**(2), 3216–3223 (2021)
3. Al-Obaidi, M.A., Ruiz-García, A., Hassan, G., et al.: Model based simulation and genetic algorithm based optimisation of spiral wound membrane RO process for improved dimethylphenol rejection from wastewater. *Membranes* **11**(8), 595–595 (2021)
4. Dolezel, P., Holik, F., Merta, J., et al.: Optimization of a depiction procedure for an artificial intelligence-based network protection system using a genetic algorithm. *Appl. Sci.* **11**(5), 2012 (2021)
5. Güler, S., Yenikaya, S.: Analysis of shielding effectiveness by optimizing aperture dimensions of a rectangular enclosure with genetic algorithm. *Turk. J. Electr. Eng. Comput. Sci.* **29**(2), 1015–1028 (2021)
6. Nayana: Electric vehicle charging with battery scheduling and multicriteria optimization using genetic algorithm. *J. Electr. Eng. Autom.* **2**(3), 123–128 (2021)
7. Khosravian, P., Emadi, S., Mirjalily, G., et al.: QoS-aware service composition based on context-free grammar and skyline in service function chaining using genetic algorithm. *PeerJ Comput. Sci.* **7**(4), e603–e603 (2021)
8. Davoudi, K., Thulasiraman, P.: Evolving convolutional neural network parameters through the genetic algorithm for the breast cancer classification problem. *Simul. Trans. Soc. Model. Simul. Int.* **97**(8), 511–527 (2021)
9. Batayneh, W., Bataineh, A., Jaradat, M.A.: Intelligent adaptive fuzzy logic genetic algorithm controller for anti-lock braking system. *Int. Rev. Model. Simul.* **14**(1), 44 (2021)
10. Ousmane, B., Moustapha, D., Adama, C.: Genetic algorithm for the pick-up and delivery problem with time window by multi-compartment vehicles. *Transp. Telecommun. J.* **22**(3), 343–352 (2021). Abidjan,cte d’ivoire eE cole supérieure africaine des TIC,Abidjan,Cote d’Ivoire eU niversité Felix Houphout Boigny, Abidjan,Cote d’Ivoire
11. Dayal, S., In, R., Gruplarnn, R., et al.: Genetic algorithm-based approach for constructing product groups in make-to-order production environments. *Yönetim ve Ekonomi Celal Bayar Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi* **20**(2), 259 (2021)
12. Ajmi, N., Helali, A., Lorenz, P., et al.: MWCSGA—multi weight chicken swarm based genetic algorithm for energy efficient clustered wireless sensor network. *Sensors* **21**(3), 791 (2021)
13. Indian, A., Bhatia, K.: An approach to recognize handwritten Hindi characters using substantial Zernike moments with genetic algorithm. *Int. J. Comput. Vis. Image Process.* **11**(2), 66–81 (2021)
14. Vadivoo, N.S., Usha, B., Sudha, K.: A comprehensive review of current microbiological detection methods of SARS CoV-2. *Indian J. Microbiol. Res.* **7**(3), 113–124 (2020)
15. Pal, M., Bulcha, M.R., Banu, M.G., et al.: Emerging role of biosensors for detection of foodborne pathogens. *Am. J. Microbiol. Res.* **9**(3), 92–95 (2021)
16. Widodo, E., Pranibilan, A., Ardilla, Y., et al.: Effect on encapsulated liquid smoke in combination with formic acid on intestinal development and microbial counts in broiler. *IOP Conf. Ser. Earth Environ. Sci.* **788**(1), 012185 (2021)



# The Application of Virtual Reality Technology in Ophthalmology

Jingying Wang<sup>(✉)</sup> and Yi Zhang

Chongqing Medical and Pharmaceutical College, Chongqing, China  
cqygz@yeah.net

**Abstract.** Virtual reality (VR) is changing the way we perceive and interact with various digital information, so that many scenes can place users in an ideal visual sensory environment through head mounted devices. In recent years, the research of VR in the field of ophthalmology is mainly reflected in clinical application and teaching. This paper analyzes the relevant research in the fields of vision training and amblyopia treatment, myopia prevention and control, eye adjustment and convergence function, strabismus diagnosis, ophthalmic surgery assistance and ophthalmic teaching, and introduces the application of VR technology in the field of Ophthalmology.

**Keywords:** Virtual reality technology · Ophthalmology · Medical treatment · Amblyopia · Myopia

## 1 Introduction

With the continuous development and progress of computer hardware and software, computer human-computer interaction interface technology based on computer image, multimedia and multi-sensor has also developed rapidly. VR and augmented reality (AR) technology are more and more applied and studied in the fields of entertainment, medical treatment, education and so on. With the innovation of technology, VR equipment has been paid more and more attention in the clinical and teaching fields of Ophthalmology.

## 2 Overview of VR and Ophthalmology

VR uses computer simulation to generate a virtual world in three-dimensional space, which provides users with visual and other sensory simulation, so that users seem to experience their environment and can observe things in three-dimensional space in real time and without restrictions. When the user moves the position, the computer can immediately carry out complex calculation and transmit the accurate three-dimensional world image back to make the user feel telepresence [1], as shown in Fig. 1

The three-dimensional display of VR is based on the principle of binocular parallax and realized by means of head mounted display equipment. From the perspective of technology, VR system has three basic characteristics: Immersion interaction conception.

The details are as follows: (1) immersion: it means that the user is in a completely virtual environment from the first perspective, rather than watching from the third perspective like other 3D display devices, so he has a stronger sense of scene; (2) Interactivity: it means that users can interact with the built virtual environment, such as performing surgical operations on a virtual human body; (3) Imaginative: with the help of those conventional unreachable or abstract scenes, users can be in any environment, so as to expand their vision and imagination. For example, viewing the anatomical structure of the eyeball from the inside of the eyeball, following the atrial flow to feel the aqueous circulation, and displaying the complex visual path from different angles [2].

VR technology can use head dynamic instrument, eye vision sensor, hand touch sensor and so on to generate feedback of simulated operation information in virtual space in real time, to improve the user's experience of the reality of three-dimensional space.



**Fig. 1.** VR and ophthalmology

### **3 Application of VR in Ophthalmic Clinical Field**

#### **3.1 Research in the Field of Myopia**

Previous studies found that watching VR stereo video can simulate far and near vision activities, so as to train ciliary muscle function and relieve ciliary muscle spasm, so as to alleviate visual fatigue and the progress of myopia. Ha et al. [3] found that wearing VR equipment for 30 min will briefly lead to the progression of myopia, but this effect can be completely recovered after 40 min. In recent years, it has been reported that VR equipment may be used to control myopia. Turnbull and Phillips [4] found that the diopter and binocular visual function (such as stereopsis and adjustment amplitude) of human eyes have no significant change after wearing VR equipment, while the choroidal thickness of human eyes will become thicker. Choroidal thickening may be related to myopic defocus, so it may delay the development of myopia. From the findings of basic research, in the animal model of myopia, hyperopia defocus can accelerate

the development of myopia, and myopia defocus can slow down the development of myopia. At the same time, increasing outdoor activities can slow down the occurrence of myopia. For example, children can effectively slow down the occurrence of myopia by moving for 3 h under the light intensity of  $>10000$  illuminance every day. For the reasons why outdoor sports can alleviate myopia, there are two aspects recognized internationally: one is that high-intensity light promotes dopamine secretion and then delays the development of myopia; Second, because high-intensity light can induce the pupil to shrink, and then increase the depth of field, so as to improve the visual blur, so as to delay the emergence of myopia. Therefore, in the next step, it is possible to better control the defocus of the surrounding retina through eye tracking technology, fixation point rendering technology and focal plane display technology in VR equipment, and then combined with VR equipment to control the brightness and spectral components, so as to control the progress of myopia, as shown in Fig. 2.



**Fig. 2.** Application of VR in myopia

### **3.2 Research in the Field of Eye Regulation and Convergence (Divergence) Function**

At present, many scholars have proposed that when wearing VR equipment, the inconsistency between accommodation and convergence may lead to functional eye diseases such as visual fatigue, dry eye, transient accommodative strabismus, video terminal syndrome and so on. Mohamed Elias et al. [5] wore VR glasses to 34 young people. They measured the binocular adjustment force and convergence and divergence function before wearing and 30 min after wearing. They found that the use of VR equipment will lead to the advance of eye adjustment. At the same time, the ratio of accommodation convergence/accommodation ( $AC/a$ ) will be reduced, and the binocular convergence and divergence function will be weakened. They also found that if VR equipment is used to see virtual close range for too long, it will lead to mild exotropia, As shown in Fig. 3. Godinez et al. [6] compared and studied the different reactions of 20 young people (aged

18–24) to wearing VR equipment and traditional computer display. It was found that VR equipment would lead to the increase of Bo direction blur point (near and far vision) and the slight increase of accommodation amplitude in the examination of convergence and dispersion range, but the difference was not statistically significant. Yoon et al. [7] also found that after wearing VR equipment for 30 min, although the ocular diopter will not change, the never point of convergence (NPC) and never point of accommodation (NPA) will increase. Although the impact of VR use on visual function is not clear, the impact on human eye adjustment function and convergence and dispersion function after wearing VR equipment is still an important direction of its safety detection in the future.



**Fig. 3.** Role of Vr In Ocular Accommodation and Dispersion

### 3.3 Research in the Field of Strabismus

Compared with amblyopia, VR is rarely used in strabismus research. In 2018, Thomsen et al. [8] found that after 6 months of training for 25 patients with intermittent exotropia (5 adults and 20 children), their strabismus degree decreased or disappeared, stereopsis



**Fig. 4.** The role of VR in strabismus

was established, and there was no change in diopter degree. Miao et al. [9] found that VR equipment can better evaluate the degree of ocular strabismus, which is basically consistent with the diagnosis results of doctors and affirmed the accuracy and effectiveness of VR equipment by comparing the diagnosis of ocular strabismus of 17 different patients (5 orthosis and 12 exotropia) by VR equipment and doctors.

At the same time, the research of moon et al. [10] found that VR training can improve the clinical diagnosis skills of ophthalmologists for esotropia and exotropia in a short time, and affirmed the effectiveness and convenience of VR application, as shown in Fig. 4. Therefore, VR equipment is expected to be applied to the auxiliary diagnosis of strabismus in the future.

## **4 The Role of Visual Rehabilitation Training in Children with Visual Impairment**

Visual impairment includes blindness and amblyopia. In the early stage of children's visual development, active and correct amblyopia treatment will produce good results. However, children in this period have poor cognitive ability, so it is difficult to cooperate with and adhere to the traditional therapy with monotonous and long training cycle. Therefore, in recent years, ophthalmologists have tried to find a new VR treatment method that can not only stimulate children's interest in training, but also improve the treatment effect.

### **4.1 Amblyopia Treatment Based on VR**

VR technicians from the University of Nottingham and ophthalmologists from Queen's Medical Center have developed an interactive "binocular processing system" to provide interactive 3D games and videos for children with amblyopia. Research shows that the system can provide a relaxed and pleasant treatment method, which can enable children to obtain ideal curative effect in a short time. Chinese ophthalmologists have also made similar explorations and developed the "vision enhancement" system software, which integrates amblyopia treatment with virtual scenes, as shown in Fig. 5.

The system adopts a variety of stimulation modes. On the one hand, it improves the visual acuity of amblyopia and makes up for the shortcomings of traditional therapy; On the other hand, help children establish normal binocular visual function and promote their visual function and healthy development of body and mind. According to the clinical report of Mian Yao, the system software of "increasing visual energy" has the advantages of strong pertinence, easy operation, diversification, and children's willingness to accept. Its training effect is better than that of traditional therapy, especially for ametropic amblyopia and mild amblyopia. When using the system software to treat amblyopia children of different ages, it is pointed out that children need to have certain hand eye coordination ability because they need to control the mouse by hand in the training process, Therefore, too young children are not suitable for using the system. In another study, they emphasized the early detection and treatment of amblyopia, and suggested that qualified families use the software for training as soon as possible.



Fig. 5. Application of VR in amblyopia treatment

#### 4.2 Research on the Types of Visual Function Defects Based on VR

The traditional types of visual impairment are divided into three types: ametropic amblyopia, anisometropic amblyopia and strabismus amblyopia based on the examination of visual acuity chart, and are divided into three grades: mild, moderate, and severe. Using the “children’s vision and intelligent VR database system based on perceptual learning”, 323 children with amblyopia were examined for visual function defects. According to the types of visual information processing defects, amblyopia was divided into “low-level visual function defect”, “high-noise visual function defect” and “high-level visual function defect”, and a good distinction effect was obtained. When diagnosing amblyopia children, we should increase the evaluation of their visual status on the basis of measuring their visual acuity level with the traditional visual acuity chart, and take this as the basis for targeted treatment to repair their visual dysfunction, as shown in Fig. 6.



Fig. 6. VR is used for visual function adjustment



## 5 Research on VR in Ophthalmology Teaching

Using VR technology to build a simulation system of normal human eye anatomical structure, the created image has both three-dimensional and realistic feeling. At the same time, it can also rotate, zoom in, zoom in, zoom out, etc., which can more intuitively observe the internal structure of the eyeball. At the same time, the research of Jin et al. [11] also pointed out that using VR technology can build various three-dimensional scenes, and then simulate the symptoms and signs of various ophthalmic diseases, such as visual blur, visual object deformation, visual field defect, etc., which is helpful to assist the teaching of students' ophthalmology courses. As long-term use of VR equipment may lead to visual fatigue, improving VR technology and equipment to reduce students' visual fatigue after use is an important direction of development in the future.

In the future, VR technology has a wide application prospect in the teaching of simulated ophthalmic diseases. It can also be used to establish a standardized patient database for ophthalmic teaching and assessment. The eye Si (vrmagic, Germany) surgical simulator is most used in ophthalmic surgery teaching. This simulator can simulate three-dimensional images in surgery under the microscope, simulate and train cataract and vitreous surgery. It has the advantages of simple and controllable use, high degree of simulation and reverse operation. It can significantly improve the technical level of ophthalmic inpatients in cataract surgery, especially capsulorhexis and anti-shaking, and has a significant correlation with the actual operation. The surgical simulation system can also support the training of vitreous surgery. Through this system, users can carry out basic intraocular micromanipulation training, such as vitrectomy, intraocular laser, posterior vitrectomy, stripping of internal limiting membrane and so on. Through training, surgical skills can be improved to varying degrees, but whether it can be successfully converted to real patient surgery remains to be further studied.

## 6 Conclusions

With the rapid development of modern medicine science and technology, more and more medical technology achievements benefit mankind. However, in the field of ophthalmic medicine education, students have some problems in the process of learning ophthalmic medicine, such as boring theoretical knowledge, shortage of experimental sites, unsatisfactory practical operation and so on. From the perspective of Ophthalmology, this paper studies the application of VR in the fields of vision training and amblyopia treatment, myopia prevention and control, eye regulation and convergence function, strabismus diagnosis, ophthalmic surgery assistance and ophthalmic teaching. At the same time, it expounds the advantages of VR technology in physics and cognition and puts forward suggestions on the application of VR technology in this field, so as to provide more references for the effective implementation of ophthalmic medical education.

**Acknowledgments.** School-level project: "Integration of socialist economic construction, political construction, cultural construction and social construction, integration of learning and innovation" The Construction of Optometric Talent Training Path of Exploration and Practice YGZCG1905; school-level key discipline project: optometric medicine ygz2021302.

## References

1. Mitrousia, V., Giotakos, O.: Virtual reality therapy in anxiety disorders. *Psychiatriki* **27**(4), 276–286 (2016)
2. Pratorius, M., Burgbacher, U., Valkov, D., et al.: Sensing thumb-to-finger taps for symbolic input in VR environments. *IEEE Comput. Graph. Appl.*, 20–24 (2019)
3. Ha, S.G., Na, K.H., Kweon, I.J., et al.: Effects of head-mounted display on the oculomotor system and refractive error in normal adolescents. *J. Pediatr. Ophthalmol. Strabismus* **53**(4), 238–245 (2016)
4. Turnbull, P., Phillips, J.R., Ocul, V.R.: Effects of virtual reality headset weVR in young adults. *Sci. Rep.* **7**(1), 16172 (2017)
5. Mohamed, E.Z., Batumalai, U.M., Azmi, A.: Virtual reality games on accommodation and convergence. *Appl. Ergon.* **81**, 102879 (2019)
6. Godinez, A., Harb, E.N., Grimes, J., et al.: Oculomotor changes after sustained Virtual Reality use. *Invest. Ophthalmol. Vis. Sci.* **60**(9), 5924 (2019)
7. Yoon, H.J., Kim, J., Park, S.W., et al.: Influence of virtual reality on visual parameters: immersive versus non-immersive mode. *BMC Ophthalmol.* **20**(1), 200 (2020)
8. Thomsen, A.S., Smith, P., Subhi, Y., et al.: High correlation between performance on a virtual-reality simulator and real-life cataract surgery. *Acta Ophthalmol.* **95**(3), 307–311 (2017)
9. Miao, Y., Jeon, J.Y., Park, G., et al.: Virtual reality-based measurement of ocular deviation in strabismus. *Comput. Methods Programs Biomed.* **185**, 105132 (2020)
10. Moon, H.S., Yoon, H.J., Park, S.W., et al.: Usefulness of virtual reality-based training to diagnose strabismus. *Sci. Rep.* **11**(1), 5891 (2021)
11. Jin, B., Ai, Z., Rasmussen, M.: Simulation of eye disease in virtual reality. In: *Conference Proceedings - IEEE Engineering in Medicine and Biology Society*, pp. 5128–5131 (2005)



# Research on the Application of BIM Technology in the Whole Process Cost Management of Construction Project

Kangyan Zeng and Zhen Wen<sup>(✉)</sup>

School of Civil Engineering, Chongqing College Architecture and Technology,  
Chongqing 401331, China  
Wendy198787@163.com

**Abstract.** Cost management, with the information technology used frequently, is the most important link in the process of engineering construction. It is, however, no longer able to adapt to the trend of information technology development by using traditional work methods. The whole process cost focuses on the whole construction process and the overall interests of the project. BIM technology, as an electronic information modeling, provides an efficient information exchange platform, on which the cost management work can be connected at any stages in series with the result of repetitive work reduced and work efficiency increased. The improvement of work efficiency and way can be realized by applying the suitable BIM software to each stage of cost management can improve work efficiency. The paper studies specifically the integration of the whole process cost management on construction engineering. There is a commercial residential project in the case part in the use of relevant BIM software to realize the systematic collaborative management for the whole process of engineering cost. On the basis of the transfer of the cost data to the whole process of construction by using the BIM software, it concludes the route and method of the whole process cost management in the use of relevant software.

**Keywords:** BIM Technology · Whole Process · Whole Process Cost Management

## 1 Introduction

With the increasingly development of the scale and the output value in recent years, the construction market has reached a high proportion of GDP in China, and even up to 12% in 2021 [1]. Nowadays, the development of new technologies is rapidly in all walks of life, so the construction technology has been updated and improved to a certain extent. The efficiency on the construction technology, nevertheless, is still at low level, because of the properties of the products and the trend of industry development. There are problems in the work, including every link in the construction process fails to transmit information efficiently, and participants of the construction work fail to cooperate with

each other, and the repetitive work is emerged at each stage. It will not only lead to low efficiency in the construction process, but also cause serious waste of resources [2].

BIM technology is an effective way to improve the informatization of the construction industry. It turns the construction process into a modern industrialized production model of an assembly line, in which BIM forms a digital production line with a powerful data in all related software and models through a shared work platform, upgrading the production process of construction products to a modern industrial-grade production model. At this stage, the cost management methods of construction engineering are neither able to adapt to the current trend of informatization development, nor to control construction investment effectively. The paper effectively integrates BIM technology into the cost management of all aspects of engineering construction based on a comprehensive and systematic analysis of the whole process engineering cost management and BIM technology [3].

## **2 Hole Process Cost Management of Construction Project**

The whole process cost management of construction engineering includes total cost management, whole lifecycle management, total factor cost management, and whole process cost management. The whole process cost management place emphasis on the work of the cost management which should be involved in advance. It runs through the whole range of construction engineering with the starting point of the construction engineering planning and decision-making stages to the end point of the completion [4].

### **2.1 Cost Management in Each Stage**

First of all, the investor of the construction project needs to take such factors as the amount of capital to be invested, opportunity cost, and actual technical level into consideration so as to make a decision on the scale and usage of the project at the investment decision-making stage. Project cost management staff should grasp the investment of the project as a whole and prepare investment estimation [5].

The design stage is the key point for the actual formation of the project cost and the most effective control. At this stage, the design unit often calculates the project cost according to its proposed construction scheme or construction drawings.

The project price calculated in the bidding stage is not only the basis for both parties to determine the contract price, but also the basis for settlement between both parties in the later stage [6].

The project construction stage is the central link in the whole project life cycle of the construction unit and the contractor. The contractor completes an actual building through the construction process, and most of the project investment will be spent at this stage. This stage is the formation process of the actual project cost [7].

The completion settlement of the project is the process in which both parties calculate all the completed construction products and pay the project price according to the specific provisions of the construction contract after the completion acceptance is qualified. The completion settlement received by the construction party is generally composed of the contract price plus or minus the adjustment amount recognized by Party A and Party B

as well as the deduction of project progress payment and quality warranty deposit paid during the construction process [8].

## 2.2 Difficulties in the Implementation of the Whole Process Cost

- (1) It is difficult to gather professionals to support the whole process cost management at the full stages of the construction process with the result of the small scale in the cost industry.
- (2) The key point of cost management focuses on the project pricing business. The quantity surveyors always take the measurement and valuation, rather than focusing on the value management of the entire life cycle of the construction project [9].
- (3) There is the situation with slow update speed for the measurement and valuation basis is not fast enough in the whole process of cost management. There is insufficient amount of similar engineering cost data in the estimation and budgetary estimation stage.
- (4) At present, the work mode of most quantity surveyors is to take the measurement and valuation in the use of the project cost software and quota data set by government, which is apt to fail to be in accordance with the actual situation [10].

Firstly, on the basis of the combination of the above-mentioned difficulties in project cost management, the quantity surveyors need to use electronic information tools to improve the accuracy and efficiency of engineering measurement; secondly, the information platform should be built for idea exchanges between different professional staff at each stage to record and analyze the whole process cost data; finally, a record carrier of engineering cost data which is practical and reliable is needed to analyze and store the engineering cost data in a structured manner [11].

## 3 Bim Technology is Applied in the Whole Process Cost to Do Bim Fusion Analysis

### 3.1 Investment Decision-Making Stage

The application of BIM includes initial modeling, model maintenance, cost estimation, etc. in the project planning stage. According to the existing data, the current 2D drawings are imported into software with BIM technology to build a 3D modeling. Generally speaking, it is the initial project modeling created the early stage. The investment estimation is taken in the use of the BIM technology with a powerful information statistics function based on this modeling. At this stage, relatively accurate engineering quantities can be obtained according to the model, and the further calculation can be taken on the installation costs of the building. At the same time, the project cost data can be used to weigh the pros and cons of different schemes, compare and optimize the schedule, so as to prepare and provide an important basis for project decision making.

### **3.2 Design Stage**

In the past, the drawings were made by different designers with different majors such as civil engineering, water and electricity, and fire-fighting pipelines. Conflicts and collisions as well as size deviation is easy to occurs between different majors and different views of the same major. The designers, auditors and other parties are unable to completely find and correct the unreasonable points, even if they spend a lot of energy to check and compare the drawings. These conflicts are manifested in the construction process, which has caused great uncertainties to the cost management, even quality and safety of the project, and result in an increase in costs. When it comes to the establishment of 3D models, the collaborative design of various professional designers, and the visual analysis of different professional components adopted in the process of the design, the conflicts caused by the drawings will be resolved in time, and the interactive check can reduce errors in the design.

The various dimensional information provided by the BIM model will also simplify the calculation of the engineering quantity in the design stage. It can be directly calculated for the engineering quantity in the use of the BIM model. The data of each component in the model is related to the calculation process of the engineering quantity. When the components in the model are changed, the engineering quantity will also be updated, so that the engineering cost data can be updated in real time. In the design stage, the cost personnel can use the BIM technology to greatly shorten the time for calculating the project quantity, realize the rapid and accurate preparation of the project estimate, and can also discover some conflicting problems that were only discovered during the construction in advance, and reduce the later engineering changes.

### **3.3 Tendering and Biding Stage**

For the tenderer, BIM can truly provide the engineering entity information required in the calculation of the engineering quantity to automate the calculation, improve the accuracy of the calculation, and allow the cost staff to change from repetitive calculation work to thinking and controlling the factors that affect the price of the project, a more scientific budget can be prepared. If bidders want to have their own bidding data, they need to introduce BIM to quickly calculate and fully store the consumption standards during the construction process. Through reuse or rapid establishment of 3D models, fast and accurate calculation of engineering quantities will no longer be a problem. In addition, the bidder can use the 3D design model to quickly locate the structural information of heavy and difficult areas, determine and adjust the construction plan according to the actual situation of the project, correctly evaluate the difficulty of the project, and make accurate quotations.

### **3.4 Construction Stage**

The key point of construction units on costs management is the management of the project construction costs. The cost targets are mostly compared with the unit price and amount of the signed contract. Generally speaking, the post-event analysis is been taken. There is a lack of cost control in the process. The construction schedule only contains the

size of the project and the completion time information, instead of changing the project plan and actual completion; most of the construction schedule of the project department is determined by the sophisticated construction management personnel. There will be deviations between the engineering quantity and the amount of labor, materials, and machinery resources calculated by the project manager and the actual value, with result of the increasing on the engineering cost in the actual construction process. It is usually more accurate for the construction schedule preparation completed by BIM technology. The time information is added to form a 4D model in the use of the 3D model of BIM. The resource consumption required by each construction process and construction node can be accurately calculated, and then cost information is added to form a 4D model. The 5D model of the project, using quota consumption data, etc., accurately calculates the number of labor and construction machinery required for each construction process. In order to prevent insufficient resource input, it is feasible to add progress information, cost information, and construction organization information into the 3D model to calculate the consumption of people, materials, and machines for the entire project.

### 3.5 Project Acceptance Stage

The settlement data of the project can be obtained by collecting and arranging the project information and data during the design and construction of the construction project. Using BIM technology to collect the information of the building in a complete and structured manner, the completion and settlement of the project can be quickly counted. It can quickly compare and calculate with the contract price, and finally form an accurate settlement price.

## 4 CASE

We take a commercial real estate project as an example. According to the whole process cost management process, we adopt the bill of quantities valuation method to calculate the engineering cost of the civil works. In the process, BIMMAKE, a kind of Glodon modeling software, is used to establish BIM model, Glodon GTJ is used to calculate engineering quantity, Glodon cloud pricing platform (GCCP) is used to calculate list unit price and project cost, and Glodon BIM5D is used to take construction simulation. At first, Glodon BIM software is used to build a 3D model in this project. The established model is imported into the Glodon GTJ2018 to calculate the construction project volume. Next, the Glodon pricing platform is used to apply the list quota, so as to, on the one hand, avoid the data loss caused by the REVIT model in the process import procedure or the errors caused by manual copying of CAD drawings, on the other hand, it is, in a large extend, to reduce the workload of the cost engineers for modeling.

Here are the specific work of in the use of BIM and related softwares to carry out the whole process cost management of the case project: at the planning stage, the total investment of the project should be estimated; at the design stage, it is mainly for the preparation of budget estimates; at the tendering and bidding stage, the bidding control price is prepared; in the mid-construction settlement, the engineering quantity calculation and engineering change control are carried out; in the completion settlement,

the engineering quantity is calculated and the claim management is carried out; and finally the cost data, the extraction and preservation of the cost target are completed.

In the investment decision-making stage, Glodon software is used for 3D modeling, on which the engineering volume is quickly calculated, and then the Glodon Index Network is used to query and check investment estimation indicators, which is quickly and accurately achieved on the investment estimation documents. This will greatly improve the accuracy of estimation and provide an accurate data source for subsequent cost management work.

At the design stage, the exact engineering quantities should be calculated on the basis of the designed construction drawings, and the cost of the project should be calculated in the use of the current bill of quantities valuation specifications and local quotas. At this stage, BIM software can be used to perform some direct conflict checks of various disciplines. For example, when we check the collision of the drainage pipes, it can be judged whether the engineering pipes collide with the frame beams; when it comes to every view, we would wonder and check that the discrepancies between the structural drawings and the architectural drawings. At the design stage, the mistakes in the drawings should be corrected as much as possible, so as to avoid the occurrence of rework and changes during the later construction, thereby avoiding the increase of the engineering cost.

In the bidding stage, the tenderer needs to use the BIM model to quickly calculate the quantities when preparing the cost documents. The software has built-in list specifications to form a complete bill of quantities; the calculation rules of list and quota have been set. There is no need for cost personnel to remember the calculation rules. The software will automatically deduct according to the drawing of component elements and use them at the same time. The quantities of two calculation rules can be obtained from the same model; the software provides multiple engineering quantity codes, which can be combined and extracted freely; we can use the Glodon cloud pricing platform to calculate the unit price of the bill of quantities, take the fee, summarize and calculate the bidding control price.

The bidder adopts the three-dimensional model provided by the tenderer to calculate the quantities faster and uses the pricing software to prepare the bidding price. The BIM model established by the bidder at this stage can be imported into Glodon BIM-5D software to prepare the schedule, and carry out engineering change cost, monthly settlement and quarterly settlement at the construction stage.

The cost management to be carried out in the construction stage includes change management, process payment management and progress management. The premise of using Glodon's change software to record the design change is to have a model file approved by both Party A and Party B to draw the changed components on the basis. It is convenient to use the change software to see the increase or decrease of the changed parts. At present, the quantities of the general list are settled according to the facts, or the changed parts can be drawn directly. The actual quantities can be counted in the progress settlement or completion settlement. Process payment management and progress management can be realized through Glodon BIM-5D. During settlement, Party A and Party B shall calculate the actual quantities on time according to the comprehensive unit price or price adjustment method signed in the unit price contract. The calculation of quantities



can use the three-dimensional model that has been used in the process to find differences. After modifying the components, you can update yourself and related quantities. With regard to the change of the comprehensive unit price in the list of quantities, the price of materials can be calculated by using Glodon Assistant to calculate the weighted average of the monthly information price or market price. The comprehensive unit price of the list can be calculated automatically by using the pricing software and then the settlement documents can be prepared according to the contract.

In this process, the same three-dimensional model has been used for data flow, which can reduce the modeling time of cost personnel in each stage and avoid data loss and error caused by repeated modeling.

## 5 Conclusions

This paper introduces BIM Technology to realize the whole process cost management of construction engineering, and mainly obtains the following research results:

- (1) Through qualitative analysis, it proves that BIM Technology has the characteristics of simulation and visualization, which can greatly improve the speed and accuracy of cost personnel in calculating quantities, shorten the time of calculating project cost and provide an effective and advanced working method for cost management.
- (2) Through quantitative analysis, the initial BIM model is established by using BIM make software, which transforms the traditional two-dimensional drawing into the three-dimensional physical drawing of what you see is what you get, strengthening the intuitiveness of the drawing and easy to understand and find design errors; the BIM calculation model is established by using Glodon GTJ2021, which realizes the rapid and accurate calculation of quantities and can correlate the design change with the calculation results of quantities in real time. After the change, the quantities can be calculated and counted quickly.
- (3) This paper analyzes the BIM software used in each stage of project construction and the use process, methods and important functions of the software, which provides practical experience for similar projects to use BIM Technology for cost management in the later stage. Using BIM Technology can greatly improve the efficiency and accuracy of cost management.

From the perspective of cost management in the whole process of construction engineering, the application of BIM should focus on the overall construction process rather than just considering a certain stage. BIM model should be continuously transferred to the whole process of cost management in order to achieve the best use effect.

## References

1. BorjeGhaleh, R.M., Sardroud, J.M.: Approaching industrialization of buildings and integrated construction using building information modeling. *Procedia Eng.* **164** (2016)
2. Chen, L.J., Luo, H.: A BIM-based construction quality management model and its applications. *Autom. Constr.* **46**, 64–73 (2014)

3. Dos Santos Jr., R.F., Lu, C.-T.: Geography markup language (GML). In: Encyclopedia of GIS (2016)
4. Yi, J.: Research on construction project cost management. *Fujian Build. Mater.* **01**, 111–113 (2021)
5. Jun, L.G.: Present situation and countermeasures of the construction project whole process. *Chin. Archit. Decor.* **01**, 160–161 (2021)
6. Qi, Z.: Application research of BIM technology in project cost management. *Neijiang Technol.* **42**(05), 59–60 (2021)
7. Yingjie, C.: Practice of BIM technology in construction project management. *Resid. Real Estate* **24**, 123 (2020)
8. Peiyuan, L.: Present situation and development suggestion of whole process cost management. *Resid. Real Estate* **24**, 41–42 (2020)
9. Cavalliere, C., Dell' Osso, G.R., Favia, F., Lovicario, M.: BIM-based assessment metrics for the functional flexibility of building designs. *Autom. Constr.* **107** (2019)
10. Li, X., Wang, C., Alashwal, A., Bora, S.: Game analysis on prefabricated building evolution based on dynamic revenue risks in China. *J. Clean. Prod.* **267** (2020)
11. Kochovski, P., Stankovski, V.: Supporting smart construction with dependable edge computing infrastructures and applications. *J. Autom. Constr.* **85**, 182–192 (2018)



# Modal Parameter Identification of Bridge Structure Based on Hybrid Genetic Algorithm

Rong Hu<sup>(✉)</sup>

School of Road Bridge and Architecture, Chongqing Vocational College of Transportation,  
Chongqing 402247, China  
mingtiandaban@163.com

**Abstract.** Bridges play an irreplaceable role in the structure of modern transportation system and play an important pivotal role in the development of politics, economy and culture. In the development of bridge engineering, the modal parameter identification (MPI) of bridge structure is particularly important. Therefore, this paper studies and analyzes the MPI of bridge structure based on hybrid genetic algorithm (HGA). Firstly, the identification method of bridge structural modal parameters and the identification content of bridge construction parameters are briefly analyzed, and the HGA is proposed. It is analyzed that the HGA mainly plays the role of fitting and Optimization in the identification of bridge structural modal parameters; Finally, based on the monitoring project of a Provincial Railway temporary bridge, combined with the finite element theory analysis of ANSYS, the genetic algorithm is applied to MPI by combining signal filtering and random decrement method. The test results show that the minimum frequency error is 1.93%, the maximum error is 9.33%, and the first three frequency errors are within 6%. When the genetic algorithm is applied to MPI, the modal order determination problem has a great impact on the results of parameter identification, The feasibility and effectiveness of HGA applied to bridge structure MPI are verified.

**Keywords:** Hybrid Genetic Algorithm · Bridge Structure · Modal Parameters · Parameter Identification

## 1 Introduction

In the process of bridge construction, in addition to considering the influence of non-uniformity of materials on structural stress, climate humidity, temperature and other uncertain factors also need to be considered. In addition, the construction method adopted is generally multi process and multi-stage construction. With the progress of construction, these factors often make the displacement and internal force of each construction stage gradually deviate from the theoretical value. The parameters adopted in the design, such as the rigidity of the cradle, the dead weight of the structure, the modulus of elasticity of materials, the shrinkage and creep coefficient of concrete and the temporary construction load, will be different from those in the actual project, which will make the state of

the actual structure in each construction stage different from the theoretical calculation. Therefore, the main parameters of the bridge structure should be calculated according to the measured data during the construction process, and then the modified parameters should be fed back to the actual construction control calculation. In order to ensure the quality and safety of bridge construction, the identification of bridge parameters in the construction stage is indispensable.

MPI of bridge structure based on HGA has been studied and analyzed by many scholars at home and abroad. Matsubaram proposed a method for identifying the parameters of passenger car tires based on the three-dimensional flexible ring model. This method can identify the modal parameters through experimental modal analysis, and compare the model parameters with the modal parameters by using the model calculation. The recalculated results using the model parameters show a good correlation with the experimental results [1]. Schfletr proposed a new off-line optimization method to solve the coverage path planning problem. For grid based environment representation, a new HGA is proposed, which uses turning start point and backtracking spiral algorithm for local search. The calculation results show that, compared with the traditional method, the path improvement rate of HGA is as high as 38.4%, and it has the same adaptability to different starting positions in the environment [2].

In this paper, a HGA is proposed by combining genetic algorithm with random decrement method and signal filtering. The MPI method based on HGA is discussed and analyzed. The performance of HGA (HGA) in avoiding the trap of local optimization and finding the global optimal solution is studied; How to identify the modal parameters of linear time invariant structural system by using the optimization function of genetic algorithm is discussed. The application of MPI based on HGA in engineering practice is discussed. Through the analysis of various analog signals and measured bridge signals, it is proved that the bridge structure MPI method based on hybrid algorithm proposed in this paper can process the bridge test signals in various environments and identify the bridge structure modal parameter information [3, 4].

## 2 MPI of Bridge Structure

### 2.1 Bridge Structure MPI Method

Frequency domain method: most frequency domain identification methods are based on fast Fourier transform. They have the advantages of mature theory, simple operation, fast identification speed and high identification accuracy. However, due to its limited frequency resolution, it is not enough to decouple the dense modes. The principle that the structural frequency response function has a maximum value at the natural frequency is used to identify the natural frequency. This method does not need to set parameters, and has the advantages of convenient operation and fast recognition speed. However, the dense modes cannot be decoupled, and the mode shapes cannot be obtained directly. Instead, the working deflection line shape is used to approximately replace the mode shapes [5].

Frequency domain decomposition method: the frequency domain decomposition method is an extension of the peak picking method. This method has certain anti noise

ability and high identification accuracy, but its decoupling ability of low-frequency dense modes is still not high.

**Polynomial fitting method:** polynomial fitting method generally carries out high-order polynomial fitting for each frequency response function, and then uses some form of averaging to obtain the overall modal parameters of the structure. But this method is easy to lead to ill conditioned matrix, and can not get high-precision fitting. In order to solve this problem, an orthogonal polynomial fitting method is proposed to improve the accuracy of modal identification.

**Time domain method:** the time domain method directly uses the system response time history signal to identify the modal parameters, and does not need to use Fourier transform to transform the signal into the frequency domain for analysis. Therefore, there is no problem of frequency resolution. However, the time domain method is difficult to determine the system order, sensitive to noise and prone to false modes. The time domain method uses the response data obtained by the random decrement method or the natural excitation technique to establish the mathematical model of the characteristic matrix equation, and uses the relationship between the system modal frequency, modal damping and the eigenvalue of the characteristic matrix to solve the modal parameters.

The random decrement method eliminates the structural response caused by random load through the sample averaging method, so as to convert the random response signal into a free attenuation signal [6]. The natural excitation technology uses the cross-correlation function of the response signals of two arbitrary measuring points on the structure to have a similar mathematical expression with the impulse response function under impulse excitation, so the cross-correlation function is used to replace the impulse response function.

**Random subspace method:** the random subspace identification method is a completely data-driven parameter identification algorithm. It does not need to obtain the free attenuation signal or impulse response function of the structure through random decrement method and natural excitation technology, and has certain anti-interference ability to noise. Stochastic subspace algorithm has been widely used in engineering because of its clear concept, perfect theory and easy programming. Generally, the stability diagram method or singular value entropy method is used to judge the system order, and the stability diagram method can also help eliminate false modes and improve the identification accuracy [7, 8].

**Time frequency domain method:** since both frequency domain method and time domain method assume that the test process is a stationary random process, it cannot meet the requirements of non-stationary signal and time-varying system parameter identification. Modern time-frequency analysis methods provide a means to analyze non-stationary signals. In addition, these methods also have excellent low-frequency dense mode decoupling ability, which is of great significance to practical projects, especially long-span bridge structures.

**Wavelet transform method:** the basic process of the wavelet transform identification method of modal parameters based on environmental excitation is as follows: firstly, the random decrement method or natural excitation technology is used to preprocess the structural response under environmental excitation to obtain the free attenuation signal or impulse response function; Then the wavelet base is constructed by using the appropriate

mother wavelet, and the processed signal is transformed by wavelet to obtain the time-frequency distribution of wavelet coefficients, on which the wavelet ridge is extracted; Finally, the modal parameters of the structure are extracted from the wavelet coefficients of the wavelet ridge. This method has excellent anti noise ability, low-frequency dense mode decoupling ability and analysis ability for non-stationary signals [9]. However, this method still needs to be further improved, such as the extraction of wavelet ridge, the elimination of endpoint effect and the design of optimal wavelet basis function.

## 2.2 Parameter Identification During Bridge Construction

Parameter identification content: parameter identification is to first determine the bridge structural parameters that have a great impact on the bridge response, then based on the error between the measured response data and the theoretical calculation data, and finally feed back the actual structural parameters to the construction control calculation, so as to timely adjust the theoretical values required for the bridge construction in the next stage. For the identification of bridge structural parameters, the main structural parameters causing the structural state deviation must be determined by some analysis method, and then the appropriate parameter identification theory or method should be used to identify the structural parameters [10]. For the general bridge structure, the main structural parameters refer to the factors that can significantly cause the change of the bridge structure state.

### MPI of vertical Bridge

Signal processing: through the preliminary analysis of the test data, it can be seen that the vertical sensors arranged at the north side span of the bridge have failed, the test results are not ideal, and the test data are unavailable. The data collected by the vertical sensors at other positions are ideal, which can be used for the analysis and calculation of MPI. Due to the large span of the main span, the vibration amplitude under environmental excitation is large, and the sensor has obvious perception of vibration, so the reliability of the collected vibration data is high. For the long-span bridge constructed by phased cantilever, the numerical analysis and construction control during construction play an irreplaceable role in the smooth construction, and the parameter identification is the difficulty and focus of numerical analysis and construction control [11]. The parameter identification in the bridge construction stage is to first analyze the main parameters that have a great impact on the bridge structure state through the parameter sensitivity analysis, and then use certain methods to estimate the error between the actual parameters and the theoretical parameters according to the error between the measured data and the theoretical data during the bridge construction, so as to identify the bridge parameters in the actual construction state, and use the identified parameters to guide the subsequent construction stage, Finally, the bridge completion state of the structure is consistent with the ideal bridge completion state [12].

## 3 Frequency Domain MPI Based on HGA

As a new optimization method, HGA (HGA) is attractive for its excellent computational performance and remarkable application effect. The combination of genetic algorithm

and computer technology has created a new research field, and constantly infiltrated into other fields to give full play to its excellent performance.

HGA (HGA) mainly plays the role of fitting and Optimization in MPI of bridge structures. The vibration of the multi degree of freedom system is assumed to be the superposition of multiple impulse responses. Through the random decrement technology and signal filtering technology, the free attenuation signal is fitted with the determined impulse response function. When the signal contains fewer frequency components, the easier the fitting optimization is and the more accurate the result is. Therefore, before using the HGA, filtering the high-frequency noise components in the signal can effectively improve the accuracy of the recognition results. At the same time, when using HGA, the setting of parameters plays an important role in the accuracy of the results.

Because the solution obtained by HGA always makes the objective function tend to the minimum value when optimizing problems, it is necessary to transform the objective function when using this algorithm. The frequency response function model of the structural system is:

$$K_{lp}(\gamma; e, \delta, \gamma_r) = \sum_{r=1}^N \frac{1}{e_r(1 - (\gamma/\gamma_r)^2 + 2\delta\gamma/\gamma_r)} \tag{1}$$

$e_r, \delta_r, \gamma_r$  ( $r = 1, 2, \dots, N$ ) is the modal parameter to be identified.  $\hat{K}_{lp}$  is the measured frequency response function, and the theoretical frequency response function is  $K_{lp}$ . The identification problem is transformed into minimizing the difference between  $\hat{K}_{lp}$  and  $K_{lp}$ . Namely:

$$\min X = \min \sum_{i=1}^T [\hat{K}_{lp}(\gamma_i) - \sum_{r=1}^N \frac{1}{e_r(1 - (\gamma_i/\gamma_r)^2 + 2\delta_r\gamma_i/\gamma_r)}]^2 \tag{2}$$

The fitness function is:

$$\begin{aligned} j &= j_{\max} - X \\ &= j_{\max} - \sum_{i=1}^T [\hat{K}_{lp}(\gamma_i) - \sum_{r=1}^N \frac{1}{e_r(1 - (\gamma_i/\gamma_r)^2 + 2\delta_r\gamma_i/\gamma_r)}]^2 \end{aligned} \tag{3}$$

$j_{\max}$  is a known quantity set before identification to ensure  $j > 0$ .

#### 4 MPI of Bridge Structure Based on HGA

Based on the monitoring project of a Provincial Railway temporary bridge, this paper studies the application of MPI based on HGA in engineering. Through the detailed project overview, the risks existing in the project construction are understood, which reflects the necessity of monitoring. Through the improvement of time domain MPI method based on modal decomposition, the identification method suitable for engineering practice is obtained, so as to improve the accuracy of identification results.

In order to monitor and evaluate the health status of the bridge during its operation, a health monitoring system was designed during the construction of the bridge. A variety

of sensors were installed at the main positions of the bridge girder, tower, stay cable, etc. if the annual data were analyzed and processed, the workload would be huge. Therefore, the data volume of a day with ideal test data quality was selected as the analysis object, with a total of 24 time history files, The sampling time of each time history data is 3600 s and the sampling frequency is 20 Hz.

#### 4.1 Finite Element Theoretical Analysis Based on ANSYS

ANSYS large-scale general finite element software is widely used in structural engineering, bridge engineering, geotechnical engineering, water conservancy engineering and other fields because of its powerful function and versatility. In order to master the dynamic characteristics of the temporary railway bridge, the finite element theoretical analysis of the temporary railway bridge is carried out by using the finite element software ANSYS, so as to compare with the identification value of the measured signal.

The bridge deck is made of in-situ reinforced concrete continuous slab, and the pier body is made of angle steel lattice column pier. The first five natural frequencies and vibration modes of the temporary bridge before horizontal and vertical bending are calculated by subspace iteration method. See Table 1.

**Table 1.** Natural frequency and mode shape of temporary bridge in transverse and vertical bending

Transverse bending		Vertical bend
stage	Frequency /HZ	Frequency /HZ
1	1.752	7.455
2	1.877	7.462
3	2.271	8.116
4	3.381	8.121
5	7.134	8.190

#### 4.2 MPI Based on Genetic Algorithm

Signal preprocessing: the measured signal is generated by environmental excitation. The vibration signal under environmental excitation is used to identify the modal parameters. Data preprocessing is required to make the signal conform to the form required by the time domain identification method. Usually, the random decrement method is used to extract the free vibration signal, or the next method is used to take the cross-correlation function as the time domain identification input data. Firstly, the signal is filtered and denoised, and then the free vibration response signal is extracted from the original signal by random decrement method, and the modal parameters of the preprocessed signal are identified.

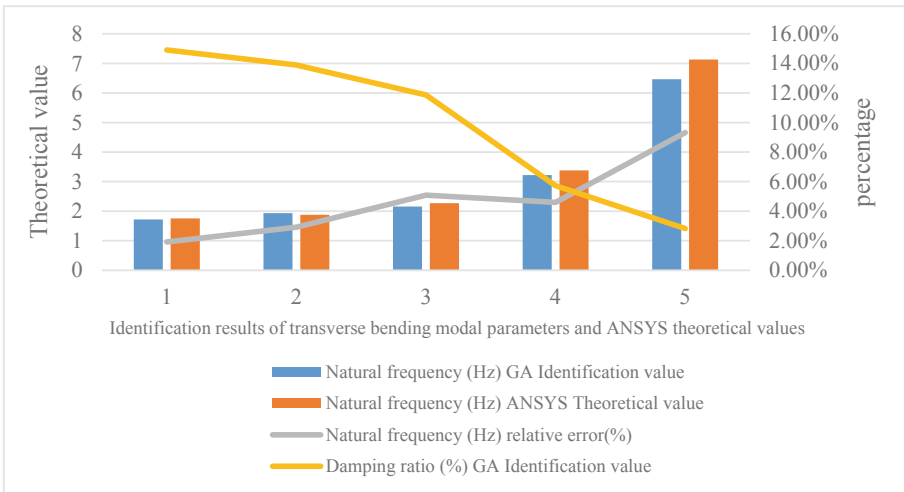


After preprocessing the measured signal, the free attenuation vibration response of each channel is obtained, and the input data required by the real-time domain identification method is obtained. A genetic algorithm parameter identification program based on MATLAB is developed to identify the modal parameters of horizontal and vertical measured signals respectively. The parameters of HGA are set as follows: population size 600, initial range [1:9], crossover probability 0.95, mutation probability 0.015, iterative evolution times 100, stop criterion using maximum evolution times, and coding method using binary coding. See Table 2 and Fig. 1 and Fig. 2 for the identification results of damping ratio and ANSYS theoretical values.

**Table 2.** Ratio of transverse bending MPI result to ANSYS theoretical value

stage	Natural frequency (Hz)			Damping ratio (%)
	GA Identification value	ANSYS Theoretical value	relative error(%)	GA Identification value
1	1.720	1.754	1.92	14.9078
2	1.932	1.877	2.91	13.8950
3	2.155	2.269	5.08	11.8640
4	3.223	3.381	4.59	5.7279
5	6.466	7.133	9.32	2.8179

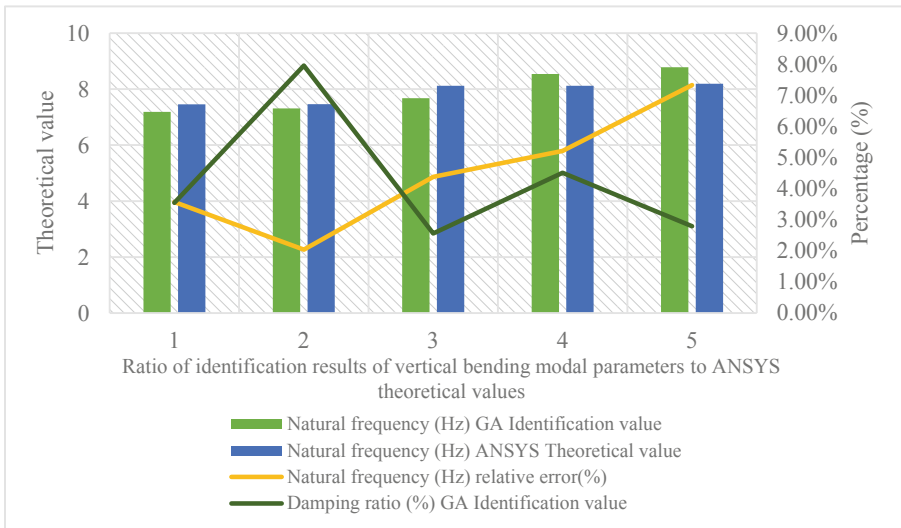
Based on time domain method and frequency domain method, this paper applies genetic algorithm to MPI of single degree of freedom and multi degree of freedom simulation signals. The results show that the maximum error of frequency and damping



**Fig. 1.** Identification results of transverse bending modal parameters and ANSYS theoretical values

ratio is 1.84% and 3.1% respectively without noise. When the noise is 20%, the maximum frequency error is 24.04% of the maximum damping ratio error. It can be seen that the frequency identification accuracy is high and the damping ratio identification accuracy is relatively low. When using genetic algorithm to identify modal parameters, it has strong anti noise ability, which reflects its strong robustness. However, the algorithm has many parameters and needs more debugging in use to minimize the error.

Comparing Fig. 1 and Fig. 2, it can be found that in the first five natural frequencies, the natural frequency of vertical bending is larger than that of horizontal bending, and the damping ratio of vertical bending is smaller than that of horizontal bending. This is mainly because the lattice columns are dense, and the transverse stiffness is smaller than the vertical stiffness, making the temporary bridge more prone to transverse vibration. Comparing the identification value of HGA with the theoretical value of ANSYS, it can be found that the relative error increases with the increase of order. Using the optimization function of genetic algorithm to identify the natural frequency, the identification accuracy of low-order frequency is higher than that of high-order frequency. As the actual structural stiffness is less than the modeling stiffness, the identification result is also less than the ANSYS theoretical calculation result.



**Fig. 2.** Ratio of identification results of vertical bending modal parameters to ANSYS theoretical values

In this paper, by combining signal filtering and random decrement method, genetic algorithm is applied to MPI, that is, HGA is applied to bridge MPI. It can be seen from Fig. 1 and Fig. 2 that the minimum frequency error is 1.93%, the maximum error is 9.33%, and the first three frequency errors are within 6%, but the error increases with the increase of modal order, so the accuracy problem when used to identify high-order modes is worth considering. Genetic algorithm is applied to MPI, and modal order

determination has a great impact on the results of parameter identification, which shows the feasibility and effectiveness of HGA applied to bridge structural MPI.

## 5 Conclusions

At present, many scholars at home and abroad have done a lot of research on MPI methods and achieved rich results. However, each identification method has certain limitations. It is particularly important to apply a new method to MPI to overcome the limitations of existing methods. For this purpose, this paper combines genetic algorithm with random decrement technology and signal de-noising technology to propose a bridge structure MPI method based on HGA. Although some achievements have been made, there are still many shortcomings worth further study: MPI is based on test signal analysis, so the identification results are greatly affected by the quality of test data, The quality of the data even directly affects the development of the identification work. How to process the signal to extract the effective information of the structure in the case of weak vibration signal and general test data quality still needs further research; In this paper, due to the limited resources, the research object is a single bridge type. Both the model bridge and the actual bridge are cable-stayed bridges. Whether other types of bridges can also achieve good identification results needs further research and verification.

## References

1. Matsubara, M., Kawamura, S.: Parameter identification of a three-dimensional flexible ring-based model of a tire using experimental modal analysis. *Int. J. Autom. Eng.* **10**(2), 133–138 (2019)
2. Schffe, T.R., Mitschke, M., Uchiyama, N.: Generation of optimal coverage paths for mobile robots using HGA. *J. Robot. Mechatron.* **33**(1), 11–23 (2021)
3. Watanabe, S., Keyaki, T., Naito, N., et al.: Automatic identification method for natural frequency of bridge piers by microtremor measurement at both sides on top of pier. *Q. Rep. RTRI* **61**(2), 103–108 (2020)
4. Silva, M.S., Neves, F.A.: Modal identification of bridge 44 of the Carajás Railroad and numerical modeling using the finite element method. *Revista IBRACON de Estruturas e Materiais* **13**(1), 39–68 (2020)
5. Fan, L., Liu, X., Cai, G.P.: Dynamic modeling and modal parameters identification of satellite with large-scale membrane antenna. *Adv. Space Res.* **63**(12), 4046–4057 (2019)
6. Naranjo-Pérez, J., Jiménez-Alonso, J.F., Sáez, A.: Parameter identification of the dynamic Winkler soil–structure interaction model using a hybrid unscented Kalman filter–multi-objective harmony search algorithm. *Adv. Struct. Eng.* **23**(12), 2653–2668 (2020)
7. Omidalizarandi, M., Herrmann, R., Kargoll, B., et al.: A validated robust and automatic procedure for vibration analysis of bridge structures using MEMS accelerometers. *J. Appl. Geodesy* **14**(3), 327–354 (2020)
8. Ahmad, M., Kumar, N., Kumari, R.: A HGA approach to solve inverse kinematics of a mechanical manipulator. *Int. J. Sci. Technol. Res.* **8**(9), 1777–1782 (2019)
9. Mathur, A.: Hybrid combination of error back propagation and genetic algorithm for text document clustering. *Int. J. Comput. Trends Technol.* **68**(11), 64–68 (2020)

10. Khalaf, J.A., Majeed, A.A., Aldlemy, M.S., et al.: Hybridized deep learning model for perfbond rib shear strength connector prediction. *Complexity* **2021**(8), 1–21 (2021)
11. Shrividya, G.: Application of HGA for successful CS-MRI reconstruction. *J. Adv. Res. Dyn. Control Syst.* **12**(3), 408–414 (2020)
12. Sun, L., Xu, Y.: MPI and finite element model updating of a long-span aqueduct structure based on ambient excitation. *J. Vibroeng.* **22**(3), 896–908 (2020)



# Overload Damage Detection Method of Motor Car Axle Based on Neural Network Algorithm

Pin Xia<sup>(✉)</sup>

College of Intelligent Manufacturing and Automobile,  
Chongqing Vocational College of Transportation, Chongqing 402247, China  
xarp@163.com

**Abstract.** Neural network is a new theoretical model. It has the ability of parallel processing. It can classify, define and optimize information and knowledge by simulating biological neural system. In this paper, a typical nonlinear deformation and damage monitoring method is trained by BP algorithm, which is based on neural network to detect the axle overload strength. Firstly, the fatigue response characteristics of the corresponding working conditions (such as low speed) under different stress states on the axle when the method is running in the motor car are studied by experimental method. Secondly, the actual working environment is simulated as the process of high-speed driving through the design model, and the detection degree of axle overload damage in this scenario is tested by the model. Finally, the test results show that the running time of the motor car axle overload damage detection model based on neural network algorithm is relatively short, and the delay time is also relatively short. The probability of checking the overload damage is basically more than 90%, which shows that the motor car axle overload damage detection rate of this model is very high and can meet the needs of users.

**Keywords:** Neural Network Algorithm · Motor Car Axle · Overload Load Loss and Damage Detection

## 1 Introduction

With the rapid development of social economy and technology, highway traffic plays a more and more important role in cities, and traffic accidents are also increasing [1, 2]. Therefore, it is particularly urgent to evaluate the safety of vehicles. In order to minimize the personal and property losses and maximize the driving speed, it is necessary to develop an action efficiency that can accurately predict the accident probability, and take corresponding measures in time to ensure that the personal and property will not be damaged. Neural network is a nonlinear system formed by a large number of neurons through simulation. It has the advantages of good approximation performance and strong fault tolerance, and is widely used in the field of traffic safety [3, 4].

Many scholars at home and abroad have done relevant research on neural networks. Neural network is a new computer-aided system. It has been widely used in the field of biological intelligence and human brain, and has been widely used in various scientific

and engineering designs [5, 6]. The research on sports injury detection technology started early in foreign countries. The United States, Germany and other developed countries have begun to use this technology for fault diagnosis and maintenance. American scholars have proposed artificial neuron simulation method to predict the damage degree of vehicle axles. Japanese scholars have developed a nonlinear finite element simulation software based on BP algorithm - fuzzy bases and artificial neural network to deal with the stress distribution and size change law of the top of the car model, and optimize the model on the computer [7, 8]. There are also some mature companies in China that are developing integrated intelligent vehicle component detection methods and research work based on artificial neural network (annr), BP algorithm and other artificial intelligence systems, and have achieved some results. The above research has laid the research foundation for this paper.

Neural network is a nonlinear system analysis method, which has the characteristics of high parallelism and good robustness. It is widely used in solving complex problems. In this paper, the intelligent traffic monitoring platform is modeled based on the principle of neural network algorithm. Firstly, the overload damage detection technology and working process of intelligent axle are introduced. Then the linear crack initiation mechanism is established based on BP algorithm under a certain working condition, and the corresponding diagnosis model and method are proposed. Finally, using the research results, the influence factors of different parameters on the deformation characteristics, contact stress distribution and crack propagation of vehicle journal are analyzed.

## **2 Discussion on Overload Damage Detection Method of Motor Car Axle Based on Neural Network Algorithm**

### **2.1 Overload Measurement Method of Motor Car Axle**

The overload load detection method of vehicle axle is mainly based on artificial neural network, which designs and learns the structure of cerebral cortex by simulating human brain neurons and external signal stimulation, so that it has better anti fatigue and strong robustness. It also includes the direct contact method [9, 10]. The test is to calculate whether the deformation occurs at the corresponding parts by manually collecting the radial tensile stress, rotation angle and other data at different positions of the upper body. However, this method can only obtain a point strain diagram with a direction parallel to the centerline of the wheel axis, and can not obtain the displacement curves and corresponding angle values of all cut-in points on the centerline diagram in the axis top plane. Because the traditional manual measuring equipment has certain limitations in on-line vehicle monitoring, and its work efficiency is also low, it requires a lot of manpower to complete data collection and other operations. At the same time, manual ranging can not meet the requirements of real-time dynamic monitoring and the defects of slow data processing speed and low accuracy. It is also common for vehicles to suffer from axle overload damage caused by various factors during driving Random occurrence.

### **2.2 Influence of Axle Overload on Motor Car**

Axle damage refers to the deformation of vehicle body caused by external force during driving, resulting in the bending of vehicle body surface or interior, wheel locking

(depression), roll and tire wear. What affects the overload fatigue life of the axle is that the changes of its main braking performance and structural parameters interfere with the test results to a great extent [11]. When there is a certain error between the vehicle motion track and the actual situation, the measured value will deviate, resulting in inaccurate measurement. During the running process of the motor car, the axle is subjected to the force between the wheel and the track, resulting in bending deformation, torsion and compression. When the vehicle body is damaged by the ground applied to the vehicle (such as the front wheel) and the steering linkage (or the rear wheel), the vibration waveform will be distorted and the vehicle body will shake or roll over. At the same time, under the driving state of the vehicle body, due to the gravity of the vehicle itself, the axle will also be bent, deformed, twisted and compressed due to excessive external force.

### **2.3 Factors Affecting Overload of Motor Car Axle**

The main factors affecting vehicle axle overload are: (1) operating conditions. This includes driving speed, number of stops, etc. In practical work, the requirements for the bearing force and stiffness of the frame are different under different working conditions. At high speed, the braking pressure is large and the deceleration is slow. At low speed, the vehicle speed is fast but the braking distance is long and there are some obstacles to restrict its normal movement or, if it is necessary to reduce the range, the detection method must be used to evaluate and determine whether there is necessary to leave enough clearance between the vehicle axle and the rail to ensure safety. (2) Load characteristics. The vehicle is subject to a variety of forces during driving, mainly gravity, wind, etc., and will also be affected by various power sources in different directions and angles. Therefore, the bearing capacity of the axle to the track is different and variable. At the same time, considering the friction resistance between the vehicle body and the bridge wall and the structural stiffness problems, the deformation of the vehicle wheels may cause the vertical vibration, bending, deflection or even fracture failure of the vehicle body, resulting in serious consequences such as vehicle safety accidents or frequent traffic accidents.

### **2.4 Neural Network Algorithm**

Neural network is a new and widely used information processing model. It simplifies and parallelizes the functions of neurons in the human brain connecting with the outside world, so that it can achieve the maximum performance goal of human brain's cognitive ability when solving complex problems. It is designed to simulate the structure of biological nervous system. When processed in the computer, the input and output signals are connected with different types of neurons. Through the storage of neural information and learning rules to achieve the automatic adaptability of the artificial system. BP routing protocol can be divided into three layers: perception layer, hidden layer and application service layer. It has a strong self-learning, self-learning and automatic adaptation system. This network is a mathematical algorithm model. By adjusting the connection weights and other parameters between a large number of internal neurons, it can intelligently learn and train the input and output sample data, mine the potential relationship between

input and output, and have the ability to calculate and predict new samples to obtain the prediction results. Neural algorithm can be used to predict the changing trend of unknown parameters (inputs) in the model without any external factors to modify the model parameters, so as to obtain the optimal results.

Input information  $X_i$  and threshold in neuron expression  $\theta_k$  constitutes a linear combination, so the threshold can be regarded as a specific input information, then input  $x_0 = -1$ , and the corresponding weight  $w_{k0} = \theta_k$  to obtain:

$$u_k = \sum_{i=0}^m w_{ki} x_i \quad (1)$$

There are two kinds of nodes in the network: input node and calculation node. The input node only receives signals, and the calculation node is the unit neuron. So the final mathematical expression of unit neuron is:

$$y_k = \phi \left( \sum_{i=0}^m w_{ki} x_i \right) \quad (2)$$

The single-layer neural network, which puts many neurons on the same computing level, is just an output layer, which can solve the linear separable problem well, but it can not deal with the nonlinear separable problem at the same time. Therefore, a single neuron can be regarded as a multi input and single output system, while a single-layer neural network can be regarded as a multi input and multi output system, but their working mechanisms are not much different in essence.

### 3 Experimental Process of Overload Damage Detection Method of Motor Car Axle Based on Neural Network Algorithm

#### 3.1 Process of Axle Overload Damage Detection Method Based on Neural Network Algorithm

It can be seen from Fig. 1 that according to the training and simulation results of neural network, combined with the actual working conditions, the detection method of axle overload damage has been deeply studied, and it is concluded that the neural network model based on BP algorithm has good anti noise ability when the factors such as the motion track, initial state and node position in the vehicle change. The intelligent traffic safety system based on BP algorithm can effectively prevent traffic accidents. The system is composed of several subsystems. Firstly, judge whether each subsystem has fault (i.e. whether it is an accident) through expert experience. Secondly, predict the axle overload damage. During the detection process, the axle will be slightly damaged due to various factors on the body surface, such as temperature, humidity, etc. According to the neural network model, the deformation of vehicle body under different working conditions is analyzed. It is mainly described by establishing the relative position relationship between the corresponding nodes between the grid and the vehicle contact area, and then using the geometric coordinate system of the grid and the contact surface to convert it into a standard state vector, and calculate the corresponding input and output values.



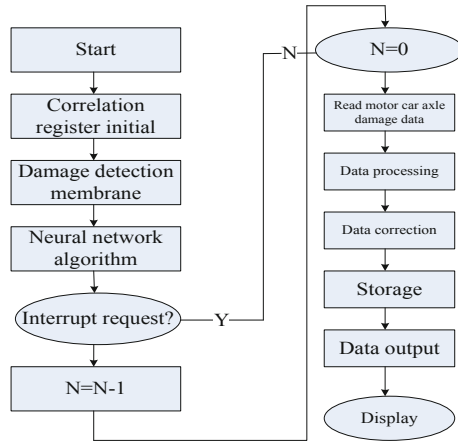


Fig. 1. Axle overload damage detection method and flow

### 3.2 Test Steps for Overload Damage Detection of Motor Car Axle Based on Neural Network Algorithm

The basic idea of neural network algorithm is to classify the nodes in the system layer by layer through input and output neurons, and aggregate the data sets of different types, sizes and attributes according to certain rules to form a parallel processing function with strong adaptability, rich information and good global optimization ability that can be combined with other topological structures. When the artificial neural network algorithm is applied in fault diagnosis, it should be determined according to the damage model and system state, and the methods used in different cases will be different. Therefore, in order to ensure the training success rate and data processing effect. First, initialize the vehicle body. That is, the vehicle body starts to learn and complete layer by layer from static to motion and from motion to stop. Secondly, the starting point of each iteration is the input and output current sampling value at the nodes of each part on the axle. After the corresponding damage model is established, it is necessary to collect the initial state and operating environment of the system. The physical model in the process of vehicle driving is established according to the kinematics theory. Secondly, the vehicle speed, acceleration and other parameter values are obtained through the training set and used as the prediction basis. The input and output are pre estimated by BP neural network algorithm and the error signal correction is calculated and analyzed. Finally, the axle overload damage detection task is realized.

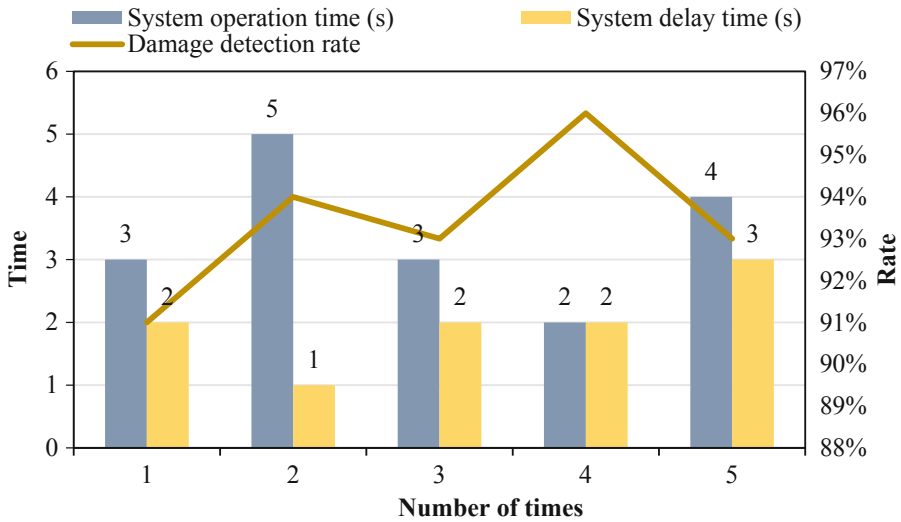
## 4 Experimental Analysis of Overload Damage Detection Method of Motor Car Axle Based on Neural Network Algorithm

Detection and Analysis of Overload Damage of Motor Car Axle Based on Neural Network Algorithm

Table 1 shows the performance test results of the neural network algorithm.

**Table 1.** Performance test results

Number of tests	Intelligent algorithm	System operation time (s)	System delay time (s)	Damage detection rate
1	Artificial neural algorithm	3	2	91%
2	Artificial neural algorithm	5	1	94%
3	Artificial neural algorithm	3	2	93%
4	Artificial neural algorithm	2	2	96%
5	Artificial neural algorithm	4	3	93%



**Fig. 2.** System test

This paper mainly studies the basic principles and related theories of neural network algorithm, and combines the actual motor car axle overload damage detection experiment to complete the intelligent traffic accident early warning, vehicle safety protection and rescue based on BP neural network. After the whole vehicle simulation platform is built, it is necessary to judge whether there is a fault according to the actual working conditions, and then compare the collected data with the operating conditions of the standard sample vehicle in the system to determine the detection method and performance indicators under the overload state of the motor car, and diagnose and evaluate different working conditions. It can be seen from Fig. 2 that the detection model of motor car axle overload

damage based on neural network algorithm has a short running time and a short delay time. The probability of checking the overload damage is basically above 90%, which shows that the detection rate of motor car axle overload damage of this model is very high and can meet the needs of users.

## 5 Conclusions

Neural network is a new and large-scale application field. It plays an important role in solving complex engineering problems and improving system performance. It is especially suitable for dynamic characteristic analysis under some nonlinear or uncertain working conditions. With the development and wide application of artificial intelligence technology and computer soft science, and the deepening understanding of neural network theory, an intelligent detection method for axle overload damage diagnosis based on artificial neuron is proposed. This paper analyzes and summarizes the neural network detection method of axle over strength damage. Firstly, the three-dimensional model is established and the body structure characteristics and vehicle driving conditions are modeled. Secondly, the nonlinear design unit is constructed by using BP neural network theory and kinematics equations to improve the identification accuracy. Finally, the fatigue response and strain degradation of the system under different types are verified by experiments, which has good application prospects and practical significance.

**Acknowledgements.** This work is partially supported by the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJZD-K202105701).

## References

1. Janabi, A.H., Kanakis, T., Johnson, M.: Convolutional neural network based algorithm for early warning proactive system security in software defined networks. *IEEE Access*. **10**, 14301–14310 (2022)
2. Janabi, A.H., Kanakis, T., Johnson, M.: Convolutional neural network based algorithm for early warning proactive system security in software defined networks. *IEEE Access* **10**, 14301–14310 (2022)
3. Setiawan, F., Liu, A.B., Lin, C.W.: Development of neuro-degenerative diseases' gait classification algorithm using convolutional neural network and wavelet coherence spectrogram of gait synchronization. *IEEE Access* **10**, 38137–38153 (2022)
4. Bharath, Y.K.: Griffiths' variable learning rate online sequential learning algorithm for feed-forward neural networks. *Autom. Control. Comput. Sci.* **56**(2), 160–165 (2022)
5. Ramalingam, P., Mehbodniya, A., Webber, J., Shabaz, M., Gopalakrishnan, L: Telemetry data compression algorithm using balanced recurrent neural network and deep learning. *Comput. Intell. Neurosci.* **10**, 4886586:1–4886586 (2022)
6. Kalphana, I., Kesavamurthy, T.: Convolutional neural network auto encoder channel estimation algorithm in MIMO-OFDM system. *Comput. Syst. Sci. Eng.* **41**(1), 171–185 (2022)
7. Abadal, S., Jain, A., Guirado, R., López-Alonso, J., Alarcón, E.: Computing graph neural networks: a survey from algorithms to accelerators. *ACM Comput. Surv.* **54**(9), 191, 1–191, 38 (2022)

8. Hosseini, S., Nezhad, A.E., Seilani, H.: Botnet detection using negative selection algorithm, convolution neural network and classification methods. *Evol. Syst.* **13**(1), 101–115 (2022).
9. Shafieenejad, I., Rouzi, E.D., Sardari, J., Araghi, M.S., Esmaeili, A., Zahedi, S.: Fuzzy logic, neural-fuzzy network and honey bees algorithm to develop the swarm motion of aerial robots. *Evol. Syst.* **13**(2), 319–330 (2022).
10. Ramakrishnan, D., Radhakrishnan, K.: Applying deep convolutional neural network (DCNN) algorithm in the cloud autonomous vehicles traffic model. *Int. Arab J. Inf. Technol.* **19**(2), 186–194 (2022)
11. Deepika, S., Senthil, S.: Credit card fraud detection using moth-flame earth worm optimisation algorithm-based deep belief neural network. *Int. J. Electron. Secur. Digit. Forensics* **14**(1), 53–75 (2022)
12. Bhullar, A.K., Kaur, R.: Sondhi, S: Modified neural network algorithm based robust design of AVR system using the Kharitonov theorem. *Int. J. Intell. Syst.* **37**(2), 1339–1370 (2022)
13. Salimi-Badr, A., Ebadzadeh, M.M.: A novel learning algorithm based on computing the rules' desired outputs of a TSK fuzzy neural network with non-separable fuzzy rules. *Neurocomputing* **470**, 139–153 (2022)



# Ground Wire Status Object Detection Based on Cloud Application

Haoran Li<sup>(✉)</sup>, Shengli Xie, and Qian Wang

State Grid Jibei Electric Power Company Limited Skills Training Center,  
Baoding Technical College of Electric Power, Baoding, Hebei, China  
lihaoran404@126.com

**Abstract.** The rapid development of my country's economy has led to a continuous increase in social electricity consumption. The increase in electricity consumption in the society as a whole not only brings about economic prosperity, but also increases the pressure on national transmission lines. Due to the increasing shortage of land in big cities, overhead lines will pose a potential threat to the safety of citizens' electricity consumption, and the cluttered overhead lines will have a greater impact on the appearance of the city, so more and more overhead lines are underground replaced by cables. The main purpose of this paper is to perform object detection on ground wire status based on cloud application. This paper mainly analyzes the theoretical basis of on-line monitoring of high-voltage cable grounding current, analyzes and solves the calculation method of ground-induced circulating current, analyzes the factors affecting the circulating current of cable sheath, and designs high-voltage cable grounding current hardware. Experiments show that the RMS value of the N600 grounding wire current changes to a maximum of 91% of the load current.

**Keywords:** Cloud Application · Ground Wire · Target Detection · Ground Wire Status Detection

## 1 Introduction

In the city, the backbone cable in the power grid is responsible for the power supply of a large area of the urban area, and the insulation problem of the cable must be highly valued. In the process of operation and maintenance of power cables, most people mainly pay attention to their insulation characteristics, and strive to improve their insulation to make the cable run more safely, but the insulation of its sheath is also very important, once the main insulation fails or is damaged. In this case, the metal sheath will also undergo changes in the electric field and generate a certain amount of induced electric energy. Therefore, in the process of monitoring and maintaining high-voltage cables, the problems reflected by the induced electric energy of the sheath should be considered at the same time. It can show the insulation characteristics of the sheath and the insulation characteristics of the entire cable at the same time. By monitoring the grounding current of the cable It can carry out early warning in a targeted manner, avoid major accidents,

and ensure the normal operation of high-voltage cables and the effective transmission of electric energy. Therefore, it is a very reliable and efficient method to carry out its own ground-induced circulating current monitoring for high-voltage cables [1, 2].

The methods and theories of ground wire detection have always been the focus of research, and the research results are numerous. For example, Bui KT proposed a monitoring device suitable for mines, which collects the induced current of the sheath, performs reduction and comparison, and obtains abnormal information [3]; On-line monitoring system of power cable sheath based on induced voltage and ground current [4]. But not much for ground wire condition detection on cloud applications.

The main purpose of this paper is to design and study the target detection system of the ground wire state based on the cloud application program. By monitoring the grounding current of the cable, it can give early warning in a targeted manner, avoid major accidents, and ensure the normal operation of high-voltage cables and the effective transmission of electric energy. Including the design of the cable circulation monitoring terminal, the device design of the field data acquisition, the software design of the monitoring device, etc. Design and implement the monitoring and management software of the upper computer, and realize the remote online monitoring function of the system through remote acquisition, which greatly facilitates the test of ground current and improves the safety factor of high-voltage cable operation. Carry out all-round tests on the system, including hardware and software, to make the system run and collect effective information to verify the reliability and stability of the system. The circulating current generated on the sheath is monitored to ensure the safe and stable operation of the main insulation and sheath insulation. Once abnormal data or damage occurs, an alarm will be issued to ensure the continuous and safe operation of high-voltage cables and greatly promote the smooth operation of the power grid [5, 6].

## **2 Design and Research of Ground Wire State Target Detection Based on Cloud Application**

### **2.1 System Design**

The system consists of on-site monitoring terminal, data relay device, radio frequency communication network, monitoring center monitoring and comprehensive analysis system, and adopts distributed control structure. A data acquisition device with wireless communication function is installed at each overhead line or indoor cable grounding node to monitor the grounding cable and induced voltage. The data relay device receives real-time data on site and manages multiple data acquisition terminals. Radio frequency network communication, connecting remote computer systems [7, 8].

The monitoring and evaluation analysis software of the upper computer is installed on the server of the monitoring center of the power company, and the upper computer communicates with the wireless receiving device of the GSM network through the serial port mode. The main functions and features of the system are:

Real-time monitoring of the voltage, circulating current and fault current induced by the single and double-circuit cable sheath, and comprehensively grasp the impact of the circulating current of the high-voltage cable sheath on the cable operation in real time.

Refer to parameters such as time and load current to comprehensively judge the characteristics of the cable grounding current.

The change law of the cable circulation under various load conditions, and the analysis of the influence on the cable operation.

Carry out early warning and over-limit alarm for the change of the circulating current of the cable operation, record the historical data of each contact on the whole point, and form data reports, historical curves, etc.

The data is transmitted to the monitoring terminal and the remote computer system by radio frequency communication technology or GSM/GPRS technology.

## 2.2 Analysis of Factors Affecting Metal Sheath Circulation

Because the grounding circulation technology of collecting high-voltage cables is a feasible method to judge the operating environment of high-voltage cables, many scientific research institutions have also explored this method. Before using the ground-induced circulating current calculation method to collect the circulating current, the factors affecting the circulating current generated by the metal sheath of the high-voltage cable must be considered [9, 10]. There are three main reasons for affecting the circulation of the metal sheath:

### (1) Power cable laying parameters, including

Laying method. The most common ones in production and life are the three-phase horizontal arrangement, the three-phase vertical arrangement, the right-angled triangle arrangement, and the equilateral triangle arrangement. The equilateral triangle arrangement can effectively reduce the induced circulating current. This is because the position of the three-phase transmission lines is relatively symmetrical, so the mutual inductance has a certain inhibitory effect. The three-phase arrangement, due to its asymmetrical arrangement, will have a greater impact on the grounding circulation of its high-voltage cables due to mutual inductance.

Metal sheath interconnection and grounding methods. In the process of laying the cable, the metal sheath needs to be grounded and interconnected to ensure the stable operation of the cable. Most high-voltage cables mainly choose the method of non-interconnection, only grounding at both ends, or connecting after crossing and then grounding, etc. In several cases, due to the large changes in the electrical parameters of the high-voltage cable sheath, such as changes in grounding resistance and wire parameters, the calculation methods will eventually be very different.

Line length. When the cable sheath is well grounded, the line length of the high-voltage cable can be disregarded, but when an external resistor whose resistance value cannot be ignored or the grounding resistance is large during the grounding process of the cable sheath, the external series resistance or contact resistance will affect the whole. The resistance of the loop should be paid attention to to a certain extent, and the influence of the grounding resistance must be considered, and the influence of the line length on the circulating current cannot be ignored.

Whether there is a return line. If the high-voltage cable has a return line, the environment has little influence on the cable, and the current forms a loop through the return line.

- (2) The size of the electric energy transmitted by the cable and the mutual inductance of each phase loop, mainly including

Load size. The three mutual inductances are related to their respective positions and should have nothing to do with the electrical energy of the cables. However, in the process of actually transporting electrical energy, due to different loads, the currents flowing through each cable are not the same, so the mutual inductance does not play a role in suppressing it. In the same way, the three mutual inductances will affect their respective mutual inductances due to their own current differences, resulting in greater inductive circulation.

Load unbalance. Load unbalance has multiple effects. For example, the choice of laying method will have a significant impact on load unbalance, but the choice of laying method is only one aspect of load unbalance. There are many single-phase power systems in the power system. Due to the existence of these irregular loads, the load will also be uneven, and due to its unstable behavior, the unbalance of the load will also vary.

Environmental factors and laying methods of high-voltage cables:

At present, cables are mostly laid in direct burial. However, the impact of this method is very significant. The cable sheath will directly form a loop with the ground through the ground wire, resulting in a large current. There are other laying methods, such as installing in the cable trench, so that the influence of the environment on the cable is relatively reduced. There are many other methods, and different types of laying methods have certain influence on the calculation [11, 12]. The error can be reduced by calculating the influence on the induced circulating current under different laying situations, so as to avoid misjudging the operation state of the high-voltage cable.

### 2.3 Algorithm Research of Ground Wire State Target Detection Based on Cloud Application

- (1) Calculation of ground induced circulating current

Most of the medium and high voltage cables in production are single-core, so the induced electromotive force on the cable sheath is related to the thickness of the wire, the distance between the multi-phase high-voltage cables, the size of the high-voltage cables to transmit power, and the distance to transmit power.

In order to collect the grounding circulating current information of the high-voltage cable, it is necessary to carry out analysis and calculation, guide the hardware design, and concentrate the distribution parameters of the cable. Through the expression of Kirchhoff's voltage law, the specific formula is obtained as follows:

$$\begin{aligned} E_A &= E_{SA} + I_{SA} [(R + jX + R_g)L + R_1] \\ E_B &= E_{SB} + I_{SB} [(R + jX + R_g)L + R_1] \\ E_C &= E_{SC} + I_{SC} [(R + jX + R_g)L + R_1] \end{aligned} \quad (1)$$



$R + jX$  is the impedance of the sheath of each cable wire, ISA, ISB, ISC are the current flowing on the sheath of each wire, EA, EB, EC are the sheath and the ground wire and the earth induced by each wire potential difference in this loop. R1 is the ground wire resistance and the sum of the ground wire resistance and the ground contact resistance.

## (2) Fault line judgment

By calculating the integral of the AEO of each line in the time domain, the relative energy value of each line from the time of the fault occurrence to the end of the simulation can be compared as the line selection criterion, which can realize the fault line selection more comprehensively and accurately. Assuming that there are  $N$  outgoing lines in the resonant grounded distribution network, the atomic decomposition method is used to decompose the transient zero-sequence current of each line after the fault into  $m$  atomic components and a residual component. Then the energy of the atomic component of the  $i$ -th line  $L_i$  at the  $j$ -th decomposition is:

$$E_{ij} = \int_{t_1}^{t_2} \psi_{ij}(x) dt \quad (2)$$

Among them,  $x$  is the optimal atom matched by the  $j$ th decomposition;  $j = 1, 2, \dots, m$ ;  $E_{ij}$  is the energy of the signal of the  $i$ -th line at the  $j$ th decomposition;  $m$  is the number of atomic components,  $t_1$  is the start time of the fault, and  $t_2$  is the end time of the simulation.

## 3 Experimental Study of Ground Wire State Target Detection Based on Cloud Application

### 3.1 Self-inductive Power Supply Circuit

Various electronic equipment applied on the high-voltage side needs to be insulated from the ground, and the safety of the equipment needs to be ensured during operation. The measurement circuit cannot be simply led out directly through the line to avoid accidents. Therefore, the power supply of the measuring equipment has become a key part of the inspection and detection system. The application of the measuring device plays a role in detecting the operation of the power grid, which is indispensable. Therefore, it is also very important to consider the related circuit research on how to obtain electricity from the high-voltage line.

The system adopts the transmission line to obtain electricity, obtains the voltage through the coil, and outputs the DC voltage through rectification and filtering. This method does not directly take electricity from the line, and has little impact on the original high-voltage circuit.

Adding a supercapacitor monitoring terminal to the power circuit design will reduce the use of the system, and replacing the battery will greatly reduce the automation of the system, and because the cables are often located in remote locations, it is difficult to replace the battery. The supercapacitor has a large capacity and can meet the power demand of the system, which has more advantages than ordinary lithium batteries. The

power supply of the monitoring terminal of the system adopts self-induction to obtain electricity, and is connected to the circuit through the current transformer. Use super capacitors to avoid instantaneous high-power power supply. Principle wiring diagram:

(1) Rectifier filter circuit

The switch-type power-taking access circuit, through the rectifier and filter circuit composed of the stabilizer tube, converts the AC power into DC power, in which D1 and C1 play the role of rectification, and L1 and C2 filter the current and voltage with large ripple to make its applied.

(2) Transient protection

As the energy device of the entire system equipment, the power supply must pay attention to its stability, and can still ensure that it is reactivated and not damaged in the face of various emergencies, and continues to supply power to the equipment. In order to avoid accidental short-circuit and high current, the iron core emits ultra-high voltage, which damages the circuit. In order to protect the circuit, the bipolar TVS1 tube is connected before the rectifier and filter circuit to limit the output impulse voltage.

(3) Voltage protection and energy discharge circuit

Since the power used by the system is roughly constant, the electric energy obtained by the mutual inductance is not stable, and its voltage and current have certain fluctuations. The filter back-end voltage  $U_{dc}$  will increase with the increase of the current. Considering the electricity safety of the entire system, it is necessary to maintain the voltage within the range of electricity consumption standards, and implement voltage protection and energy discharge on the basis of voltage. If the voltage is too large, the voltage comparator output is turned on at a high level, and the excess energy is directly discharged; when the  $U_{dc}$  is low, the output is low, and the bleeder circuit stops working, and there is no need to protect the subsequent circuits.

(4) Energy storage circuit

Use the super capacitor SC as an accessory to store energy, avoid rapid charging and discharging of energy, and ensure low temperature operation. The data transmission adopts wireless communication, and the power is large, which affects the normal operation of other circuits. Adding SC can effectively solve the problem.

### 3.2 Real-Time Data Monitoring System

The master station sends the data read from the slave station to the host computer through the RS232 serial port, and the host computer displays it through a series of calculations. Including real-time data, alarm information, historical records, settings four interfaces.

- 1) Real-time data: The real-time data interface is responsible for displaying real-time data, real-time curves and operating status. The staff can understand the grounding situation of high-voltage cables in various regions by viewing the real-time operating status. The numerical display is clear and accurate, and the curve display can clearly see the changes of high-voltage cables in various regions, which is more intuitive.

## (1) Real-time data display

The current value of each point of the detection line can be displayed, and the data is updated every ten seconds.

## (2) Real-time curve display

The current change of the current route can be detected in real time and updated every ten seconds. Any circuit can be selected and its current data for one hour can be obtained and displayed.

## (3) Running status light

Communication indicator: The upper computer communicates with the master station, and the green indicator flashes.

Communication alarm light: The line communication failure, the red communication alarm light is on.

Fault alarm light: When the line is faulty, the red fault alarm light is on.

2) Alarm information: The alarm information bar is responsible for displaying the running status of each line of each substation. If an alarm occurs, the corresponding position will be displayed in red, and the event record box will record and display the alarm information.

3) History: You can set the start time of communication, the completion time of communication, the always time and fault time of each line, and the moving cursor can also display the current current value.

4) Equipment test: The ground current online monitoring system designed in this paper is mainly tested in the following aspects:

(1) Real-time monitoring of the circulating current of the cable sheath and fault grounding current, to grasp the influence of the circulating current of the high-voltage cable sheath on the operation of the cable in real time.

(2) Refer to parameters such as time and load current to comprehensively judge the characteristics of the cable grounding current.

(3) Change law of cable circulation under various load conditions, and analyze the influence on cable operation.

(4) Carry out early warning and over-limit alarm for the change of the circulating current of the cable operation, record the historical data of each contact on the hour, and form a data report, historical curve, etc.

(5) The data is transmitted to the monitoring terminal and the remote computer system using radio frequency communication technology or GSM/GPRS technology.

## 4 Experimental Analysis of Ground Wire State Target Detection Based on Cloud Application

### 4.1 Induced Current Value of Cable Metal Sheath

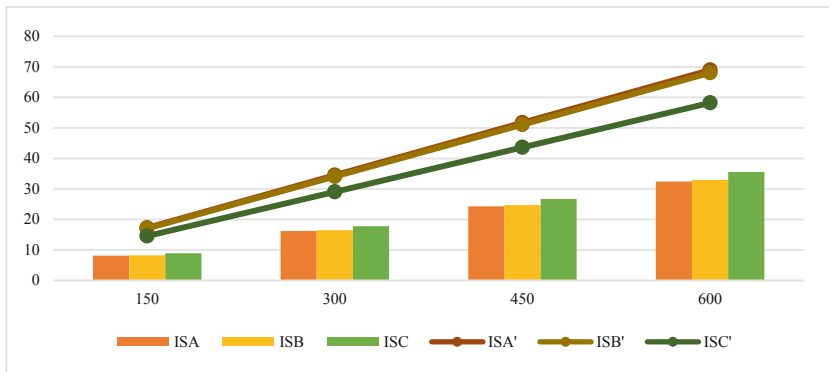
The length of each section of the test cable at the test site is 500 m, and the distance between the two phases is 35 cm. Under different load currents, two different three-phase

**Table 1.** Time Circulation Current Values (A) At Different Arrangements

Iload	Right triangle arrangement			horizontal arrangement		
	ISA	ISB	ISC	ISA'	ISB'	ISC'
150	8.1025	8.2373	8.8952	17.2432	17.0233	14.5474
300	16.2049	16.4747	17.7905	34.4865	34.0466	29.0468
450	24.3074	24.7120	26.6857	51.7297	51.0698	43.6422
600	32.4099	32.9493	35.5809	68.9729	68.0931	58.1956

arrangements are tested, and the detected induced current values are as follows. Table 1 Show:

It can be seen from Fig. 1 that under different load current values, the metal induced current is always larger in the horizontal arrangement than in the vertical arrangement. Therefore, the vertical arrangement is more conducive to protecting the metal sheath of the cable.

**Fig. 1.** Time circulation current values (a) at different arrangements

## 4.2 Analysis of the Influence of Load Unbalanced Current

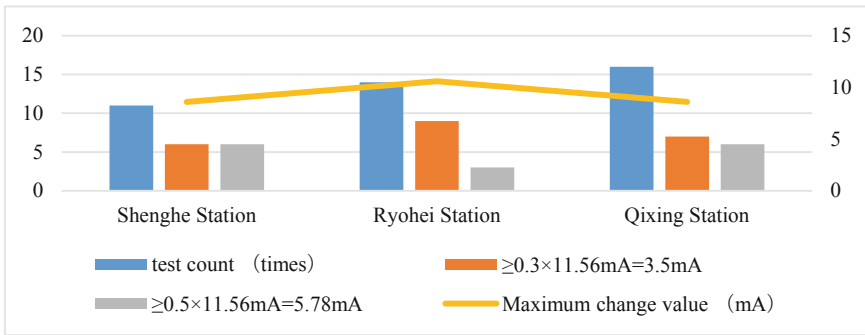
In this paper, the 110 kV Shenghe Station, the Liangping Station and the 220 kV Qixing Station are loaded with an unbalanced load current of  $57.8V/5\text{ k}\Omega = 11.56\text{ mA}$ , and the grounding current of the N600 under the condition of the load unbalanced current is calculated. The details are shown in Table 2 below:

As shown in Fig. 2, when there is an unbalanced load current, the variation characteristics of the RMS current of the N600 grounding wire are calculated. In the 41 experiments carried out, there were 22 times when the rms value of the N600 ground wire current changed more than 30% of the RMS load current compared with the normal situation, and 15 times when the ground current changed by more than 50% of the

**Table 2.** N600 ground current with unbalanced load current

Substation name	Shenghe station	Ryohei station	Qixing station
test count (times)	11	14	16
$\geq 0.3 \times 11.56 \text{ mA} = 3.5 \text{ mA}$	6	9	7
$\geq 0.5 \times 11.56 \text{ mA} = 5.78 \text{ mA}$	6	3	6
Maximum change value (mA)	8.6	10.6	8.6

unbalanced load current. In addition, in the experiment, the RMS value of the N600 grounding wire current has a maximum change of 91% of the load current.



**Fig. 2.** N600 ground current with unbalanced load current

## 5 Conclusions

On the basis of drawing on similar domestic technologies, according to the actual situation of the site, based on cloud applications, using sensors, wireless communication, inductive power acquisition technology, etc., through wireless radio frequency networking, the online monitoring and analysis of the grounding current of urban cables is realized.. The theoretical basis of on-line monitoring of high-voltage cable grounding current is analyzed, the calculation method of grounding induced circulating current is analyzed, and the factors affecting the circulating current of cable sheath are analyzed. Design the grounding current hardware of high-voltage cables, including the design of the cable circulating current monitoring terminal, the device design of on-site data acquisition, and the software design of the monitoring device. Design and implement the monitoring and management software of the upper computer, and realize the remote online monitoring function of the system through remote acquisition, which greatly facilitates the test of ground current and improves the safety factor of high-voltage cable operation. Carry out all-round tests on the system, including hardware and software, to make the system run and collect effective information to verify the reliability and stability of the system.

The system basically meets the power grid's requirements for monitoring the grounding current of high-voltage cables, and can analyze faults in a timely and effective manner, providing necessary support for the safe and stable operation of the power grid.

## References

1. King, P.: Pruitt's scandals cloud methane rule redo - senators. *Greenwire*, 1–2 (2018)
2. London, P.: Saro Cutty Sark Cloud & Windhover. *Aeroplane* **46**(12), 73–86 (2018)
3. Bui, K.T., Vo, L.V., Nguyen, C.M., et al.: A fault detection and diagnosis approach for multi-tier application in cloud computing. *J. Commun. Netw.* **22**(5), 399–414 (2020)
4. Gawande, P.S., Deshmukh, A., Rathod, A.L., et al.: Detection and identification of vehicle's noparking area using iotand cloud – a review. *J. Res. Eng. Appl. Sci.* **05**(1), 15–18 (2020)
5. Gilmer, E.M.: Groundwater policy still muddled until Supreme court rules. *Greenwire*, (APR. 23), 1–4 (2019)
6. Tabada, M.T., Loretero, M.E., Lasta, F.F.: Investigation on the performance of a multi-wire water level detection system using contact sensing for river water monitoring. *SN Appl. Sci.* **2**(1), 1–8 (2019). <https://doi.org/10.1007/s42452-019-1887-0>
7. Voevodina, E.: Commissioning of the New sMDT monitored drift tube detectors for the phase-1 upgrade of the ATLAS muon spectrometer. *Phys. At. Nucl.* **84**(4), 562–570 (2021). <https://doi.org/10.1134/S1063778821040281>
8. Naval, S., Laxmi, V., Rajarajan, M., et al.: Employing program semantics for malware detection. *IEEE Trans. Inf. Forensics Secur.* **10**(12), 2591–2604 (2017)
9. Coppo, M., Bignucolo, F., Turri, R., et al.: Analysis of frequency distribution of ground fault-current magnitude in transmission networks for electrical safety evaluation. *Electr. Power Syst. Res.* **173**(Aug.), 100–111 (2019)
10. Etwire, P.M., Buah, S., Ouédraogo, M.: An assessment of mobile phone-based dissemination of weather and market information in the Upper West Region of Ghana. *Agricul. Food Secur.* **6**(1), 1–9 (2017)
11. Sonner, S.: Green groups oppose off-roaders' lawsuit over grouse plan. *Greenwire*, (Mar. 20), 1–1 (2019)
12. Bagi, T.: Commonly used ground wire of distribution networks under various operating and load conditions. *Acta Electrotechnica et Informatica* **20**(1), 3–8 (2020)



# Design of Multimedia Assisted Course Based on Immersion Japanese Teaching

Xiaomin Zhi<sup>1</sup>, Liang Gao<sup>2</sup>(✉), and Guohe Qin<sup>2</sup>

<sup>1</sup> College of Foreign Language, Jiangxi Agricultural University, Nanchang, Jiangxi, China

<sup>2</sup> College of Foreign Language, Shaoguan University, Shaoguan, Guangdong, China  
2567017976@qq.com

**Abstract.** In recent years, with the continuous development of economic globalization, people's exchanges at home and abroad have become more and more extensive, international Japanese teaching has developed rapidly, and the number of people studying Japanese at home and abroad is also gradually increasing. How to improve the quality of Japanese teaching and design auxiliary courses is a major problem to be solved urgently in the development of Japanese teaching career. Based on the related applications of immersion Japanese teaching, this paper studies the design of multimedia-assisted courses. Immersion teaching focuses on creating a target language environment, takes the course content as the teaching goal, and the second language as the teaching tool, which realizes the simultaneous acquisition of course knowledge and language skills. The advantages of immersion teaching mode help to solve the problems existing in traditional Japanese teaching. The research on immersion teaching has important theoretical significance and practical guiding significance for the development of Chinese teaching in Japan. Therefore, teachers should be proficient in various classroom teaching skills such as multimedia technology, and continuously improve their teaching level and ability in Japanese teaching. The final result of the research shows that when the number of participants in the platform five multimedia assisted courses is 67, the proportion is 20.6%. The number of users of multimedia-assisted courses in Japanese teaching is always larger than that of traditional courses, which indicates that the design of multimedia-assisted courses based on immersion Japanese teaching is feasible and has promotion significance.

**Keywords:** Immersion Teaching · Japanese Language Teaching · Multimedia Assistance · Curriculum Design

## 1 Introduction

With the rapid development of science and technology, multimedia technology based on computer and network technology has become an important tool for teaching and learning. The country's investment in language education is increasing, and Japanese classroom teaching has also rapidly reformed teaching methods and teaching methods [1]. Multimedia technology has been gradually popularized and applied in Japanese teaching classrooms in my country. Multimedia technology can make Japanese teaching

more intuitive, vivid, and more interesting. If multimedia technology can be reasonably applied in Japanese teaching, it can attract students' interest in learning foreign languages, effectively strengthen students' understanding of Japanese knowledge, and optimize students' knowledge of Japanese learning effect.

In recent years, many researchers have explored the research related to multimedia-assisted course design, and achieved good results. For example, Jong S Y believes that the previous studies on oral Japanese courses lacked the combination of quantitative and qualitative methods externally, and did not grasp the overall design framework of oral language courses internally [2]. Samavati T believes that the analysis of curriculum needs should highlight the needs of students, the design of curriculum objectives should comprehensively consider the particularities of the country, localities and institutions, and the selection of curriculum content and organizational design should be based on reality and pay attention to academic truth [3]. At present, scholars at home and abroad have carried out a lot of research on the application of multimedia-assisted curriculum design. These previous theoretical and experimental results provide a theoretical basis for the research in this paper.

Based on the relevant theoretical basis of immersion Japanese teaching, this paper systematically designs the current Japanese multimedia auxiliary courses, and has achieved relatively fruitful research results. Although the research of multimedia technology-assisted teaching in China started late, the development progress and speed are very fast, and the pertinence is stronger. There are not many specific studies on the application of multimedia in Japanese teaching courses, nor is it in-depth enough. However, multimedia technology is an important part of modern information technology, and its characteristics and achievable functions are very suitable for the needs of Japanese teaching.

## **2 Related Theoretical Overview and Research**

### **2.1 Research on Multimedia Assisted Course Design**

#### **(1) Multimedia Assisted Curriculum Development**

Multimedia technology plays a very important role in information technology. It is a technology that can process both text and data information, as well as images, audio, video and other media [4, 5]. It can change the teaching content from simple teaching materials to teaching materials, and cooperate with rich information resources. By using multimedia to present teaching content, it provides a variety of methods and ideas for analyzing and solving problems, making it concrete, simplified and easy to understand. This can not only help improve the teaching effect of teachers, but also improve the learning effect of students in the classroom.

#### **(2) Teaching Technology of Multimedia Courses**

Compared with traditional teaching, classroom education and teaching assisted by multimedia information technology breaks the conventional classroom teaching. It not



only activates the classroom teaching atmosphere, but also uses multimedia technology to create situational teaching to assist students in completing the knowledge system construction of the content they want to learn [6]. To a certain extent, abstract knowledge can be concretized, and through sound, light, color, shape, shadow and other senses of hearing and vision, it can give enough stimulation to make students feel immersed in the situation, so as to help break through teaching difficulties. It also broadens students' horizons; it also stimulates students' interest in learning and allows students to learn actively; its intuitive image features can reduce individual differences and take care of students with different cognitive levels; in this way, we can give full play to the potential of our teaching objects and enrich Teaching content improves its classroom efficiency.

### (3) Construction of Multimedia Auxiliary Course Environment

From the perspective of language subjects, it is not like mathematics, physics and other rational subjects that have complex formulas, principles and theorems, nor is it like politics, history and other liberal arts subjects that require long-term understanding and recitation of theoretical content [7]. Relatively speaking, the most prominent performance of language teaching is the context. Teachers can choose different multimedia technologies to assist teaching according to their different needs. Teachers can display language content intuitively by playing media such as videos and pictures. In the process of learning, help students to feel the different feelings brought by language more vividly, and cultivate students' sense of language [8]. Through various forms of multimedia technology, the scenes and stories described in the language can be restored almost completely, so that these languages can be presented to students more intuitively, and learning with background and content will achieve a multiplier effect.

The multimedia-assisted course design based on immersion Japanese teaching, in daily teaching, can be established by establishing exclusive classrooms for immersion classes. Teachers can arrange classrooms full of Japanese characteristics according to their own preferences, teaching needs, and students' hobbies [9]. At the same time, they will also post commonly used classroom terms, classroom rules, etc. in the form of Japanese text or pictures in a fixed position in the classroom to facilitate regular teaching and classroom management. In this way, students can be fully immersed in the Japanese environment. In this way, we can optimize the content of immersion teaching, help students build confidence, and arrange immersion teaching hours reasonably.

## 2.2 Theoretical Introduction to Immersion Teaching

From the perspective of different countries and regions, immersion teaching is developing all over the world, but North America has developed the most rapidly. As the birthplace of immersion teaching, its teaching concept is authoritative and its teaching model is relatively mature [10]. Immersion teaching aims to create a complete target language environment, allowing students to immerse themselves in the target language environment and acquire the second language naturally [11]. Its outstanding feature is that the second language (target language) in the immersion teaching is different from the traditional language teaching that regards the second language as a direct teaching goal, but acts as a language tool for teaching the content of various courses. The teaching

goal is based on the content of each subject (Mathematics, science, sociology, art) are the main ones, while mastering the knowledge and culture, the language skills of the second language are acquired subtly.

The core of immersion language teaching is implicit cognition. Compared with the learning of course content, the process of language acquisition is more implicit. It has similar characteristics of implicit cognition [12]. Mainly include: 1. The conscious state of language learning is in an unconscious state. In immersion teaching, language is the medium for learning course content, and it is automatically acquired along with knowledge learning; 2. From the perspective of whether there is a purpose, the knowledge of course content is the direct purpose, and language is the direct purpose. Learning is an indirect purpose; 3. The key point of acquisition lies in the understandable input of meaning, and the goal of acquiring the knowledge carried by the language is the intelligible input of the meaning of the target language; 4. The method of acquisition is mainly acquisition, but cannot be separated from it Learning; 5. The intelligibility generation method of meaning is mainly based on the construction of scenarios, the creation of tasks, and the method of solving problems; 6. The information processing method is procedural, and the input and output of knowledge is mainly based on real communication training; 7. The sense of language is in plays a leading role in language acquisition.

### 3 Experiment and Research

#### 3.1 Experimental Method

In the training process of the application and propagation path model in the Japanese course, the process influencing factors in the training sample of the teaching process working hours are input, and the calculation of the hidden layer neurons and the output layer neurons is performed to judge the predicted teaching hours and statistics obtained. Whether the error between the actual processing hours of the process meets the requirements:

$$\omega_i = \frac{\lambda \sum_{i=1}^n x_i}{n} \quad (1)$$

$$I_i = \sum_{j=1}^m w_{ij}x_j + \theta_i \quad (2)$$

In the above formula, F represents the total score, the data x represents the reference evaluation, the indicator w represents the weight, the coefficient  $\lambda$  represents the existing error, and the indicator n represents the total number. w is the weight between the input layer and the hidden layer of the initial work-hour prediction model.

### 3.2 Experimental Requirements

This experiment is based on immersion Japanese teaching, however, from the results of the survey and analysis, there are some errors in the current Japanese teachers' understanding of the status of multimedia education. For example, some experienced teachers may think that their teaching mode is summed up after years of teaching experience. Teachers will think that a set of relatively mature teaching methods has certain practicality, so they will reject multimedia to assist Japanese classroom teaching. Some young Japanese teachers have relatively high information literacy, but lack some teaching experience and need to use multimedia courseware to provide certain help. Therefore, young teachers will overemphasize the status of multimedia in the classroom, resulting in the abuse of multimedia, and students will feel the classroom. The larger the knowledge capacity, the lower the student's learning effect, but the poor student's learning effect.

## 4 Analysis and Discussion

### 4.1 Analysis of Multimedia Auxiliary Course Layout

The experiment is based on the design of the multimedia auxiliary course for immersive Japanese teaching. By testing the layout of the multimedia auxiliary course in each teaching and research platform, the experimental data is as follows:

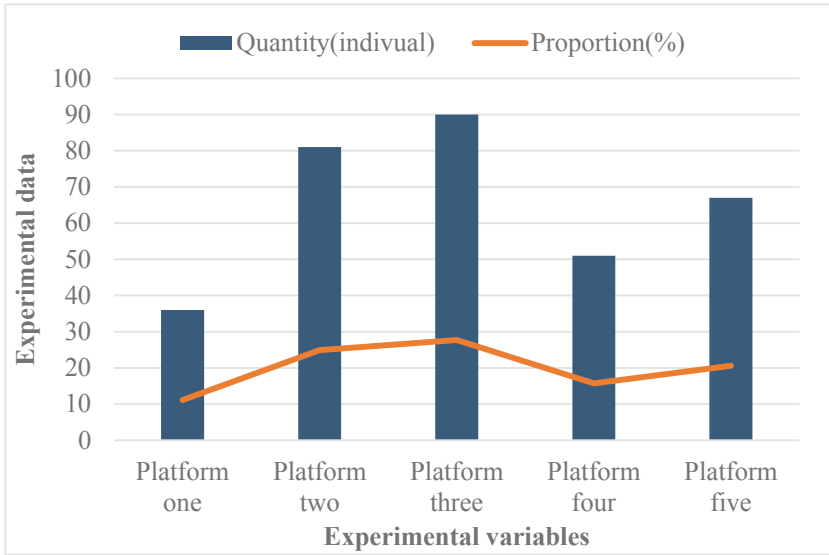
**Table 1.** Multimedia-assisted course layout analysis table

Item	Quantity (individual)	Proportion (%)
Platform one	36	11.1
Platform two	81	24.9
Platform three	90	27.7
Platform four	51	15.7
Platform five	67	20.6

It can be seen from the data analysis in Table 1 and Fig. 1 that, it can be seen from the results that when the number of participants in the platform 1 multimedia auxiliary course is 36, the proportion is 11.1%. When the number of participants in the platform 2 multimedia auxiliary course was 81, it accounted for 24.9%. When the number of participants in the platform three multimedia auxiliary courses is 90, the proportion is 27.7%. When the number of participants in the platform 4 multimedia auxiliary courses was 51, the proportion was 15.7%. When the number of participants in the platform five multimedia auxiliary courses was 67, the proportion was 20.6%.

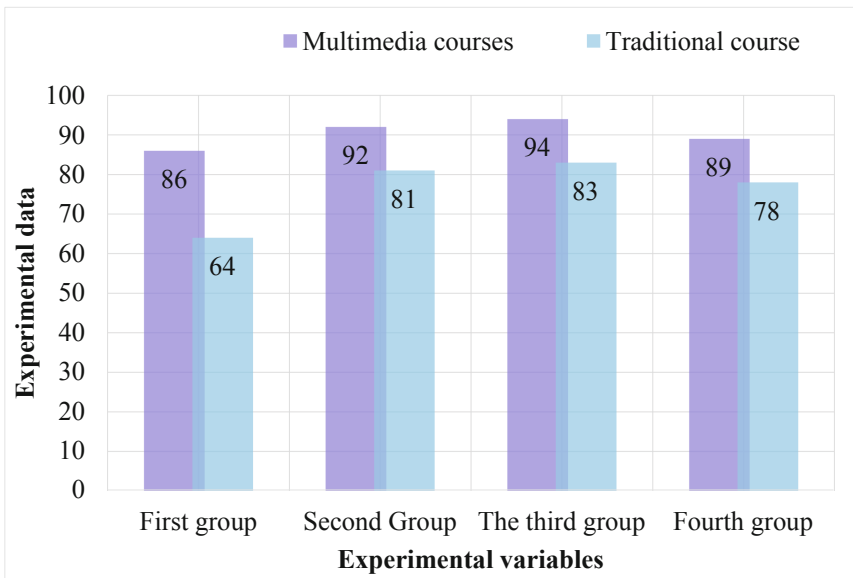
### 4.2 Analysis of the Use of Multimedia Auxiliary Courses

By analyzing the use of multimedia auxiliary courses in Japanese language teaching in colleges and universities, this experiment compares and analyzes the use of multimedia



**Fig. 1.** Layout analysis diagram of multimedia-assisted courses

auxiliary courses in Japanese teaching in four colleges and universities and the use of traditional courses. The experimental data is shown in the figure below:



**Fig. 2.** Analysis of the usage of multimedia-assisted courses

As shown in Fig. 2, through the data analysis of the use of multimedia auxiliary courses in Japanese teaching in colleges and universities, it can be seen that the number of participants in the multimedia auxiliary courses of Japanese teaching in four groups of colleges and universities is 86, 92, 94 and 89 respectively. The number of users of traditional courses was 64, 81, 83 and 78 respectively. In the data of the four groups of colleges and universities, the number of users of Japanese teaching multimedia auxiliary courses is always more than that of traditional courses, indicating that the design of multimedia auxiliary courses based on immersion Japanese teaching is feasible and has promotion significance.

## 5 Conclusions

Based on the research background of immersion Japanese teaching, this paper first studies the design and application of multimedia auxiliary courses, and in the experiment of analyzing the use of multimedia auxiliary courses in colleges and universities, the number of users of multimedia auxiliary courses in Japanese teaching is always larger in the data of the four groups of colleges and universities. In terms of the number of users of traditional courses, it means that the design of multimedia-assisted courses based on immersion Japanese teaching is feasible and has promotion significance. Practical research should further explore the specific improvement path of Japanese teaching mode and learning mode, further exert the auxiliary role of network technology and mobile technology in course teaching and learning, further innovate specific auxiliary methods, and strengthen the management and control of auxiliary effects. In terms of research methods, it integrates the three perspectives of students, teachers and curriculum environment, improves the position of qualitative research methods and action research methods in curriculum design research, and combines quantitative research methods.

## References

1. Abbas, S.: Development of e-learning oriented inquiry learning based on character education in multimedia course. *Euro. J. Educ. Res.* **9**(21), 40–74 (2020)
2. Jong, S.Y., Chen, G., Tam, V., et al.: Design-based research on teacher facilitation in a pedagogic integration of flipped learning and social enquiry learning. **5**(2), 45–53 (2022)
3. Samavati, T., Farvardin, M.: Efficient deep learning-based estimation of the vital signs on smartphones. **3**(1), 12–59 (2022)
4. Davis, T., Frederick, T.V.: The impact of multimedia in course design on students' performance and online learning experience: a pilot study of an introductory educational computing course. *Online Learn.* **24**(3), 1–22 (2020)
5. Ruan, R.: Design of the in-depth intelligent learning system based on the college English teaching. **5**(1), 21–87 (2019)
6. Si-Yu, H.E., Zhao, Y., Humanities, S.O.: Research on the teaching design of cultural experience course for international students based on the regional culture of eastern Guizhou—taking Tongren polytechnic college as an example. *Heilongjiang Sci.* **5**(15), 9–32 (2019)
7. Thohir, M.A., Sukarelawan, M.I., Jumadi, J., et al.: The effects of instructional design based web course on pre- service teachers' competencies. *Int. J. Eval. Res. Educ. (IJERE)* **10**(1), 230–236 (2021)

8. Sjøberg, J., Brooks, E.: Discourses of digital game based learning as a teaching method. In: International Conference on Design, Learning, and Innovation, vol. 2, no. 1, pp. 3–18. Springer, Cham (2021)
9. Judge, P.K., Buxton, J.A., Sheahan, T.C., Phetteplace, E.R., Kriebel, D.L., Hamin Infield, E.M.: Teaching across disciplines: a case study of a project-based short course to teach holistic coastal adaptation design. *J. Environ. Stud. Sci.* (2), 12–36 (2020). <https://doi.org/10.1007/s13412-020-00610-z>
10. Ma, J.W.: The construction of website of English excellent course based on the blackboard content design. *52882–890* (2019)
11. Sung, H.Y., Chen, S.H.: The screen shows movement – movement is interesting!" exploring effects of multimedia stories on preschool children's story comprehension and enjoyment. *Libr. Hi Tech.* **37**(5), 12–36 (2019)
12. Kim, S.A., Cho, G.P.: Teacher efficacy, subjective happiness, school organizational culture of high school teachers, effects on teaching immersion. (3), 1–10 (2019)



# Analysis and Research of Intelligent Distribution UAV Control System Based on Optical Flow Sensor

Shuai Sun<sup>1</sup>(✉), Yanqi Han<sup>2</sup>, Yue Li<sup>2</sup>, Wanting Zhao<sup>2</sup>, Qifeng Gao<sup>2</sup>, Yang Yang<sup>2</sup>, and Heng Li<sup>2</sup>

<sup>1</sup> Engineering Training Center, Jilin University, Changchun, China  
ss88311@jlu.edu.cn

<sup>2</sup> School of Mechanical and Aerospace Engineering, Jilin University, Changchun, China

**Abstract.** UAV (Unmanned Aerial Vehicle) delivery is one of the solutions for logistics companies to improve delivery efficiency in recent years. The GPS and optical motion capture or other positioning methods currently used by distribution UAVs have some issues such as high cost and relatively fixed operating environment. The UAV system based on optical flow sensor can realize the functions of autonomous positioning, real-time obstacle avoidance and hover correction with high precision and low cost, which can solve the shortcomings of traditional positioning methods and can be widely used. Many scholars have already researched UAV systems using optical flow sensors. We summarize the results of previous research on these systems and analyze them in the hope that they can provide development ideas for UAV delivery.

**Keywords:** Optical Flow Sensors · Smart Delivery · Distribution UAVs

## 1 Introduction

In recent years, China's e-commerce industry has developed rapidly, courier business volume has risen sharply. The 2022 National Postal Management Conference pointed out that China's express business volume in 2021 was 108.5 billion pieces and is expected to be 122.5 billion pieces in 2022, which will increase by about 13% year-on-year. However, Chinese logistics service industry shortcomings are serious, it is difficult to meet the development needs of the e-commerce industry, the city's last mile and other issues still need to be resolved.

Compared with traditional delivery methods, delivery drones take up less space, and have a higher degree of intelligence, are more flexible and have more mature technology, which can have reduced transportation costs and improved delivery efficiency, so they have also gained the favor of many logistics. As early as 2013, the U.S. company Matternethad tested the drone delivery network. In China, JD, SF Express and other companies also carried out tests of delivery drones in Suqian, Ganzhou and other places respectively.

Although delivery UAVs have many advantages, the problems faced in delivery such as gathering flight and land in high-density buildings and targeted delivery to high-rise households and remote areas are still pain points in development. To address these problems, this paper researches and analyzes the control system of intelligent delivery drones based on optical flow sensors [1, 2].

## 2 Concept and Application Examples of Distribution UAVs

Delivery UAVs (UAV Express) are unmanned low-altitude vehicles that can carry packages and automatically deliver them to their destinations using radio-controlled equipment and self-contained program control devices. Most delivery UAVs are equipped with a variety of high-precision sensors and advanced control algorithms and have a black box to record their status information. Most drones also have a runaway protection function that automatically maintains a precise hover when the drone goes out of control to ensure the safety of the UAV.

The automated drone courier system utilizes UAVs to replace manual courier delivery, aiming to automate, unmanned and informatized courier delivery and improve courier delivery efficiency and service quality in order to alleviate the contradiction between courier demand and courier service capacity. At present, numerous companies have started to try to apply delivery drones.

### 2.1 Amazon UAV Delivery Express

Amazon made its first commercial drone delivery flight in the countryside and suburbs of Cambridge, England, after receiving permission from the U.K. Civil Aviation Authority 2016 in 2007, and it took a total of 13 min from the time of the customer completed the order to the time the goods were received, with the entire process requiring no human control and the drone returning automatically after completing the delivery is shown in the Fig. 1.



**Fig. 1.** Amazon UK's first single drone delivery courier



## 2.2 Swift Ant UAV and China Post Joint

On September 19, 2016, the domestic startup company Swift Ant UAV and China Post Zhejiang Anji Branch jointly opened China's first drone express mail route. This express mail route using UAV delivery can reduce both delivery time and cost to half of the original. The UAV has a built-in self-developed intelligent control system, which can achieve fully automatic flight path and accurate landing through high-precision sensors and machine vision. The flight speed and angle will be automatically adjusted according to the wind speed and weather conditions during the flight, and the whole process only requires the staff to monitor the flight status and position of the drone through the cloud system, which largely realizes the unmanned delivery is shown in the Fig. 2.



**Fig. 2.** China Post uses drones for delivery

## 2.3 Initial Applications in Emergency Blood Distribution

In the emergency situation where medical resources are insufficient, traffic is inconvenient and time dependence is very strong, the use of drones can achieve rapid delivery of emergency equipment, drugs, blood and various specimens within the county, and the medical security of large events, medical emergencies of emergencies, etc. can be used to efficiently “grab time” by UAVs. For example, in the delivery of Zhejiang II Binjiang, the one-way land transportation time is 10 min, the UAV is 6 min, which shows a better timeliness. The time efficiency is good for the hospital round-trip emergency blood collection mode, the overall timeliness of drone delivery is significant, and is conducive to



**Fig. 3.** Zhejiang Blood Center uses drones to deliver blood

hospital manpower and transportation cost accounting, and the temperature fluctuation of blood component has no significant impact on the medical effect is shown in the Fig. 3.

### 3 Research and Analysis of Intelligent Delivery Drones Based on Optical Flow Sensors

#### 3.1 Optical Flow Positioning and Optical Flow Sensors

The concept of optical flow was first introduced by biologist Gibson in 1950. When we look at a moving object, the external environment creates a series of continuously changing images on the retina. These images “flow” from the front of the retina (imaging plane) as if light were “flowing through”, hence the term “optical flow”. Optical flow is the apparent motion of the image with constant grayscale values of the image luminance pattern, which contains information about the motion of the pixels in the image, including velocity vectors and coordinate positions. Since the optical flow contains the approximate information of the target motion, it can be used by the observer to determine the approximate motion of the observed target and thus achieve the basic localization function of the moving target. Optical flow localization is usually performed with the help of an image acquisition device (e.g. a camera), which captures images of the surrounding environment. Assume that  $I(x, y, t)$  is a pixel point  $(x, y)$  at  $t$  time, the pixel point has moved on two frames of the image,  $\delta y, \delta t$ , since the luminance information of the same point is constant, the following equation can be obtained.

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t) \quad (1)$$

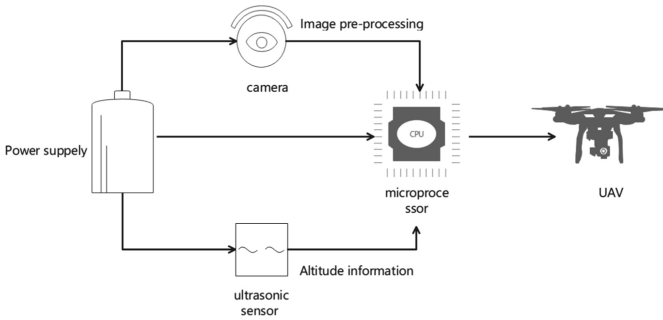
Based on this equation, the instantaneous velocity and direction of movement of each pixel point of the image can be calculated (Lucas-Kanade algorithm). These instantaneous velocities constitute a two-dimensional instantaneous velocity field called the optical flow field. The computed optical flow field vector enables precise positioning of the observed target and more accurate and smooth control of the target, as well as attitude control of the target object.

The optical flow sensor is a device that detects the movement of an object and its position by capturing images of changes in the target operating environment and inputting pixel movement information to the main control board. The conventional optical flow sensor consists of four parts: power supply, camera, ultrasonic sensor, and microprocessor, and its working principle is shown in the Fig. 4.

#### 3.2 UAS Based on Optical Flow Sensor

Autonomous positioning and real-time obstacle avoidance system

The UAV positioning problem mainly refers to the use of its own sensors to determine the position and attitude information of the UAV in the flight environment relative to the inertial coordinate system. The widely used positioning method is mainly based on the Global Positioning System (GPS) [3]. The principle of autonomous positioning is different from that of GPS positioning, GPS sends the position information of the UAV in



**Fig. 4.** Schematic diagram of the working principle of optical flow sensor

the BeiDou coordinate system to the UAV, which is an absolute coordinate positioning. Autonomous positioning based on airborne equipment does not rely on the assistance of external information, and it cannot obtain absolute position information, but can only obtain relative position information of the UAV relative to the environment by capturing environmental information through airborne sensing equipment. If environmental information (e.g., environmental maps) is available, relative positioning can be achieved relatively easily. However, in unknown complex environments, UAVs need to use simultaneous localization and map building method (SLAM) to recover the environmental structure during motion, while using the recovered environmental structure to estimate their own positional information. The accurate attitude estimation is the prerequisite and foundation for achieving complex flight tasks such as UAV obstacle avoidance, trajectory planning and target tracking, while obstacle avoidance capability is particularly important for UAVs, and UAVs obstacle avoidance systems play an increasingly critical role in UAV applications [4].

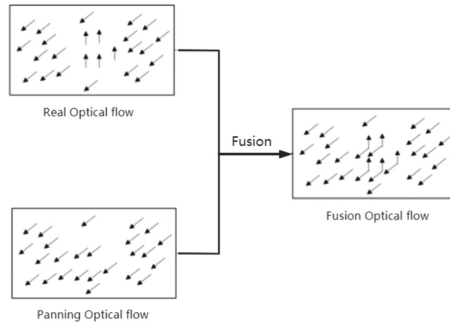
SLAM is widely used in various fields. Among them, in the field of UAVs, SLAM can be used to build local 3D maps quickly and can combine with geographic information systems (GIS) and visual object recognition technologies, which can assist UAVs to identify roadblocks and plan paths for automatic obstacle avoidance. The sensors currently used in SLAM are mainly divided into two categories: Lidar-based laser SLAM (LidarSLAM) and vision-based VSLAM (VisualSLAM).

MeihuiCao et al. of Tianjin University used optical flow sensors to obtain the velocity and position information of UAVs for autonomous control of UAVs under the loss of visual simultaneous localization and map building (SLAM) to achieve autonomous flight control of quadrotor UAVs. Although the visual SLAM algorithm has high accuracy, its dependence on map information in the environment is strong, and it is easy to cause the loss of map information in a structured environment [5]. In order to ensure the autonomous control of man-machine in the case of visual SLAM loss, Meihui Cao et al. introduced optical flow sensor as an auxiliary position information measurement unit and detected the matching ratio of feature points of visual SLAM algorithm by controlling the program: when the matching ratio is lower than 30%, it will switch to optical flow method to control automatically; and when the matching ratio is higher than 30%, it will switch to visual SLAM again. In the experiment to verify the flight control under the loss of visual SLAM map, the position control error of the UAVs automatic control

direction is about  $\pm 0.1$  m, which proves the effectiveness of the optical flow sensor as an auxiliary position information measurement unit under the loss of visual SLAM.

On this basis, ZhengkangJinet al. from Jiangnan University used the Lucas-Kanade based optical flow algorithm to achieve UAV positioning control in environments that are not covered by GPS signals, such as indoor environments, greatly reducing the positioning error due to time accumulation with the help of traditional inertial navigation systems only. The Lucas-Kanade algorithm is the most common and popular optical flow algorithm. Since it is based on the Taylor series of image signals, this method is called differencing, which is the use of partial derivatives for both spatial and temporal coordinates.

A new monocular vision quadcopter obstacle avoidance method based on the fusion of pyramidal LK (Lucas-Kanade) and translational optical flow is proposed by Hai Zhao et al. at Northeastern University. Supposing there are relatively moving objects in the scene, as shown in Fig. 5. The upper left part of the figure shows the real optical flow generated by the pyramid LK, where the vertical upward optical flow is the suspected obstacle, and the lower left part shows the translational optical flow derived from the single-strain transformation, and the fused optical flow is formed by superimposing the two onto the same image.



**Fig. 5.** Real optical flow and panning optical flow fusion

The angle of the fused optical flow on the obstacle (the offset between the pyramidal LK optical flow and the translational optical flow) is obviously different from the other unobstructed parts, and the characteristics of one fused optical flow can be deduced by comparing the offsets. Therefore, by establishing the correspondence between the fused optical flow and the obstacle, it is possible to determine the obstacle.

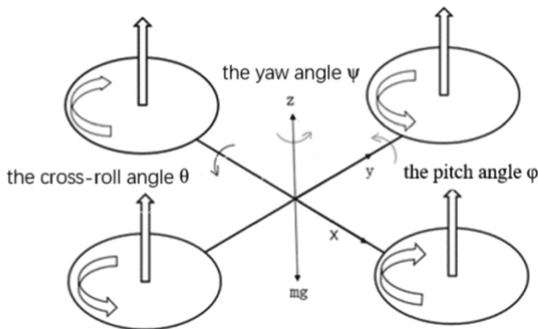
Using the above optical flow algorithm, UAVs can basically achieve autonomous positioning and obstacle avoidance.

### 3.3 Hover Correction System

The stability of the aircraft hovering in the air is an important indicator of the performance of this aircraft, and the main factors affecting the hover control are the low cost, the drift of the laired measurement unit itself and the error generated by the mechanical structure

of the airframe itself. And this error often leads to the vehicle can't hover at the specified position and height. In order to solve the problems of aircraft hovering, scholars have conducted the following research.

Using the PX4FLOW optical flow sensor and taking advantage of its high sensitivity, high pixel count, high data update speed, and the ability to acquire data in real time through the ground station software QGround Control, Qiang Lu et al. proposed a method to detect the horizontal movement speed of a quadrotor relative to the ground based on this sensor and use the MTI-G high-precision inertial measurement unit is used to compensate for the attitude angle of the quadrotor, and finally realize the air hover correction of the quadrotor. During the flight of the quadrotor UAV,  $\varphi$  is the pitch angle,  $\theta$  is the cross-roll angle and  $\psi$  is the yaw angle, as shown in Fig. 6. The attitude control of the quadrotor in flight is the control of  $\varphi$ ,  $\theta$  and  $\psi$ . The hover control does not require high accuracy of yaw angle  $\psi$ . The  $\varphi$  and  $\theta$  acquired by MTI-G have some error with the actual angle, so it will produce a slow drift in a certain direction. The optical flow sensor can detect the speed of horizontal movement, which can compensate the attitude angle measured by MTI-G. Experiments show that the control algorithm can increase the attitude angle compensation according to the velocity feedback to overcome the drift phenomenon. The experimental analysis shows that the optical flow sensor-based hover correction method can effectively overcome the drift of the measurement unit itself and the error generated by the mechanical structure of the fuselage and achieve stable hover control of the quadrotor.



**Fig. 6.** Schematic of pitch angle  $\varphi$ , cross-roll angle  $\theta$  and yaw angle  $\psi$  during the flight of quadrotor UAV

In response to the hovering instability of a miniature quadcopter caused by the larger drift error of the low-cost inertial sensor measurement unit, Jie Xia et al. from Wuhan University of Engineering proposed an optical flow sensor-based fixed-point and fixed-height hovering correction system. Of these, the optical flow sensor PMW3901 is used to collect the displacement of the aircraft in the x and y axes during hovering with respect to the initial position as the error term and input to the corresponding series PID control to correct the aircraft's fixation error. Table 1 shows a comparison of the parameters before and after the PMW3901 was involved in the fixing correction.

**Table 1.** Comparison of parameters before and after PMW3901 participation in fixed-point calibration

Parameter	Variable	
	Before PMW 3901 participates in fixed-point correction	After PMW 3901 participates in fixed-point correction
Average offset pixel	x: > 100 y: > 100	x: 48.1175 y: 47.1177
RMSE	No valid estimate	x: 0.1052 y: 0.1026
var	No valid estimate	x: 5.424 0e-04 y: 4.357 5e-04
Flight performance	The relative fixed-point drift is random, and the drift distance is large, so that PMW 3901 cannot be measured	The aircraft can basically realize fixed-point flight without obvious deviation

The experimental data shows that the PMW3901 optical flow sensor can effectively solve the hovering instability problem of the miniature quadcopter after participating in fixed point correction under the specified working conditions, and the vehicle can basically achieve fixed-point flight with strong stability.

### 3.4 UAS Applications Based on Optical Flow Sensors

The traditional UAV positioning systems use two types of systems, GPS positioning system and optical motion capture system. The GPS system is suitable for outdoor open areas. Once indoors or in places with many buildings, the accuracy of positioning will be affected due to signal blocking and interference. The GPS spatial positioning accuracy is 1–2 m, which cannot reach the level of accurate positioning. Optical motion capture is suitable for indoor positioning, which requires cameras to be installed in all corners of the room to obtain the motion information of the UAV and feed it back to the drone, thus achieving the purpose of positioning. This type of positioning is more costly and only suitable for indoor use [6–8]. While systems equipped with optical flow sensors are able to run in unknown and GPS-free environments, such as indoor, caves, tunnels, etc. Compared to conventional sensor systems, camera-based optical flow sensing systems are widely used due to their accurate positioning, lighter mass, lower cost and smaller size [9]. The research on optical flow and optical flow sensors is relatively mature now, and the above-mentioned UAS equipped with optical flow sensors have achieved good results in their applications.

Among them, the UAV autonomous positioning and real-time obstacle avoidance system equipped with optical flow sensors can achieve better results with or without GPS. Yong Pang et al. constructed a Marine rescue system based on quadcopter by realizing the autonomous positioning of UAV through the combination of GPS and optical flow sensors. Zhou Zhou et al. used optical flow sensors to calculate the image

displacement frame by frame, which in turn achieved high accuracy positioning of the UAV in the absence of GPS signals indoors and achieved smoother control. Chaofan Yu et al. used the optical flow sensor and the improved potential field method to calculate the relative motion of the UAV and obstacles to plan the optimal flight path of the UAV, and successfully achieved the autonomous obstacle avoidance flight of the UAV. And ZengYou Han et al. used data from optical flow sensors and IMU modules to achieve accurate estimate of UAV velocity and position when the flight environment was unknown.

In terms of hovering control systems based on optical flow sensors, Wuyang Zhang et al. from Shanghai University of Engineering and Technology used optical flow sensors in combination with ultrasonic sensors to achieve autonomous fixed-point hovering of a small quadrotor UAV in a GPS-free environment and achieved good control results, and ArreolaL et al. also achieved low-cost hovering by fusing dense optical flow algorithms, GPS and inertial components [10].

## 4 Development Prospects

In recent years, the development of the UAV industry has received a huge boost thanks to both policy and funding. With the integration of new technologies, new products with various functions will continue to emerge, and the potential and value of intelligent delivery UAV will surely be gradually explored, even overturning the traditional logistics industry pattern.

In the future, intelligent delivery drones will develop in a more professional, scientific and efficient direction. At the national level, unified industry standards and a perfect delivery system will be developed. At the social level, publicity will be strengthened so that the public can understand and accept the convenience and efficiency of UAV delivery; At their own level, the software and hardware environment will be optimized to address their own limitations. Although UAV delivery still suffers from many constraints, UAV delivery will definitely bring a major contribution to social development, as long as we seize the opportunity to develop properly.

**Acknowledgment.** Cooperative Education Program of the Ministry of Education (220504011055045).

## References

1. Patruno, C., Nitti, M., Petitti, A., Stella, E., D’Orazio, T.: A vision-based approach for unmanned aerial vehicle landing. *J. Intell. Rob. Syst.* **95**(2), 645–664 (2019)
2. Wubben, J., Fabra, F., Calafateet, C.T., et al.: Accurate landing of unmanned aerial vehicles using ground pattern recognition. *Electronics* **8**(12), 1532 (2019)
3. Wu, Z., Jie, C., Wu, J., Wang, Y., Liu, C.: Detecting genuine communities from large-scale social networks: a pattern-based method. *Comput. J.* **57**(9), 1343–1357 (2013)
4. Wang, L., Liu, X., Wang, H.X., Wang, X.B.: Research on obstacles avoidance for UAV SLAM. *Appl. Mech. Mater.* **3485**(635–637:1329–1334) (2014)

5. Grabe, V., Bulthoff, H.H., Giordano, P.R.: Nonlinear ego-motion estimation from optical flow for online control of a quadrotor UAV. *Int. J. Robot. Res.* **34**(8), 1114–1135 (2015)
6. Lucas, B.D.: An iterative image registration technique with an application to stereo vision (darpa), *Proc Ijcai* **81**(3), 674–679 (1981)
7. Duchon, A.P., Warren, W.H.: Robot navigation from a gibsonian viewpoint. In: *Proceedings of IEEE International Conference on Systems, Man and Cybernetics, IEEE* (2002)
8. Serrano, F.J., Moreno, V., Curto, B., Lves, R.: Semantic localization system for robots at large indoor environments based on environmental stimuli. *Sensors* **20**(7), 2116 (2020)
9. Gageik, N., Strohmeier, M., Montenegro, S.: An autonomous UAV with an optical flow sensor for positioning and navigation. *Int. J. Adv. Rob. Syst.* **10**(10), 341 (2013)
10. Arreola, L., Montes, D., Flores, A., Sanchez, J., Flores, G.: Improvement in the UAV position estimation with low-cost GPS, INS and vision-based system: application to a quadrotor UAV. In: *Icuas18 the International Conference on Unmanned Aircraft Systems*, pp. 1248–1254 (2018)





# Method of Improving Logistics Distribution Efficiency by Combining Small Multi-rotor UAV with the Visual Recognition System

Shuai Sun<sup>1</sup>(✉), Weilun Xia<sup>2</sup>, Qiankun Zhu<sup>2</sup>, and Zeping Han<sup>2</sup>

<sup>1</sup> Engineering Training Center, Jilin University, Changchun, China  
ss88311@jlu.edu.cn

<sup>2</sup> School of Mechanical and Aerospace Engineering, Jilin University, Changchun, China

**Abstract.** Analysis of the Application of Small Multi-Rotor UAV in Logistics Distribution, Combining the Visual Recognition with Small Multi-Rotor UAV, Improving Multi-Rotor UAV Survey Capability through Visual Recognition System, put forward several logistics distribution schemes for multi-rotor UAV, solve the “Last-mile” problem of traditional logistics

**Keywords:** UAV · Visual Recognition · Logistics Transportation

## 1 Introduction

With the development of computer science and internet technology, UAV is playing a more and more important role in people’s life. It can be used in many aspects such as transportation and cargo transportation, agriculture, forestry and crop cultivation, and military fields to a large extent.

Nowadays, UAVs are very effective in many industries, especially in logistics industry and E-Commerce Platform. Base on the existing platform system, the E-Commerce Platform complete this series of processes with the support of large data, robotic sorting and, UAVs distribution. The usage of UAV in transportation industry mainly effect of cooperation between cargo transportation and branch research. By using this, it can improve the recognition ability in those remote area effectively. SF Express Company and JD Logistics Distribution Center have played the role of UAV well, which not only reduces the distribution costs, but also improves the efficiency of freight distribution. Small multi rotor UAVs can take off and land vertically or hover with the powerful maneuver ability. It is mainly applicable to low altitude, low speed, vertical take-off, landing and hovering tasks. Therefore, it can be applied in the distribution of most scenarios [1]. Small multi rotor UAVs can be remotely controlled by radio equipment or cruise at a fixed point by relying on its own geographic information, which greatly improves transportation efficiency. At the same time, aiming at the distribution problems in some areas with poor road conditions, small multi-rotor UAV has great application prospects and excellent application potential. In this paper, the visual recognition system based on OpenMV will be combined with several new transportation schemes of small

multi-rotor UAVs to solve the “last mile” problem which perplexes the logistics and transportation industry.

## 2 Brief Introduction About Advantages of Small-Multi Rotor UAV Cooperating with Visual Recognition in the Distribution Process

### 2.1 Introduction of Visual Recognition Landing System Based on OpenMV

Traditional UAV landing positioning technology depends on inertial navigation system and global positioning system. However, with the development of UAV technology and the improvement of UAV technical requirements, there are some problems in navigation, such as accumulated errors of inertial components, too sensitive to initial values, and GPS not always available. The traditional navigation accuracy can no longer meet the requirements of the precise landing of UAVs.

Computer visual navigation is beginning to cut a striking figure of science and technology. Computer vision technology uses installed cameras to obtain digital image information from the external environment and uses computer algorithms to analyze the effective information of the acquired image, such as shape, scale, etc. The UAV runs and arrives at its destination based on this valid information. At this stage, there is a UAV location method based on QR code recognition. The style of the April tag (as shown in Fig. 1) is similar to a QR code and is not complicated. The processor can reduce the amount of calculation required and satisfy the real-time requirements, so the accuracy landing is greatly improved [2].

The main process of UAV loaded with such devices is to use GPS to let the UAV land close to the charging base station, using the visual device to identify the April label icon on the charging base station, and then realize fixed-point landing.



Fig. 1. April tag's family

The main process of UAV loaded with such devices is to use GPS to let the UAV land close to the charging base station, using the visual device to identify the April label icon on the charging base station, and then realize fixed-point landing.

When the UAV reaches the location specified by GPS, GPS locates the first ground wire near the target point and turns on the OpenMV camera. After successfully receiving the OpenMV startup flag, the UAV enters the self-stabilizing mode (exit if the number of retries is exceeded) [3]. In self-stabilizing mode, the throttle and angle control of the UAV are not locked and can be controlled flexibly. The system analyzes the serial port processing function, and PID calculation adjusts the UAV attitude in time. After landing successful, turn off the motor to complete landing. If there is an error in data reception during this process, then request OpenMV to reinitialize. Through this process, the UAV can automatically land at a fixed point (refer to Fig. 2 for the detailed process).

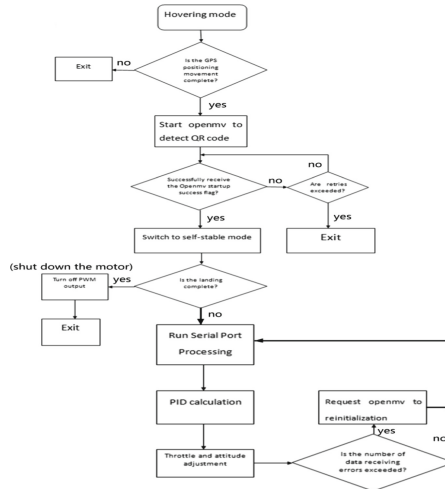


Fig. 2. UAV fixed-point automatic landing process

## 2.2 Advantages of the Combination of UAV and Visual Recognition

- 1) Higher positioning accuracy: Compared with ordinary GPS, the positioning accuracy of UAV equipped with visual recognition is higher, even reaching to millimeter level.
- 2) Faster recognition speed: The visual recognition system has the characteristics of high efficiency and high degree of automation. It also has a very fast response speed and can achieve a very high recognition speed.
- 3) Stronger security performance: The UAV equipped with visual recognition system can maintain a safe distance when detecting the target, which greatly improves the overall safety and prevents the occurrence of damage.
- 4) Stronger endurance and stability: The UAV equipped with visual recognition system can work stably for a long time because it improves the charging efficiency.
- 5) More detailed software control: The location, batch and shelf life of inventory goods can be carefully managed by using back-end management software such as intelligent warehouse management platform.

## 3 Several New Automatic UAV Distribution Schemes

### 3.1 Star Structure Distribution Scheme

In this section, we will introduce a simple automated UAV distribution scheme. This scheme is inspired by the star structure in the computer networks and the method ants transport food in nature (as shown in Fig. 3).

Detailed process: When the goods arrive at the first-level transfer station (the initial unloading transfer station when the express arrived in a city), the detection system immediately detects the information in the QRcode on the goods, judges the delivery location, and plans a nearest route through the internal intelligent algorithm. Before

distribution, the UAV performs self-inspection. If there is no abnormality, the UAV will send it to various secondary transfer station (some medium-sized transfer stations in the city). After the goods are delivered, the time data is written in the QR code of the goods, the storage area is addressed using the visual recognition system, and the goods are temporarily stored in chronological order. After that, the UAV will be tested for remaining power. If the power is sufficient, it will directly return to the first transfer station. If the power is insufficient to support the return, it will use the positioning system and the visual recognition system to find the wireless charging pile and stop accurately to complete the charging. The equipment self-inspection and process are shown in Fig. 4.

The secondary transfer station assigns tasks according to the UAV equipment. At this stage, the process is the same as above, and the UAV sends a message to the recipient to prompt the delivery. After completing the overall dispatch process, the UAV returns to the secondary transfer station.

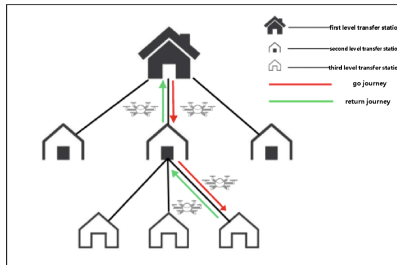


Fig. 3. Overall layout of Star structure Distribution

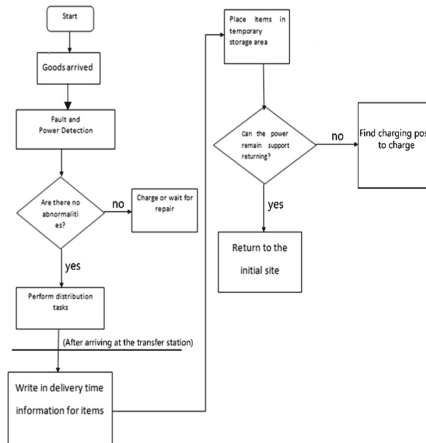


Fig. 4. Schematic diagram of UAV status during transportation

### 3.2 Truck Delivery and UAV Distribution Scheme

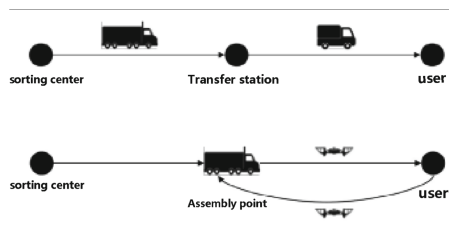
In general, end-of-line distribution will set up first-level distribution stations in major regions. According to actual needs, set down market segments and multi-level distribution stations will be set up. Trucks will depart from sorting center to various distribution sites and then distribute downward from these sites. This section will present a method to improve traditional delivery by using trucks and UAVs to deliver multiple packages [4].

There are now randomly distributed customer demand points within a defined flat area. Considering the limitation of the UVA maximum load capacity and flight distance, replace multi-level distribution stations with regional optimal nodes, and complete all distribution tasks with UAVs instead of distributors. For this reason, we make the following assumptions:

Trucks have enough capacity to carry packages and small multi-wing UAV equipment, all the UAVs are equipped with Prime Air, with a maximum load of 2.3 kg, the full working radius is 8 km.

The UAV can only be released and retracted at the optimal assembly point but cannot take off and land on a moving truck.

Without calculating the loading, maintenance, and launch time of the UAV, the default UAV unit can return to the charging point of the truck through the visual recognition system and GPS, then they are able to complete the next mission (Note: No-fly restrictions due to weather, policies, laws and regulations are not considered) Starting from the first assembly truck, all distribution tasks are completed and returns to the first assembly point to form a closed loop are shown in Fig. 5.



**Fig. 5.** Schematic diagram of truck transport method

There are many advantages: By using the truck to transport UAVs and using UAVs to deliver goods can reduce a large number of construction costs associated with transit stations. Users are classified by using intelligent algorithms, then assemble, and finally the optimal delivery route is set using the dynamic planning system. It can reduce a lot of labor costs and save the total delivery time. Moreover, installing wireless charging devices on a specific section of the road will greatly improve its endurance. The UAV carries a visual identification system, which will make it possible for the UAV to deliver continuously.

## **4 Development Trend and Prospect of Small Multiwing UAV in Distribution**

### **4.1 Development Trend of Small Multiwing UAV'S Distribution**

The combination of electronic fence technology and logistics cloud monitoring system: Electronic fences provide no-fly zones for UAVs, and logistics cloud monitoring systems provide routes, weather, and other services for UAVs. At the same time, real-time data collected from the UAVs will be transmitted to the airline company as a supervisory authority to supervise the logistics UAVs. In the future, the regulatory cloud systems will be open and shared among different logistics companies. Through the monitoring system, application of those data could be directly from the air traffic control department. After the application is approved, the cloud system plans the route according to the actual situation. UAVs from different logistics companies can access monitoring cloud systems for information [5].

- 1) By using visual recognition system, the logistics and distribution UAVs have a longer life, more accurate delivery, and more efficient operation.
- 2) In the future, the logistics UAVs can be charged and packaged by using the “UAV + Mobile UAV Center” instead of being limited by endurance and load. Through the tag system and the visual recognition system, package could be delivered accurately and effectively [6], which solves the “last-mile” problem of UAV logistics really, accelerates the integration of policy and technology.
- 3) At present, there are no applicable laws and regulations for UAV logistics and distribution, and most UAVs operations still follow the general aviation standards. “Unmanned Distribution” has never been a simple technical issue, but a systematic issue that requires the cooperation of the government, upstream and downstream industry chains, and all sectors of society. Therefore, we should actively carry out policy research, standardize UAV service and formulate targeted regulatory policies [7]. At the same time, we would research and issue corresponding industrial policies to support production, sales, operation, and other aspects, and accelerate the integration and innovations of policies and technology [8].

### **4.2 Prospects for Small Multi-rotor UAV Combined with Visual Recognition System in Delivery**

In recent years, people have seen the “magical power” of UAV from live TV broadcasts many times. Despite in the situation that complete disruption of power, transportation and network, the unmanned aerial vehicles hovering over the disaster area can always bring back the local disaster situation, landform and other conditions at the first time. It not only allows the masses to intuitively understand the situation of the people in the disaster area, but also provides a large number of reliable basis of the relevant departments to make decisions in the disaster relief and rescue work [9]. Now, the people’s travel and driving route navigation, rainfall probability forecast, pollution index prompt and other big data applications from UAVs have become normal services that available to everyone [10].

Additionally, using aerial views, public entertainment can be enhanced. It is generally believed that under the guidance of the country's policy of opening up the low-altitude flight, the area of unmanned aerial vehicle application will be further expanded.

**Acknowledgements.** Cooperative Education Program of the Ministry of Education (220504011055045).

## References

1. Salama, M.: Joint optimization of customer location clustering and drone-based routing for last-mile deliveries. *Trans. Res. Part C Emerg. Technol.* **114**(C), 620–642 (2020)
2. Shu, P.: Data-driven shuttle service design for sustainable last mile transportation. *Adv. Eng. Inform.* **49**, 101344 (2021)
3. Feng, Q., Liu, J., Gong, J.: UAV remote sensing for urban vegetation mapping using random forest and texture analysis. *Remote Sens.* **7**(1), 74–78 (2015)
4. Watanabe, Y., Fabiani, P.: Optimal guidance design for UAV visual target tracking in an urban environment. *IFAC Proc. Vol.* **43**(15), 69–74 (2010)
5. Comaniciu, D., Meer, P.: Mean shift: a robust approach toward feature space analysis. *IEEE Trans Pattern Anal. Mach. Intell.* **24**(5), 603–619 (2002)
6. Neshati, M., Beigy, H.: Expert group formation using facility location analysis. *Inf. Process. Manage.* **50**(10), 361–383 (2014)
7. Wang, X.: Analysis on the development trend of future UAV equipment technology. *Acad. J. Eng. Technol. Sci.* **2**(1), 114–121 (2019)
8. Wei, F.J.: Assessment and response of UAV development. *Appl. Mech. Mater.* **3468**(644–650) (2014)
9. Yamaguchi, S.: State-of-the-art and future direction of UAV technologies. *IEEE Cons. Electr. Mag.* **3**(11), 4–5 (2021)
10. Wang, G.: UAV data link system: a survey. *Sci. J. Intell. Syst. Res.* **3**(11) (2021)



# Prediction Model of Popular Tourist Attractions Based on Big Data Fusion Algorithm

Yan Wang<sup>(✉)</sup>

Dean's Office, Wuhan Business University, Wuhan, China  
706870720@qq.com

**Abstract.** With the continuous development of the tourism industry, it has become a crucial issue to grasp the needs of tourists and accurately select suitable tourist attractions and related tourism services. The prediction model of popular tourist attractions can well solve the problem of decision-making of attractions in tourism activities. For tourists, it can help them choose the scenic spot products with the highest tourism utility, and for travel agencies, it can also improve customer satisfaction and efficiency. In this paper, a prediction system for popular tourist attractions is established, and a big data fusion algorithm is introduced into the system to collect users' browsing data of scenic spots. By comparing the prediction accuracy of the system designed in this paper with the gray prediction model, the prediction system based on the big data fusion algorithm is verified. The prediction accuracy of popular attractions is higher.

**Keywords:** Big Data Fusion Algorithm · Popular Tourist Attractions · Prediction Model · Prediction Accuracy

## 1 Introduction

As the core decision of a tourist activity, the choice of tourist attractions largely determines the quality of the tourist activity. Since tourism requires a relatively large investment of time cost, capital cost, physical cost, etc., the expectation of choosing the best tourist attractions is stronger, and the selection method is particularly important. Reasonable selection of scenic spots can enable tourists to choose tourist attractions that meet their own needs, which is of great significance in all aspects.

So far, many scholars have studied the prediction model of popular tourist attractions based on big data fusion algorithm, and achieved good results. For example, a scholar designed a tourist attraction recommendation system using knowledge-based recommendation. The system simulates the way of a tour guide and provides users with interesting tourist information according to the historical behavior of user operations. However, if you want to obtain accurate recommendation information, It is necessary to update the knowledge base from time to time to adapt to the changing interests and preferences of users, but it is difficult to implement [1, 2]. A scholar believes that online word-of-mouth information has an achievement effect on tourism decision-making. Good



word-of-mouth information can enable decision-makers to make decisions more accurately and quickly. At the same time, it also establishes a good image for scenic spots and promotes the optimal development of scenic spots. Information can change decision makers' perception of scenic spots, eliminate the desire to travel, and reduce the attractiveness and competitiveness of scenic spots at the same time [3]. Although there are many related literatures on the prediction model of popular tourist attractions, the prediction model established at present cannot meet the needs of tourists and travel agencies.

This paper expounds the key steps of the prediction model of key tourist attractions, and then proposes several big data fusion algorithms and gray prediction models. After designing a prediction system for popular tourist attractions based on big data fusion algorithms, the system is used to predict the gold of last November. Weekly word frequency search rate of popular scenic spots, and comparing the prediction accuracy of the system and the gray prediction model, the experiments prove that the system designed in this paper is feasible to predict popular scenic spots.

## **2 Predictive Model Establishment and Big Data Fusion Algorithm**

### **2.1 Construction of a Tourist Attraction Prediction Model**

In some current studies, some scholars select the scenic spots and tour routes with the highest evaluation scores from the perspective of scenic spot evaluation [4]; some scholars have studied the optimization algorithm for maximizing the satisfaction of customers' personalized needs [5] Some scholars use the grey relational analysis method to recommend tourist destinations to tourists based on their past tourist information and data [6]. However, looking at the research results in recent years, scholars have studied the selection of scenic spots from different angles, and there is no systematic method and model. In this regard, in order to facilitate tourists to choose tourist attractions, this paper summarizes several stages of establishing a tourist attraction prediction model:

**Tourist demand expression stage.** When tourists have travel motives, they will make a preliminary judgment on their own needs. Therefore, when tourists first come to travel agencies for consultation, they will first put forward some of their own established constraints, and travel agency staff will recommend several types according to the requirements of customers. Corresponding tourist attractions [7].

**Tourist demand mining stage.** After the tourist motivation is generated, tourists will collect relevant information based on past experience or through various channels to form a general set of demand conditions. However, tourists are often inexperienced or have limited information, and the demand raised in the demand expression stage may not be enough. To be comprehensive, it is necessary to tap the potential needs of tourists at this time. According to the tourists' personal information and past travel data, the tourists' preferences are analyzed, and the tourist attractions that the tourists may like are recommended [8].

**Requirement re-determination stage.** When the set of alternatives filtered according to the needs of tourists and the set of recommended alternatives are generated, the demand data of tourists may be affected by the set of alternatives and make changes. At this time, the needs of tourists need to be detailed finalized [9].

The final selection stage. When all the alternatives and the final needs of tourists are determined, the alternatives can be selected. The selection of tourist attractions can be regarded as a multi-attribute decision-making process, according to the attributes of each scenic spot and its importance. Select the best spots [10].

## 2.2 Big Data Fusion Algorithm

The generation and application of data fusion is based on multiple levels of theoretical and practical knowledge, such as decision theory, network technology [11]. At this stage, there are still many debates on the theory in this area, and the existing systems all have loopholes to varying degrees. As the application scope of data fusion becomes more and more extensive and its importance gradually becomes prominent, many scholars have begun to match some specific fusion theories with real application scenarios and propose feasible algorithms [12].

- 1) Reliability data fusion: According to the reliability of different data sources, assign different weight values, rather than regard each data source as equally important. That is, more weight is given to data sources with high reliability, and less weight is given otherwise. This method is a reliable data fusion algorithm.
- 2) Bayes fusion: Bayes' rule refers to updating the previous likelihood estimate if an observation is added on the basis of determining the likelihood ratio. According to Eq. 2.1, when new observations are added, the posterior probability can be obtained based on the prior probability of a given hypothesis.

$$P(E_i|A_j) = \frac{P(A_j|E_j) \cdot P(E_i)}{P(A_j)} \quad (1)$$

In the above formula,  $E_i(i = 1, 2, \dots, n)$  represents the hypothesized event space,  $A_j(j = 1, 2, \dots, n)$  represents the event space constructed by the observed values,  $P(E_i)$  represents the prior probability, the sum of the probability of obtaining event  $E_i$  in different situations;  $P(A_j)$  represents the normalization constant,  $P(A_j|E_i)$  represents the probability of acquiring the observed value  $A_j$  when  $E_i$  appears;  $P(E_j|A_i)$  represents the probability that the hypothetical event  $E_i$  occurs while acquiring  $A_j$ .

Feature data fusion: Feature-level data fusion uses feature-based information extraction for monitoring each data node (such as the location, status, etc. of the extracted object) according to the specific situation. After the extraction is completed, the vector of the feature will be formed. The vector is fused, and then it is described to a certain extent according to the fusion situation.

## 2.3 Grey Prediction Model

The gray system model uses a certain data column to predict the effect size of the future time, and also has a good prediction effect on the data column with unknown information. When analyzing the structure of the grey prediction system, it is found that many data can only be verified by the experimenter's own logic theory to prove the correctness of the relationship. But in fact, the gray prediction model that can be established is not

the only one, it can only be a model established by the experimenter from a certain perspective or aspect, in order to establish the correctness of the model. The three basic forecasting models based on gray system include GM (1, 1) model, DGM(1, 1) model and Verhulst model.

The gray GM (1, 1) prediction model takes popular tourist attractions as model samples, denoted as  $X^{(0)}$ , and its sequence form is shown in formula (2).

$$X^{(0)} = \left( x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n) \right) \quad (2)$$

$X^{(0)}$  includes  $n$  popular attractions.

### 3 Design of a Prediction System for Popular Tourist Attractions Based on Big Data Fusion Algorithm

When tourists choose tourist attractions, they usually browse the tourist attractions information on the tourism APP first. The information is transmitted to the big data platform, and the recommendation function of the APP can recommend the attractions for them according to the user's browsing traces. The prediction system designed in this paper can collect the information of user browsing.

#### 3.1 Design of Data Acquisition Module

There are two main ways of data collection: API and crawler. The acquisition of network data is generally realized by the method of web crawler, that is, the entry URL is set in the program of the web crawler, and then the crawler program will store the desired target webpage content locally through the designed data fusion algorithm. For other valid URLs in the web page, it will be used as the entry address of the web page data to be crawled next, and the web crawler will not terminate until the crawling ends or the set end condition is met.

#### 3.2 Extract the Design of Key User Modules

In the network search data, as long as you master the basic computer operations, you can experience and apply it. Therefore, any user can publish information and eliminate those data that are meaningless for our research, which can not only reduce the overall data processing volume, It can also improve the correctness of the experiment.

#### 3.3 Design of Tourism Information Dissemination Trend Analysis Module

The design of this module is based on the premise of obtaining high-frequency tourism keywords. If we can analyze the way of information dissemination, then we can get the specific situation of information dissemination. In this paper, according to whether the obtained high-frequency tourism keywords are forwarded by a large number of users, they will become highly popular keywords on the Internet in a certain period of time in the future.

## 4 Experimental Analysis

### 4.1 Application of Popular Attractions Prediction System

Many people will find a good strategy on the travel APP before traveling, then the big data will record the user's search and browsing records. In this experiment, the data from September 21 to September 29 last year was intercepted on a travel APP, and the data during this period was used to predict the popularity of tourist attractions in the next period of time. The experiment predicts the data from September 30 to October 11 of the year. These data can be obtained through the data collection module in the tourism popular scenic spot prediction system. We can directly read the data to process the analysis. We take out the top five keywords from the high-frequency tourism keyword collection, and the five keywords represent the five most searched attractions, that is, popular attractions. Then calculate the percentage of the search term frequencies of these five popular attractions in all searched attractions word frequencies in each time period. The word frequency growth trend of each keyword is obtained as shown in Fig. 1.

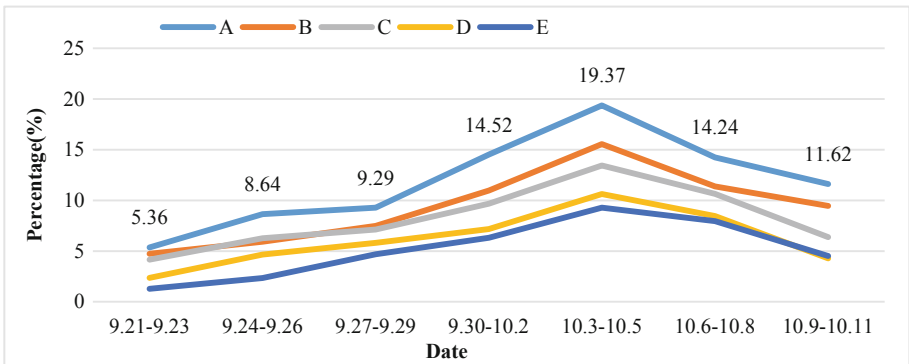


Fig. 1. Percentage of search term frequencies for popular tourist attractions

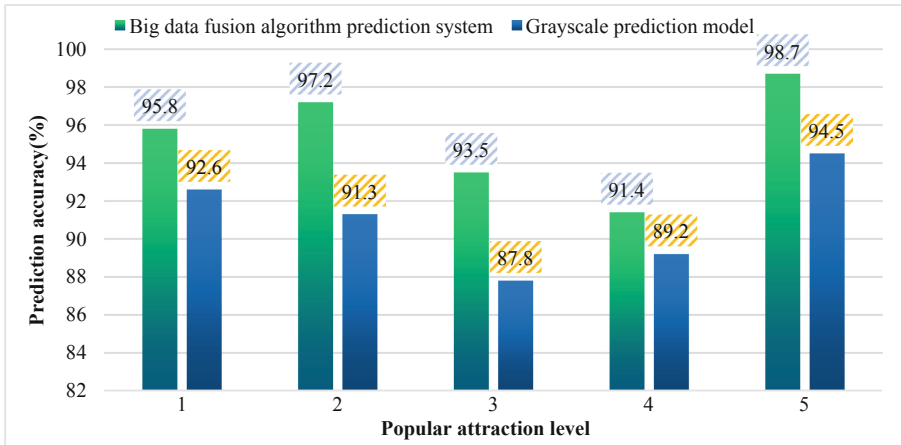
The abscissa of Fig. 1 is the date of every three days from September 21 to October 11 as a time period, and the ordinate is the percentage of the word frequency of high-frequency tourism keywords in all keyword word frequencies. From Fig. 1, we can clearly see that the popularity of each scenic spot keyword during the period from September 30 to October 11. Attraction A is still the most popular tourist attraction, and its popularity remains the highest. Attraction B and Attraction C The popularity of these two attractions is similar, and attractions D and E are the least popular among the 5 attractions. Secondly, the popularity of these tourist attractions peaked from October 3rd to October 5th. In the following period, the popularity of various tourist attractions has declined, which is also in line with the reality that people travel during the Golden Week of November. Behavior is consistent. The results of this prediction are consistent with the actual attractions that people travel during the Golden Week. The prediction results can provide a reference value for everyone to make travel decisions during the Golden Week. Try to avoid the peak travel period and choose the appropriate travel time and attractions, which can not only make oneself play happy but also bring certain convenience to others' travel.

### 4.2 Comparison of Prediction Accuracy

In this experiment, five levels of popular scenic spots were selected. The data samples of each level of scenic spots are shown in Table 1. This sample is also the scenic spot data searched by users on some travel apps. There are 60, 75, 60, 50, and 80 sample data of popular attractions at level 1, level 2, level 3, level 4, and level 5, respectively. Then compare the prediction accuracy of the prediction system based on the big data fusion algorithm and the gray prediction model proposed in this paper, and the results are shown in Fig. 2.

**Table 1.** Data sample of popular tourist attractions

	Number of samples
Level 1 Popular	60
Level 2 Popular	75
Level 3 Popular	60
Level 4 Popular	50
Level 5 Popular	80



**Fig. 2.** Prediction accuracy results

As shown in Fig. 2, the prediction accuracy of the popular scenic spot prediction system based on the big data fusion algorithm proposed in this paper is higher than that of the gray prediction model, and the prediction accuracy of the system increases with the increase of the sample size of people’s scenic spots. The increase in scenic spot information browsed on the APP will allow the system to collect more tourist tendencies, and then predict the tourist attractions trends. This also shows that the prediction system designed in this paper is effective in predicting the tourism trend of popular scenic spots.

## 5 Conclusion

In the era of big data, tourists can obtain a large amount of data through many channels when choosing tourist attractions. Most of these data are meaningless or unhelpful. However, this paper can collect tourists' search records before traveling by designing a prediction system for popular tourist attractions., and recommend attractions according to the needs of tourists. The research on the forecast model of tourist attractions is conducive to subdividing customer needs, matching the needs of tourists to the greatest extent, and helping tourists make travel decisions.

## References

1. Buevich, A., Sergeev, A., Shichkin, A., Baglaeva, E.: A two-step combined algorithm based on NARX neural network and the subsequent prediction of the residues improves prediction accuracy of the greenhouse gases concentrations. *Neural Comput. Appl.* **33**(5), 1547–1557 (2020). <https://doi.org/10.1007/s00521-020-04995-4>
2. Hall, C.M., Ram, Y.: Measuring the relationship between tourism and walkability? walk score and English tourist attractions. *J. Sustain. Tour.* **27**(1–3), 223–240 (2019)
3. Kankhuni, Z.: Exploring the relationship between travel motivations and preferred tourist attractions: A case of New Zealand and The United Kingdom. *Ottoman J. Tour. Manag. Res.* **4**(1), 378–395 (2019)
4. Hernández-Santaolalla, V., Sanz-Marcos, P.: Following death: suicide as tourist attraction through popular culture. *J. Popular Cult.* **52**(6), 1290–1311 (2019)
5. Kaharuddin, S., Al, W., Sibolga, I., et al.: Determinants of tourist attraction of the heritage tourism. *J. Environ. Manag. Tourism*, **XII**(2(50)), 507–514 (2021)
6. Kim, N., Choi, W.H.: The impact of tourism identity of tourist streets, as tourist attractions, on city tourism satisfaction: focus on city travellers at Hwangridan-gil, Gyeongju. *J. Hosp. Tourism Stud.* **22**(3), 15–28 (2020)
7. Kim, Y., Son, Y.-H., et al.: The role of tourist behaviour in the determination of tourist attractions. *Int. Rev. Spat. Plann. Sustain. Dev.* **6**(4), 62–75 (2018)
8. Kim, S.T.: An impact of tourism risk perception on intention to visit tourist attractions : focusing on mediating effect of tourist attitude. *J. Tourism Enhance.* **8**(4), 77–95 (2020)
9. Wahurwagh, R.A., Chouragade, P.M.: Popular place prediction and image recommendation using hierarchical multi-clue modeling for tourist. *Int. J. Comput. Sci. Eng.* **7**(4), 969–972 (2019)
10. Ayoubi, R., Kaboli, S.: A fast vacuum arc detection method based on the neural network data fusion algorithm for the high-voltage dc power supply of vacuum tubes. *IEEE Trans. Plasma Sci.* (99), 1–10 (2020)
11. Abdunabiyev, S.: A fusion detection algorithm targeted to gene expression outliers using Rna-Seq data. *Acta Turin Polytech. Univ. Tashkent* **8**(2), 4 (2018)
12. Zopetti, N., Ceccherini, S., Carli, B., et al.: Application of the complete data fusion algorithm to the ozone profiles measured by geostationary and low-Earth-orbit satellites: a feasibility study. *Atmos. Measure. Tech.* **14**(3), 2041–2053 (2021)



# Application of Genetic Algorithm in Rail Transit Comprehensive Monitoring System

Hui Lin (✉)

Shanghai Interlink Road and Bridge Engineering Co., Ltd, Shanghai 200000, China  
472609100@qq.com

**Abstract.** Over the years, various countries have been committed to studying how to optimize the entire monitoring system, especially for urban rail transit systems. At present, with the continuous development of rail transit, great changes have taken place in its operating environment, working principles, and technical indicators. Rail transit monitoring has played a large role in the development of rail transit. It can effectively solve the problem of failures in the operation of existing equipment and reduce the interference caused by the mutual influence of various subsystems on the line. However, due to the complexity and dynamic characteristics of rail transit, it is inevitable that there will be loopholes in monitoring. At the same time, genetic algorithm searches for the global optimal value by simulating biological evolution and natural selection mechanism, and it also has a good development prospect in the application of global optimization problem. This article adopts experimental analysis method and data analysis method to better understand the data transmission flow of the integrated monitoring system through experimental research, so as to analyze the performance of the system. According to experimental research, the monitoring system used in this experiment has the ability to process a large amount of information at the same time under extreme conditions, and will not cause network paralysis; and while ensuring the stable operation of the system, it also leaves a certain amount of space for future operation transformation.

**Keywords:** Genetic Algorithm · Rail Transit · Integrated Monitoring System · Control Strategy

## 1 Introduction

The rail transit integrated monitoring system is a kind of intelligent, advanced and integrated features, and can realize unified control and management of the entire city's rail lines, thereby effectively improving the overall operating efficiency. At the same time, genetic algorithm is a random search optimization method that simulates biological evolution and natural selection. It has the characteristics of fast global convergence and strong robustness. It is widely used in many fields in actual engineering. In order to ensure the safe operation of the entire rail transit, it must be comprehensively monitored and optimized to realize the efficient, safe and stable operation of the overall rail transit

system. Therefore, this article focuses on the application of genetic algorithm in rail transit integrated monitoring system to improve its overall operating efficiency.

At present, the research results on rail transit and genetic algorithms are relatively rich. For example, Sun Wang pointed out that the particle swarm genetic algorithm has a good effect on optimization problems, which is conducive to shortening the return interval of rail transit trains to the terminal station and improving operation efficiency [1]. Zhang Dan believes that in recent years, urban rail transit has developed rapidly. The rail transit monitoring system has played an important role in ensuring the safe and efficient operation of trains [2]. Wang Ning proposed that rail transit and conventional public transportation are the two most important components of the public transportation system. At the same time, genetic algorithms have a greater impact on the optimization of public transportation routes [3]. Therefore, this article combines genetic algorithms to conduct in-depth research on the rail transit integrated monitoring system, which has important practical significance and research value for improving the operating efficiency of the urban rail transit system and ensuring the safe operation of the urban rail transit system.

This article mainly discusses these aspects. First, the genetic algorithm and its related research are explained. Then, it discusses the rail transit integrated monitoring system and related research. In addition, the application of genetic algorithm in rail transit train control strategy is also introduced. Finally, in order to better understand the performance of the integrated monitoring system, an experimental study was carried out for the system, and the experimental results and analysis conclusions were drawn.

## 2 Related Theoretical Overview and Research

### 2.1 Genetic Algorithm and Related Research

Genetic algorithm is a search optimization algorithm based on natural selection and genetic mechanism. In the case of an optimization problem, it is calculated and solved by imitating the natural evolutionary law of the population. It can also solve the research problems of nonlinear programming and analysis of complex systems.

In nature, gene coding and selection are all operated by imitating animal behaviors to obtain optimal solutions. This optimization method has the characteristics of randomness, strong self-organization ability and good robustness, but it also has some shortcomings. For example, the search space is too large, which can easily lead to premature convergence and insufficient individual diversity [4, 5].

The genetic algorithm first assumes that a group consists of many genetically coded individuals, where each individual is actually a unit with a unique chromosome. Under the premise of survival of the fittest and survival of the fittest, the first generation population was created, which will evolve from generation to generation. In each generation, individuals are selected according to the suitability of individuals in the problem area, and genetic calculations derived from natural genetics are used to combine hybridization and mutation to form a population that represents a new collection. This process will cause the descendants of the population to adapt to the environment more like natural evolution than the previous generation, and find the optimal individual.



Although genetic algorithm has many advantages compared with traditional optimization algorithms, research shows that genetic algorithm also has its shortcomings. One is that the genetic algorithm is too slow to deal with certain problems. Secondly, genetic algorithm is prone to premature phenomenon that makes the algorithm fall into local optimal solution.

As a widely used direct search algorithm, genetic algorithm has applications in many fields, especially in automatic control, planning and construction, combination optimization, disease treatment, image processing, signal processing, artificial life and other fields. At present, genetic algorithms are mainly used for traveling salesman problems, vehicle route optimization problems, route optimization problems, and workshop planning problems [6, 7].

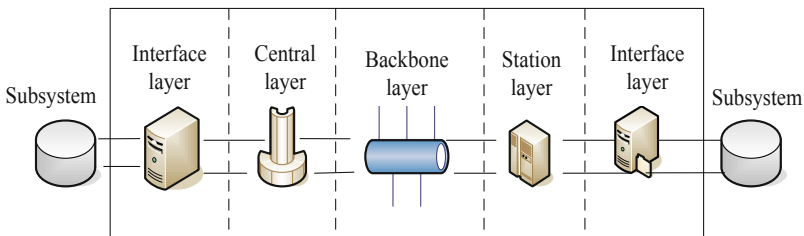
## 2.2 Rail Transit Integrated Monitoring System and Related Research

With the rapid development of cities and rapid population growth, it has also brought about the problem of traffic congestion. The development and maturity of rail transit technology has made the development of rail transit another important choice to solve traffic congestion. Among them, the integrated monitoring system plays an important role in rail transit. The rapid development of rail transit construction puts forward higher requirements for the integrated monitoring system.

In recent years, integrated monitoring systems have been popularized in major cities across the country. This method is now used in many integrated subway monitoring systems. In this way, the automatic train control system (ATC) works autonomously, which is extremely beneficial to the safe operation of the subway.

At present, the rail transit integrated monitoring system widely used in China is an integrated monitoring system with equipment monitoring as the main body. It can be said that the integrated monitoring system is becoming the development trend of the national urban rail transit automation system. Developed countries often adopt the most advanced technology and equipment, and achieve a high degree of integration, which also reflects a country's comprehensive scientific and technological strength, as well as the level of operation and management.

The main function of the rail transit integrated monitoring system is to realize real-time monitoring and automatic regulation of trains during operation through effective control of the entire urban road network, sections and stations. It can also monitor the



**Fig. 1.** Network Structure of Comprehensive Monitoring System

information between different road sections and key nodes [8, 9]. The network structure of the integrated monitoring system is shown in Fig. 1.

The rail transit integrated monitoring system is the nerve center of rail transit. The control center is equivalent to the brain of this central system. It detects the operating status of the entire rail transit system and adjusts its operation based on this information.

The function of the integrated control system is related to the operation mode of the rail transit system. The different operation modes of the rail transit system determine the operation mode of the integrated control system. The integrated monitoring system has these functions.

First, under normal circumstances, the control center of the integrated monitoring system is usually responsible for monitoring the entire line and various related professional systems. Turn on or turn off various devices according to the given work process and working mode, and display information such as power system, lighting system, environmental control system and guidance information according to train operation information, passenger flow information of each station, and environmental monitoring information of each station.

Second, the integrated monitoring system plays an extremely important role in the disaster model. When the detection device detects the occurrence of a disaster, the detection device sends an alarm message to the system, and the system automatically switches to the appropriate disaster operation mode according to the type of information received. Alarm information is displayed on the monitoring interface of the integrated monitoring system. In addition, it also includes information such as video images of the disaster area, equipment status, train operation status and location. In the disaster mode, the integrated monitoring system conducts a series of system coupling tests according to the disposal measures of the disaster management center, which greatly improves the ability of the rail transit system to resist disasters.

Third, when the main equipment system (such as power system, traction system, etc.) fails, the system automatically enters the failure mode and plans according to the nature of the failure. According to the requirements of fault management measures, reset the control mode of the equipment operation authority and revoke the remote control authority of the equipment. At the same time, notify the maintenance personnel to eliminate the fault according to the maintenance plan, and the system will return to normal operation after the fault is eliminated [10, 11].

### **2.3 Application of Genetic Algorithm in Rail Transit Train Control Strategy.**

When using the genetic algorithm, the train model should be combined with the control strategy to control the “operation” of the train model in advance. Write down the various parameters that can be reflected by the control reserve in order to identify the advantages and disadvantages of the control strategy, so that the train should run in a safer and more energy-efficient manner. The genetic algorithm uses the deterministic selection method. After the mutation operation, the best retention selection method is adopted to obtain the most adaptable chromosomes in the child population and the parent population, so as to extract the best chromosomes from the previous generations. The increase in the calculation speed of the algorithm also increases the convergence of the algorithm [12].

The specific calculation method is shown in formula (1) (2).

$$W_u = R \times d_u / \sum_u d_u \quad (1)$$

$$T = R - \sum_u \text{floor}(W_u) \quad (2)$$

Among them,  $W_u$  is the expected survival number of the  $u$ -th individual in the next-generation population,  $R$  is the size of the population, and  $d_u$  is the fitness of the  $u$ -th individual. The integer part of  $R$  is the number of survival of each individual in the next population, and then the individuals are sorted in descending order according to the decimal part, and the first one is added to the next population.  $\text{floor}(W_u)$  function is the floor function.

### 3 Experiment and Research

#### 3.1 Experimental Background

The stability of the integrated monitoring system directly affects the safe operation of rail transit. The integrated monitoring system must not only ensure a reasonable structure in the planning stage, but also pay attention to the safety of the system, so as to ensure the safety of citizens using rail transit. Under normal circumstances, the system can operate normally. Therefore, we must first analyze and test to verify the performance of the system and troubleshoot in time to deal with disasters or extreme situations. Therefore, it is very necessary to understand the performance of the system.

#### 3.2 Experimental Environment

In this experiment, the test tools will use C306 front-end processor, SystematICS, human-machine interface and excel. The C306 front-end processor is the data center of the distributed data acquisition system of the rail transit integrated monitoring system. It distributes data to different geographic locations through its own serial port, Ethernet, fieldbus and other communication media according to a specific communication protocol. The front-end processor C306 is composed of a power supply module, a CPU module, a serial port module, and an MMI module. These modules are connected through the backplane bus of the C306 device.

#### 3.3 Experimental Process

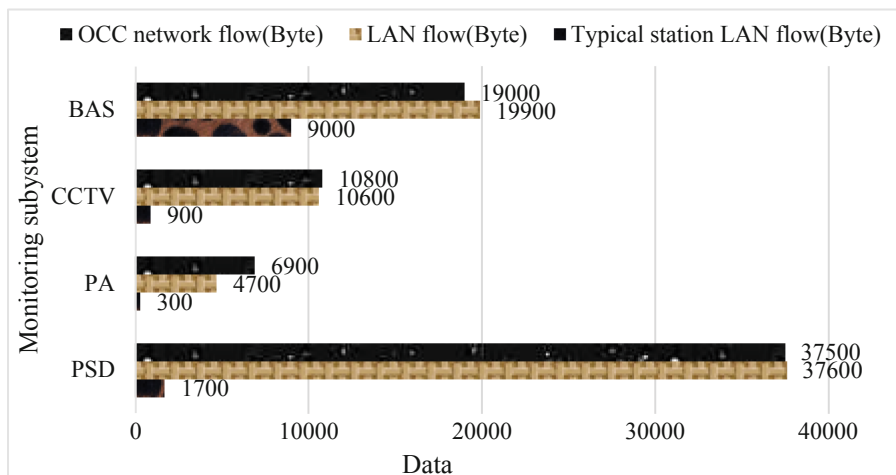
The purpose of this experiment is to better understand the data transmission flow of the integrated monitoring system, so as to analyze the performance of the system. Therefore, this experiment tests the typical station LAN traffic, LAN traffic, and OCC network traffic of the monitoring subsystems PSD, PA, CCTV, and BAS. Some test results are shown below.

## 4 Analysis and Discussion

In this experiment, in order to understand the data transmission flow of the integrated monitoring system, the typical station local area network traffic, local area network traffic, and OCC network traffic of the monitoring subsystems PSD, PA, CCTV, and BAS were tested. The test results are shown in Table 1.

**Table 1.** Data Transmission Flow of Comprehensive Monitoring System

Monitoring subsystem	Typical station LAN flow(Byte)	LAN flow(Byte)	OCC network flow(Byte)
PSD	1700	37600	37500
PA	300	4700	6900
CCTV	900	10600	10800
BAS	9000	19900	19000



**Fig. 2.** Data Transmission Flow of Comprehensive Monitoring System

As shown in Fig. 2, the local area network traffic of the monitoring subsystems PSD, PA, CCTV, and BAS are 37600, 1700, 10600, and 19900 bytes, respectively. It can be seen that the monitoring system for this experiment has the ability to process a large amount of information at the same time under extreme conditions without causing network paralysis; and while ensuring the stable operation of the system, it also leaves a certain amount of space for future operation transformation.

## 5 Conclusion

With the rapid development of rail transit and related technologies, in order to ensure the safe operation of the entire rail transit, it is necessary to monitor the operating

conditions of the equipment in the subway station in real time to achieve dynamic control. At the same time, its integrated monitoring system needs to be optimized to ensure the safe operation of rail transit. Genetic algorithm has been widely used at present, it is by imitating the survival of the fittest mechanism in nature, and continuously optimizing the solution space structure to obtain the optimal parameters. Therefore, this article combines genetic algorithm to conduct in-depth research on rail transit integrated monitoring system, which has important practical significance and research value for promoting the informatization construction of urban rail transit.

## References

1. Rosell, F., Codina, E., Montero, L.: A combined and robust modal-split/traffic assignment model for rail and road freight transport. *Eur. J. Oper. Res.* **303**(2), 688–698 (2022)
2. Musina, A., et al.: The psychophysiological status of rail traffic operators and modern approaches to its correction. *Public Transp.* **14**(3), 635–653 (2021). <https://doi.org/10.1007/s12469-021-00272-2>
3. Besinovic, N., Wang, Y., Zhu, S., Quaglietta, E., Tang, T., Goverde, R.M.: A matheuristic for the integrated disruption management of traffic, passengers and stations in urban railway lines. *IEEE Trans. Intell. Transp. Syst.* **23**(8), 10380–10394 (2022)
4. Vujanic, R., Hill, A.J.: Computationally efficient dynamic traffic optimization of railway systems. *IEEE Trans. Intell. Transp. Syst.* **23**(5), 4706–4719 (2022)
5. Divis, R., Kavicka, A.: Reflective nested simulations supporting optimizations within sequential railway traffic simulators. *ACM Trans. Model. Comput. Simul.* **32**(1), 1:1–1:34 (2022)
6. Marcelli, E., Pellegrini, P.: Literature review toward decentralized railway traffic management. *IEEE Intell. Transp. Syst. Mag.* **13**(3), 234–252 (2021)
7. Bärmann, A., Martin, A., Schneider, O.: Efficient formulations and decomposition approaches for power peak reduction in railway traffic via timetabling. *Transp. Sci.* **55**(3), 747–767 (2021)
8. Zinder, Y., Lazarev, A.A., Musatova, E.G.: Rescheduling traffic on a partially blocked segment of railway with a siding. *Autom. Remote. Control.* **81**(6), 955–966 (2020). <https://doi.org/10.1134/S0005117920060016>
9. Mathew, S., Pulugurtha, S.S.: Assessing the effect of a light rail transit system on road traffic travel time reliability. *Public Transp.* **12**(2), 313–333 (2020). <https://doi.org/10.1007/s12469-020-00234-0>
10. Ghasempour, T., Nicholson, G.L., Kirkwood, D., Fujiyama, T., Heydecker, B.: Distributed approximate dynamic control for traffic management of busy railway networks. *IEEE Trans. Intell. Transp. Syst.* **21**(9), 3788–3798 (2020)
11. Zhilyakova, L.Y., Kuznetsov, N.A., Matiukhin, V.G., Shabunin, A.B., Takmazian, A.K.: Locomotive assignment graph model for freight traffic on linear section of railway. the problem of finding a maximal independent schedule coverage. *Autom. Remote. Control.* **80**(5), 946–963 (2019). <https://doi.org/10.1134/S0005117919050126>
12. Andreasson, R., Jansson, A.A., Lindblom, J.: The coordination between train traffic controllers and train drivers: a distributed cognition perspective on railway. *Cogn. Technol. Work* **21**(3), 417–443 (2018). <https://doi.org/10.1007/s10111-018-0513-z>



# Construction of Inland VHF Communication System Based on Opposite-Reason Model

Chen Kong<sup>1</sup>, Zhongli Yi<sup>1,2</sup>, Xiqi Liu<sup>3</sup>, Shengbo Qin<sup>1,2,3</sup>, and Shanshan Wang<sup>1,2</sup>(✉)

<sup>1</sup> Transport Planning and Research Institute, Ministry of Transport, Beijing, China  
wangss@bjtu.edu.cn

<sup>2</sup> Laboratory of Transport Safety and Emergency Technology, Beijing, China

<sup>3</sup> Guangxi Maritime Safety Administration of the People's Republic of China, Nanning, China

**Abstract.** This paper analyzes the influencing factors of the construction effect of inland river VHF system from the aspects of topographic factors, background noise, navigation facilities and antenna installation. Based on the classical Reason Model, an improved Opposite-Reason Model is proposed to construct the inland river VHF system. Under the condition that all influencing factors are met, the effect of inland river VHF system will reach the best. On this basis, the construction effect of inland river VHF communication system is analyzed from the aspects of safety supervision, emergency disposal, law enforcement and service.

**Keywords:** Inland River · VHF · Topographic Factors · Background Noise · Coverage Prediction

## 1 Introduction

As an important means to ensure the safety of ship navigation, VHF communication system is the most frequently used communication mode for ships entering and leaving the port, production command and dispatching, berthing and ship intersection. It is an important part of safety supervision and emergency disposal system [1–4]. VHF communication equipment is the communication equipment specified by the International Maritime Organization (IMO) and China's ship inspection specifications. Due to its simple equipment and convenient use, it has been widely used in China's coastal and inland rivers. At present, all kinds of motor ships sailing on China's inland river trunk lines are basically equipped with VHF radio communication equipment.

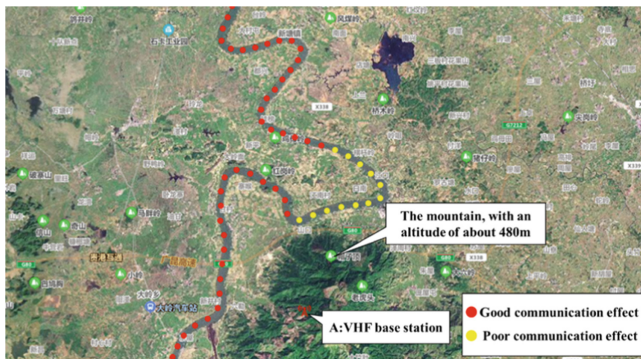
China's inland waters have complex terrain, many mountains and deep valleys, and are subject to the basic construction conditions such as power supply and communication. The construction of VHF system has some problems, such as difficult layout and huge construction cost [5, 6]. The existing VHF base station mainly relies on the existing infrastructure construction of the maritime system. It is generally close to the water surface and at a low altitude. It is vulnerable to the shelter of mountains on both sides, river crossing bridges, ship locks and other tall facilities, and the coverage is small in actual use; Moreover, most of the existing base stations are located in urban and rural population gathering areas, and the electromagnetic environment is becoming worse

and worse, which also affects the communication effect of the existing VHF system to a certain extent. Combined with relevant engineering experience, this paper puts forward the construction model of inland VHF communication system.

## 2 Analysis of Influencing Factors

### 2.1 Topographic Factors

According to the division of VHF frequency band by the International Telecommunication Union (ITU), the working frequency band of water mobile radio communication is 156–174 MHz and the wavelength is 1.72–1.92 m. According to the principles of physics, when the size of the obstacle is about equal to or smaller than the wavelength, VHF radio wave has a certain diffraction ability. However, if the size of the obstacle is much larger than the VHF wavelength, it will block the VHF radio wave and form a coverage blind area. Therefore, in the actual project, the location with higher terrain shall be selected as far as possible to ensure the coverage in combination with the VHF communication coverage requirements when the construction conditions allow.



**Fig. 1.** VHF base station A coverage effect

Figure 1 shows the coverage effect test results of an inland river VHF base station in China. The red dot in the figure indicates that the communication effect is good, the yellow dot indicates poor communication quality, i.e. covering the blind area. The altitude of VHF base station a is about 450 m and the antenna height is about 18 m. Theoretically, the water area is within the effective coverage of base station A. However, there is a high mountain with an altitude of 480 m on its north side, which blocks the VHF base station, resulting in a coverage blind area of about 7.5 km, which is about 6 km away from the VHF base station.

Figure 2 shows the typical geographical environment of China's inland rivers, with high mountains on both sides and deep valleys on the river. This terrain greatly restricts the coverage effect of VHF base stations. Therefore, topographic factors should be considered in the layout and construction of inland river VHF communication system.



Fig. 2. Typical geographical environment of inland river

## 2.2 Background Noise

With the development of social economy and the continuous expansion of urban scale, the electromagnetic environment around VHF base station is deteriorating, and the communication effect is greatly affected. The main factors affecting the electromagnetic environment around the VHF base station include transmission lines, substations, automotive electrical equipment, etc. At the same time, with the increase of construction projects, welding equipment, electrical cranes, generators and other equipment and facilities are frequently used in the construction process, resulting in the obvious deterioration of the regional electromagnetic environment in VHF frequency band and the increase of background noise power. The increase of background noise power is more obvious in specific periods (such as dense construction in white days).

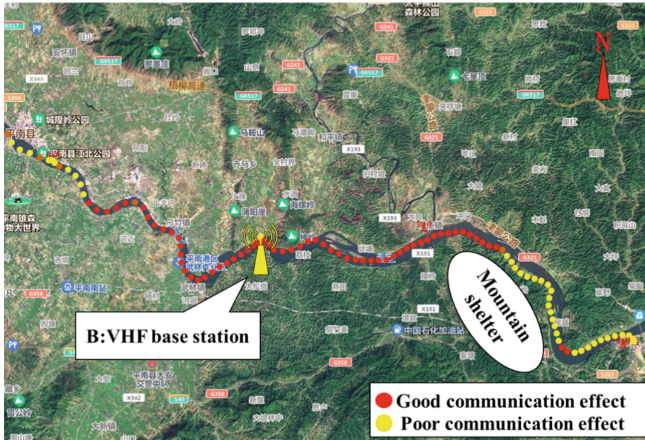
In practical engineering, the comparison test results of background noise between urban base stations and mountainous base stations show that the background noise in urban areas is significantly higher than that in mountainous areas. Figure 3 shows the real ship test results of VHF base station B coverage effect. The altitude of base station B is about 90 m and the antenna height is about 80 m. The coverage blind area on the east side is caused by high mountains. Theoretically, if there is no shelter in the west of the base station, it should be well covered within 25 km, but there is a coverage blind area about 10 km away from the base station, which is caused by the strong background noise around the city, resulting in the reduction of communication quality and reliability.

VHF band has a long communication distance, and a large number of sailing ships carry out daily and emergency communication through VHF. The ship VHF radio transmission power is small. When the ship navigation area is far from the base station, the ship communication signal power received by the base station is small. With the increase of noise power around the base station, the received signal-to-noise ratio will decrease, which will affect the actual communication effect. Therefore, in the process of inland VHF base station layout and system construction, we should pay attention to the impact of background noise.

## 2.3 Navigation Facilities

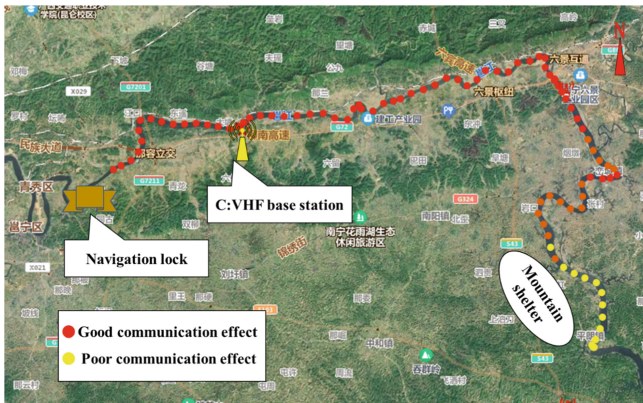
As a navigation facility, ship lock is common in inland river areas. Taking Guangxi, China as an example, the navigation mileage of Xijiang trunk line is 570 km, with





**Fig. 3.** VHF base station B coverage effect

27 ship locks of Laokou ship lock, Xijin ship lock, Guiping ship lock, Datong gorge hub and Changzhou water control hub, which has improved the channel conditions of the reservoir area to a certain extent, promoted the large-scale, standardization and specialization of ships, and improved the safety of ship transportation.

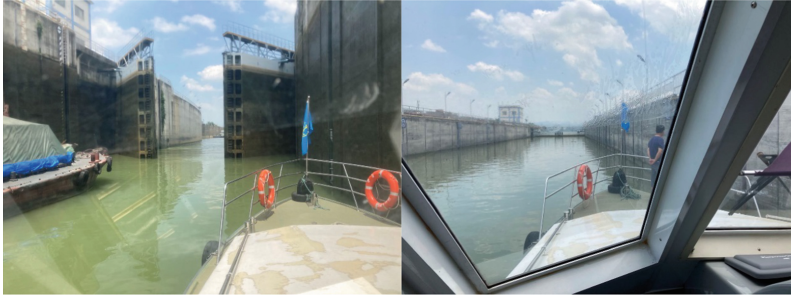


**Fig. 4.** VHF base station C coverage effect

Figure 4 shows the actual ship test results of the coverage effect of inland VHF base station C. the altitude of base station C is about 106 m, and the antenna erection height is about 30 m. An inland river ship lock is located in its west side, about 13 km away from base station C, which directly cuts off the communication of VHF base station C in the west side [7].

The ship lock is often higher than the water surface, and the navigable ship is usually lower. As shown in Fig. 5, the signal of the ship passing through the ship lock and near the ship lock is often lost due to the shielding of the ship lock facilities. Therefore, in the

construction of inland river VHF communication system, it is necessary to fully consider the shielding of navigation buildings such as ship lock to the signal, reasonably select the base station and avoid the impact as far as possible.



**Fig. 5.** Inland river trunk ship lock

## 2.4 Antenna Installation

VHF communication system generally uses vertically polarized whip antenna, and the antenna installation shall comply with the following requirements:

- When the VHF base station needs to set up more than two antennas, in order to increase the isolation of the two antennas, the isolation must be considered to reduce the interference between them. The minimum distance required between antennas is calculated according to the following formula:

$$I = 39.557 \log H + 22.263 \quad (1)$$

where  $I$  represents antenna isolation (dB),  $H$  represents the minimum distance required between the bottom of the receiving antenna and the top of the transmitting antenna (m).

The antenna isolation  $I$  is calculated by the following formula:

$$I = 137 + 10 \log PT + 20 \log S - In \quad (2)$$

where  $PT$  represents the radiated power (W) of jamming transmitter, taken as 50 W;  $S$  represents the sensitivity of the interfered receiver ( $\mu V$ ), taken as 0.35  $\mu V$ ;  $In$  represents a certain anti-interference index (dB) of the interfered receiver, taken as 95 dB.

The antenna shall be installed on the riverside to avoid shielding the antenna by the tower itself.

On the premise of ensuring the antenna erection height, the feeder length of each station shall not be greater than 50 m to reduce the signal attenuation caused by feeder loss.

A certain transverse distance shall be ensured between the antenna and the support tower, and the antenna support cross arm shall not be less than 500 mm. In the actual

ship test of inland river VHF communication system, it has occurred that the transverse distance between the test antenna and the tower is too small, resulting in communication interruption or poor communication quality.

## 2.5 Other Factors

In addition to the above three factors such as terrain, background noise and navigation facilities, the construction of inland river VHF communication base station also needs to consider transmission, support tower, land and other factors:

- The base station should be located at a location convenient for transmission or renting (or self built) transmission lines.
- Try to rely on its own resources or public resources for antenna erection and equipment placement. In principle, no new land or iron tower will be built. Focus on other communication stations with complete supporting facilities and convenient transportation.
- Connect with the existing projects, and put forward the coverage of VHF communication system in combination with the construction planning of inland water transportation on shipping trunk lines.

## 3 Construction of Inland River VHF System Based on Opposite-Reason Model

### 3.1 Classical Reason Model

In 1990, James Reason constructed the Reason Model, which believed that isolated factors could not lead to accidents, which were caused by the simultaneous destruction of multiple defense systems, as shown in Fig. 6. The Reason Model mainly includes light and defense system. The light is the risk factor in production and life, and the defense system is represented by cheese. These cheese layers overlap to make up for their respective defects and vulnerabilities. Under normal conditions, the location and size of vulnerabilities are constantly changing. When the holes in each cheese are aligned, an accident opportunity hole will be generated, resulting in the failure of the whole system. In this case, the light will cross the defense system through the holes, resulting in the occurrence of accidents. This process is essentially the role of continuous accumulation of negligence, so it is also known as cumulative behavior effect.

### 3.2 Classical Reason Model

Based on the classical Reason Model, this paper puts forward the Opposite-Reason Model, which believes that a single good factor will not directly lead to the success of the system, and the excellent construction effect of the system is realized by the satisfaction of multiple good influencing factors at the same time. As shown in Fig. 7, the Opposite-Reason Model mainly includes illumination and advanced system. Illumination is the influencing factor in a system. The advanced system is represented by cheese. These

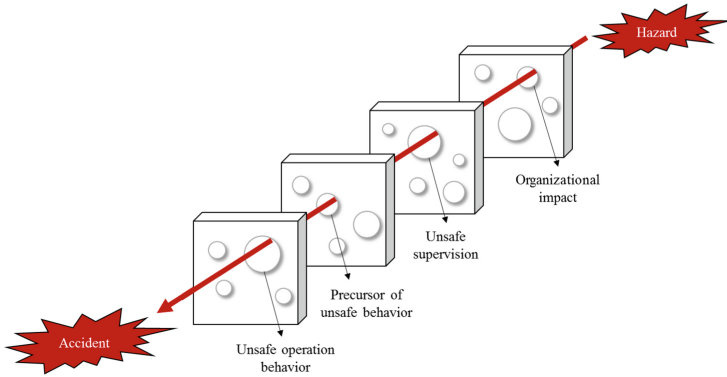


Fig. 6. The Reason Model

cheese layers overlap, and the location and size of vulnerabilities are constantly changing. When the holes in each cheese are aligned, an opportunity hole for system construction will be generated. In this case, the illumination will pass through the advanced system through the hole, and finally realize the construction of the system.

**3.3 Inland River VHF system based on Opposite-Reason Model**

Combined with the analysis of influencing factors in Chapter 2, this paper constructs the construction model of inland river VHF communication system based on the Opposite-Reason Model, as shown in Fig. 7. A successful inland river VHF communication system must meet the conditions of topographic factors, background noise, navigation facilities and antenna installation at the same time, so as to achieve good regional ship shore communication coverage effect.

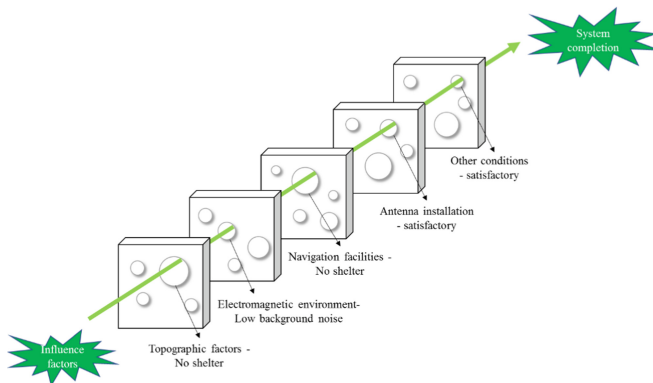


Fig. 7. The Opposite-Reason Model

The construction of inland river VHF communication system needs to follow the Opposite-Reason Model proposed in this paper, consider a variety of influencing factors,

and carry out system construction when each factor meets the conditions, so as to have a good communication effect.

## **4 Analysis of Construction Effectiveness**

### **4.1 Meet the Safety Supervision on Inland Water**

With the continuous promotion of the development strategy of inland water transportation, inland water transportation has ushered in new development opportunities. The main and tributary channels of inland rivers have been continuously expanded and reconstructed, and the trend of large-scale ships is significant [8]. The contradiction between the increasing pressure of inland water safety supervision and the insufficient capacity of the existing VHF communication supervision means of the maritime system is prominent. In order to meet the development needs of inland river safety supervision and implement the inland river water transportation development strategy and relevant planning, it is necessary to vigorously build the inland river VHF communication system.

### **4.2 Improve Emergency Response Capacity**

The frequent occurrence of inland river accidents is closely related to its complex navigation environment. The main inland waterway is mostly curved and narrow, and the local ship flow and ship density are too large, which has a certain adverse impact on the safe navigation of ships. VHF communication includes distress alarm, search and rescue communication and on-site communication. It is an important link of maritime distress search and rescue, and runs through the whole distress rescue process [9]. With the accelerated pace of inland shipping construction, the traffic flow of inland ships is increasing, and the pressure of water safety supervision continues to increase. Inland waters are facing new requirements to enhance maritime service capacity, emergency rescue capacity and communication support capacity.

### **4.3 Adapt to Maritime Law Enforcement Mode**

The mode of maritime supervision has changed from static supervision to dynamic supervision, from pre supervision to in-process and post-supervision. The maritime department is responsible for the emergency disposal of water emergencies, water search and rescue, navigation order management, cruise law enforcement and traffic control [10]. The development of relevant work requires the reliable support of VHF communication system, which can achieve timely feedback and rapid disposal, and run through the whole process of post rescue work. Therefore, strengthening the construction of VHF communication system is an effective measure to adapt to the reform of maritime law enforcement mode.

#### 4.4 Improve Information Services

Service is one of the essential attributes of inland shipping. Inland shipping faces all production departments of the economy along the river. The service process runs through all aspects of social production and circulation, and is closely related to the life of the people along the river. A healthy shipping environment requires rich and accurate information services. Water safety information can be broadcast through VHF communication system, and ship users can obtain the necessary information to the greatest extent in multiple periods, so as to predict the traffic situation in relevant waters, so as to be vigilant in complex navigation environment and consciously abide by navigation rules.

### 5 Conclusion

As an ideal mode of transportation, inland shipping has the characteristics of strong stability, environment-friendly and energy saving. With the increase of inland river navigation mileage and ship traffic volume year by year, the pressure of safety supervision of inland river navigation waters is increasing. Combined with the experience of inland river VHF communication system construction project, this paper constructs the inland river VHF communication system based on the Opposite-Reason Model from the perspectives of topographic factors, background noise, navigation facilities and antenna installation, so as to provide reference for construction or decision-making departments.

### References

1. Chuanbao, D.U., Mao, C., Sun, D., et al.: Effect characterization and modeling of rf low noise amplifier for VHF communication system after HEMP PCI Test. 2020 IEEE MTT-S. In: International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO), pp. 1–4. IEEE (2020)
2. Bashkuev, Y.B., Angarkhaeva, L.K., Buyanova, D.G.: The surface impedance of the thin ice - sea and thick ice - sea structures in Vlf-Vhf radio wave bands. *J. Commun. Technol. Electron* **66**, 1148–1154 (2021)
3. Wei, M., Hu, C., Estévez, D., et al.: Design and flight results of the VHF/UHF communication system of Longjiang lunar microsattellites. *Nat. Commun.* **11**(1), 3425 (2020)
4. Kim, J.U., Park, J., Lee, C., et al.: Target performance analysis of tactical voice communication on VHF narrow-band in combat network radio system. *J. Korea Inst. Mil. Sci. Tech.* **24**(1), 107–114 (2021)
5. Shin, J.: Distributed antenna systems for VHF maritime communications. *Wireless Pers. Commun.* **114**(2), 1623–1633 (2020). <https://doi.org/10.1007/s11277-020-07442-8>
6. Syah, R.A., Dutono, T., Santoso, T.B., et al.: HF and VHF/UHF transverter system for disaster area communication. In: 2020 International Electronics Symposium (IES), pp. 163–168 (2020)
7. Yi, Z., Zhan, J., Kong, C., Wang, S., et al.: Prediction method and verification of inland waters VHF coverage effect based on ITU-R P.1546–6 modified model. In: The 11th International Conference on Communications, Circuits and Systems (2022)
8. Rong, S., Li, H., Xiang, Q., et al.: Synchronization for VHF data exchange system. *J. Phys. Conf. Series* **1952**(4), 042035 (8pp) (2021)
9. Yan, Q., Li, W., Li, J., et al.: Real-time air-to-ground data communication technology of aeroengine health management system with adaptive rate in the whole airspace. *Math. Prob. Eng.* **2021**(7), 1–13 (2021)
10. Yi, Z., Wang, F., Wang, S., et al.: Design of high frequency digital transceiver in coastal radio station and shipborne. *J. Phys. Conf. Ser.* **1920**(1), 012059 (5pp) (2021)



# Intelligent On-Line Monitoring System of General Electrical Equipment Based on Remote Communication

Xinhe Cui<sup>✉</sup>, Fei Xia, Fanbo Meng, Xiaobo Huang, and Xin An

Information Communication Branch of State Grid Liaoning Electric Power Co., Ltd, Shenyang, Liaoning, China  
cuixinhe1993@163.com

**Abstract.** With the in-depth development of smart grid and the continuous improvement of the demand for power supply reliability, the scale of power communication network is expanding rapidly, the power communication equipment is diversified, the proportion of traditional power communication services such as voice, stability and protection is declining, and the remote control business, information management business and intelligent communication business of transmission and distribution network are increasing. The development of power communication will also usher in a qualitative leap. Based on the analysis of the intelligent on-line monitoring system of telecommunication general electrical equipment, this paper discusses the functional requirements of the intelligent on-line monitoring system, the hardware structure of the monitoring terminal, and the design and implementation of the database server. Through the remote on-line condition monitoring algorithm of the communication tower, the TCP throughput and delay data of openflow switch and goose switch are tested, The results show that the measured TCP throughput of openflow switch is about 500 mbit/s, and the maximum is no more than 570 mbit/s. The measured TCP throughput is in line with the reality. The TCP throughput of goose Ethernet switch in the same network environment is obviously not as large as that of openflow switch. Therefore, the throughput performance of openflow switch is good; For the high-speed switch, the delay of 1518 bytes and below is required to be less than 1ms in the least ideal case. Openflow switch has smaller delay, so it is feasible to apply it to the on-line monitoring system of substation equipment.

**Keywords:** Telecommunication · General Electric Equipment · On-Line Monitoring System · Intelligent System

## 1 Introduction

With the expansion of modern industrial scale and the improvement of complexity, the requirements for the reliability and safety of key equipment are also increasing. Therefore, the remote online status monitoring system has been paid more and more attention. At present, the output signal and communication protocol of the sensor do not form a unified standard. The traditional remote online condition monitoring terminal

needs to be customized in terms of hardware and software according to the selected sensor, which leads to its weak expansibility and universality. Therefore, this paper studies the intelligent on-line monitoring system of general electrical equipment based on remote communication.

Based on the analysis of the intelligent on-line monitoring system of general electrical equipment in telecommunication, many scholars at home and abroad have studied it. Kravets o J discusses a method of monitoring and Optimization in telecom networks with packet routing. The mathematical model of data network for monitoring traffic structure is described, and the optimization task is determined. It mainly focuses on the model development based on Queuing Theory in order to generate standby capacity in the system. Aiming at the optimal parameter selection of the monitoring system, the model puts forward the solution technology of similar tasks. The obtained optimization task is complex enough, and subsequent attempts should aim to find conditions that can introduce some constraints, so as to simplify the solution without obvious loss of the adequacy of the model, which is particularly important for asynchronous monitoring tasks [1]. Novilla a uses fuzzy logic method to design a monitoring system based on Internet of things. It includes monitoring the normal activities of manufacturing machines to establish a reference for their status; Then, the data collected from the sensor are monitored and analyzed in real time. The network based on fuzzy logic is used for state detection. The system adopts host, network, Ethernet module, embedded system gateway, sensor, microcontroller (MCU) and other components, which are integrated through the Internet to realize the monitoring system, and uses cloud computing technology and Internet of things devices to safely store, monitor and analyze the data collected by these devices [2].

The remote communication general-purpose electrical equipment intelligent on-line monitoring system designed in this paper solves the problems of universality of data acquisition related to electrical equipment and long-distance wireless real-time data transmission. The database server is used for data storage, which effectively solves the problems of reliability and access convenience of data storage, and fully verifies the reliability of the remote monitoring system in long-time experiments and practical applications. Through the analysis of the remote monitoring system, the overall scheme of the monitoring system is determined. A set of universal monitoring data acquisition module is designed to effectively solve the problems of slow data transmission speed and data loss [3].

## **2 Intelligent On-Line Monitoring System for Telecommunication General Electrical Equipment**

### **2.1 Functional Requirements of Intelligent Online Monitoring System**

At present, the management of the control layer supports two ways. One is the monitoring terminal equivalent to the industrial master computer, through which users can communicate with the server to complete the monitoring, query and control of the equipment; The other is to log in to the server anytime and anywhere through the user's mobile phone, query and display the operation status of the authorized power station



unit, obtain the operation status information parameters, and also start and shut down the equipment [4, 5]. All management operations of hydropower station monitoring system are concentrated on the monitoring terminal. Intelligent terminal based monitoring functions:

- 1) Data acquisition function: The monitoring terminal can query the operation status and technical parameters of the generator set through the server, including vibration signal, partial discharge signal, air gap parameters, temperature, excitation device, accident fault information, pressure, speed, etc. At the same time, it can also measure and record the AC and DC signals input by other sensors (current transformer, voltage transformer, etc.).
- 2) Data processing function: Through the statistics and calculation of the state information transmitted by the unit equipment, draw various curves, make three-dimensional or low-pressure hydro generator unit potential spectrum, analyze and predict the fault according to the trend chart and historical data, and send the prediction results to the dispatching center to complete the functions of fault prediction, analysis and diagnosis.
- 3) Display function: The display function of the monitoring terminal is divided into two types: real-time display and query display. Real time display refers to several status information always displayed on the main display interface, including recording time, system status, water level, power factor, active power, fault status and other information [6]. Query display is to manually query various data and system operation status parameters, including electromechanical voltage and current, grid voltage and current, power generation, excitation information, power supply voltage, temperature and other information.
- 4) Control function: According to the system requirements or instructions issued by the superior, the control cabinet of the generator set in the power station will be started, shut down, on-off, emergency shutdown, fault removal, prohibition of start-up, permission of start-up and other actions. For generator closing and power factor regulation, the monitoring system can complete the control function through the designed automatic quasi synchronization device and power factor regulation device [7].
- 5) Printing function: The computer of the monitoring terminal is connected to the printer and has the functions of timing printing, accident printing and command printing. Timing printing is to set the time in advance and automatically print the operation data within the specified time. Accident printing is that the system prints out fault data when the generator unit fails. Command printing is to print historical data or statistical data tables and data charts according to random commands.

## 2.2 Hardware Structure of Monitoring Terminal

ARM processor module: the processor module is the data operation and control core of the system. It is composed of ARM processor, memory and data storage chip. The memory adopts SDRAM type smart chip to provide high-speed data access space during system operation. The data storage chip adopts NAND flash to provide the system with large capacity and data storage space without loss in power failure.

**Sensor interface module:** the system realizes communication with various sensors through the sensor interface module. In addition to the voltage signal, other types of sensors can be directly connected to the monitoring terminal, and the voltage signal has no unified standard. It needs to be transformed into a 5 V power down signal through an external conditioning circuit to be connected to the monitoring terminal.

**Power module:** in the embedded system, because different state chips have different requirements for voltage value and stability, the power module needs to convert the voltage. The power supply module is composed of switching power supply module and DC/DC conversion chip, which converts the external battery voltage into various stable DC voltages to realize the normal power supply of the system.

**Configuration parsing module:** the configuration parsing module parses the configuration file provided by the user [8, 9]. The file contains a series of table items. W is used to describe the communication protocol of the sensor, and customize the sampling method and processing algorithm of the sensor. The module contains JSON parser, lexical analyzer and parser. The conversion from text file to internal data structure of the program is realized through compilation technology.

**Data acquisition module:** the data acquisition module realizes a universal data acquisition framework, and realizes the complete data acquisition process combined with the communication protocol and sampling control parameters provided by the configuration file. The event driven model is used in the module to realize the whole process. The model is implemented based on Linux multiplexing mechanism to realize the timely response to data acquisition and improve the efficiency of the program.

**Data processing module:** the data processing module realizes the data processing algorithms commonly used in condition monitoring, such as Fourier transform for time-frequency domain transformation, extreme value processing for monitoring extreme conditions, median processing for eliminating accidental errors, mean processing for reducing state deviation, etc. In addition, the data processing module also provides port limit alarm monitoring. Once the state quantity deviates from the normal threshold, it is necessary to send alarm data to the monitoring [10]. The selection of data processing algorithm and the threshold of port limit can be set in the configuration file.

**Data storage module:** the data storage module provides local persistent data storage services for other modules in the software. Since the operating conditions of monitoring terminals are often in industrial sites or remote areas, it is difficult to ensure a stable operating environment. In case of power failure, restart and other situations, the unsent data, especially the alarm data, need to be written into the memory for persistent storage, and can not be deleted until it is successfully uploaded to the server.

**Network communication module:** the monitoring terminal and the monitoring server use HTTP protocol for communication. In the network communication module encapsulates the implementation of HTTP protocol, realizes the network communication service, and provides a unified call interface for other modules. Due to the complex and uncontrollable external network conditions, the network communication module adopts a certain retransmission strategy to ensure the successful upload of data.

**Software and configuration update module:** in some application scenarios, the deployment cost of the monitoring terminal is relatively high and the deployment volume is relatively large. Therefore, when the terminal software needs to be updated, it needs to be updated automatically through the network to reduce the operation and maintenance cost of the

monitoring terminal. The software update module regularly monitors whether there are software updates or configuration updates. During the update, it ensures that the current running task of the program ends gracefully and the state recovery after heavy shoulder [11].

### 2.3 Design and Implementation of Database Server

The server database has a strong portability feature, which can eliminate the need to redistribute the processing data storage space. The main functions and advantages of the database server are as follows: the routine data maintenance of the database mainly includes the storage of real-time data of the tidal current power generation device and the use of third-party software to export the data of the tidal current power generation device for data analysis. If the database is configured correctly, it can be backed up or restored in time. The database has strong parallel operation ability, and different users may access the database server at the same time. Therefore, the database server is required to have strong parallel operation ability and be able to handle different events at the same time. Reduce the programming amount of client and server and shorten the development cycle. The database provides a set of software interfaces used in data transmission and basic operation: API. Set up a special database administrator. Only with the authorization of the administrator can you call the real-time monitoring data in the database, which effectively improves the security of the monitoring data. The system performance is improved and easy to expand. The database server can reduce the network overhead, coordinate the work of all parts, and avoid competition and waste of resources [12]. Support horizontal expansion and some vertical expansion of multiple servers with the same type of processor.

## 3 Remote On-Line Condition Monitoring Algorithm for Communication Tower

In the remote online condition monitoring project of communication tower, it is necessary to conduct real-time condition monitoring on the static inclination, maximum swing angle and swing frequency of communication tower. Therefore, a dual axis inclination sensor is selected in the project. According to the two output angles  $A_1$  and  $A_2$  of dual axis inclination, the spatial inclination of communication tower can be calculated, and its positive and negative are consistent with  $A_1$ . The calculation is as shown in formula (1) (2):

$$\gamma = \arcsin \sqrt{\sin^2 \gamma_1 + \sin^2 \gamma_2} \quad (1)$$

$$\sin^2 \gamma + \cos^2 \gamma = 1 \quad (2)$$

In each monitoring cycle (set as 600 s), 1024 spatial dip angles are collected with a sampling frequency of 10 Hz, and the mean value of these 1024 sampling points is calculated respectively. The static inclination, maximum swing angle and swing frequency

of the communication tower can be obtained by maximum calculation and Fourier transform. At the same time, the normal threshold ranges of the three values are set as Yang, [0, 1], [0, 5], [0, 2], and the units are degrees, degrees and Hertz respectively. Once the normal values are exceeded, an alarm message will also be sent to the monitoring.

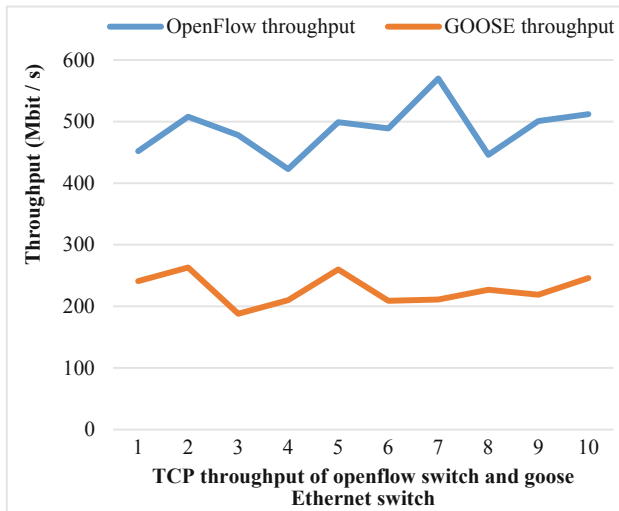
## 4 Experimental Test and Analysis

### 4.1 Throughput Measurement

Measurement of TCP throughput of openflow switch: the measurement is repeated for 10 times, each lasting for 60s. The measurement results are shown in Table 1 and Fig. 1:

**Table 1.** TCP throughput of openflow switch and goose Ethernet switch

	1	2	3	4	5	6	7	8	9	10
Open Flow throughput	452	508	478	423	499	489	570	446	501	512
GOOSE throughput	241	263	188	210	260	209	211	227	219	246



**Fig. 1.** TCP throughput of openflow switch and goose Ethernet switch

It can be seen from the figure that the measured TCP throughput of openflow switch is about 500 mbit/s, and the maximum is no more than 570 mbit / s. Because this rate is much smaller than the maximum load of other devices in the measurement environment, the measured TCP throughput is in line with the reality. The TCP throughput of goose Ethernet switch in the same network environment is obviously not as large as that of openflow switch, so the throughput performance of openflow switch is better.

## 4.2 Time Delay Measurement

From the current experimental results, the performance of openflow is mainly measured in forwarding, and the throughput is taken as the main parameter. But in general, there will be packet delay. The standard special message marking method is used to measure the delay. Ensure that the test lasts long enough, and the synchronization with physical time must be repeated several times. In the switch performance test, the delay under different message lengths should be tested. Each sending process lasts 30 s, and the message lengths are 64, 128, 256, 512, 1024, 1280 and 1518 bytes respectively. The corresponding delay measurement results are shown in Fig. 2:

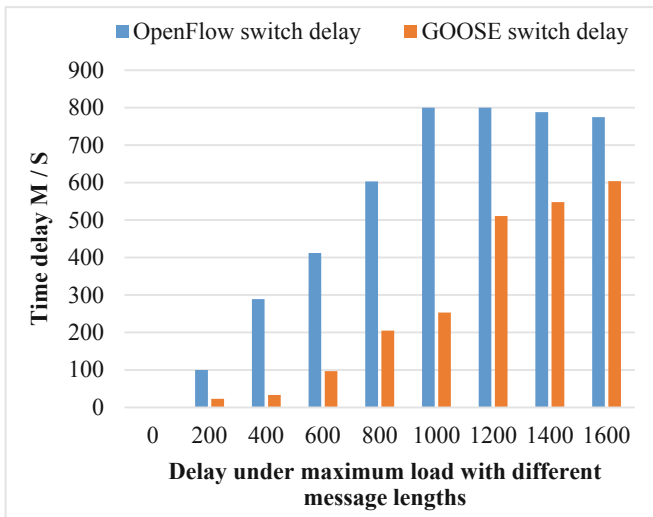


Fig. 2. Delay under maximum load with different message lengths

It can be seen from the above figure that the delay varies with the message length. For high-speed switches, the delay of 1518 bytes and below is required to be less than 1ms in the most unsatisfactory case. It can be seen from the figure that the results meet the benchmark requirements, and the openflow switch has smaller delay. Although the delay of industrial Ethernet switch in intelligent substation is also smaller than the benchmark, it can be seen from the figure that the openflow switch has smaller delay. If the switch does not have the function of network management, it is easy to lead to network storm. The special working mode of openflow network can just prevent the occurrence of network storm, so it is feasible to apply it to the on-line monitoring system of substation equipment.

## 5 Conclusions

As an auxiliary system of power generation device, remote monitoring system is of great significance to ensure the reliable operation of electrical equipment. The monitoring

terminal developed in this paper has been tested by actual projects, and can well meet the function, standardization and universality of remote online condition monitoring in all aspects, but it still needs to be improved: configuration file generation tool. At present, the configuration file is generated by manual editing, which will inevitably lead to wrong input. If the application program with graphical interface is used, the configuration file is automatically generated by Wizard, which can simplify the operation, improve the correctness and writing efficiency of the configuration file, and control the output; At present, the monitoring terminal can realize one-way equipment status information acquisition. If it can output signals to the standby at the same time to form closed-loop control, or realize remote control based on the network, it will make the terminal have greater application value.

**Acknowledgements.** This work was supported by State Grid Liaoning Electric Power Co., Ltd. Science and technology project funding (Project No.: 2020YF-73).

## References

1. Kravets, O.J., Danilina, E.I., Malikova, I.I., et al.: Monitoring optimization of telecommunication networks for generation of reserve capacity in system. *Nexo Revista Científica* **34**(2), 993–1007 (2021)
2. Novilla, A., Balute, A., Gonzales, D.B.: The use of fuzzy logic for online monitoring of manufacturing machine: an intelligent system. *Circ. Comput. Sci.* **2**(11), 31–39 (2017)
3. Faisal, N., Zaveri, A.A., Sami, N., et al.: Intelligent greenhouse monitoring system. *Int. J. Sci. Eng. Res.* **9**(3), 1234–1239 (2018)
4. Ali, S., Ahmad, A., Faheem, Y., Altaf, M., Ullah, H.: Energy-efficient RRH-association and resource allocation in D2D enabled multi-tier 5G C-RAN. *Telecommun. Syst.* **74**(2), 129–143 (2019). <https://doi.org/10.1007/s11235-019-00643-y>
5. Kovshova, I.O., Dubovyk, N.A., Zubko, T.L.: Administration of telecommunication company staff potential in Ukraine. *Int. J. Manag. Hum.* **4**(10), 135–140 (2020)
6. Hadialnashia, I.A., Mohamed, I.I., Almelian, M., et al.: The impact of partial discharge on the dielectric properties of refined bleached and deodorized palm oil (RBDPO). *J. Telecommun.* **10**(No. 1–3), 81–85 (2018)
7. Tahir, M.J., Latiff, I.A., Mazliham, M.S., et al.: Symmetrical and asymmetrical fault currents: evaluation to enhance the performance of 220 KV grid station. *J. Telecommun.* **10**(1–3), 147–152 (2018)
8. Liu, J., Li, W.: Intelligent video monitoring system based on wireless communication. *Revista de la Facultad de Ingenieria* **32**(4), 188–195 (2017)
9. Mwakatumbula, H.J., Moshi, G.C., Mitomo, H.: Consumer protection in the telecommunication sector: a comparative institutional analysis of five African countries. *Telecommun. Policy*, **43**(7), 101808.1–101808.8 (2019)
10. Yilmaz, A.: Comparative study for identification of multiple alarms in telecommunication networks. *Turk. J. Electr. Eng. Comput. Sci.* **25**(2), 677–688 (2017)
11. Wijekumar, K., Meyer, B.J., Lei, P., et al.: Supplementing teacher knowledge using web-based Intelligent tutoring system for the text structure strategy to improve content area reading comprehension with fourth-and fifth-grade struggling readers. *Dyslexia* **26**(2), 120–136 (2020)
12. Shinde, N., Ansari, S.: Intelligent bus monitoring system. *Int. J. Comput. Appl.* **168**(3), 27–30 (2017)



# Sports Video Moving Target Detection and Tracking Based on Particle Filter and Related Algorithms

Xiaocheng Wang<sup>1,2</sup>(✉)

<sup>1</sup> Faculty of Physical Education, Baotou Teachers College, Baotou 014030, Inner Mongolia, China

60516@bttc.edu.cn

<sup>2</sup> Inner Mongolia University of Science and Technology, Baotou 014030, Inner Mongolia, China

**Abstract.** How to improve the real-time, solve target tracking under complex background has always been a hot research topic. The purpose of this paper is to study sports video moving target detection and tracking based on particle filtering and related algorithms. Firstly, several common target detection algorithms are introduced. The principle of particle filter and particle filter tracking algorithm are introduced in detail, and its advantages and disadvantages are analyzed in detail. In order to improve the real-time performance of the algorithm, the average weight of the particles is used as the judgment condition to judge the distribution of the predicted particles. Only when the average weight of the particles is less than a certain threshold, it means that the average probability of the particle state being the true state of the target is small. At this time, the particle swarm optimization algorithm is used to optimize the predicted particles, and then the global extreme value when the algorithm converges is output as Estimation of the target state. The experimental results show that the tracking accuracy of the improved algorithm in this paper is 97%.

**Keywords:** Particle Filter · Sports Video · Moving Target · Detection and Tracking

## 1 Introduction

Visual information is an integral part of human perception. With the development of computer science, the use of computers to replace the human visual system to process environmental information has become important [1, 2]. Computer vision is an emerging discipline that studies how to effectively use computers to realize a human-like visual system, so that computers can analyze and understand the objective world like humans [3, 4]. In real life, people tend to be more interested in moving objects because dynamic images provide richer information than static images. It is significance to study moving target detection [5, 6].

Animation detection technology and monitoring technology are two key technologies to understand intelligent video surveillance [7]. Shahraki H proposed a small infrared

detection method called Grayscale Absolute Glass Disposable Window (SW-AAGD). The algorithm, derived from the interpretation of the fully differential algorithm, has a strong potential for mass gain and the ability to reduce background clutter in infrared images. A common challenge with gray is that due to the thin edges of small objects, a full difference algorithm is required in selecting the appropriate object and bottom window. To calculate the effectiveness of the proposed algorithm, the algorithm is applied to many real-time images, including the real world, noise-to-noise ratio (SCR), background noise factor (BSF), and receiver performance characteristic (ROC) in the calculation. The results demonstrate the effectiveness of membership assignment in the general function of search algorithms [8]. Wa Ng X proposed an automatic search method for HFSWR based on multidirectional writes. First, perform a clutch and statistical analysis to understand the different clutch areas over time. Analysis of actual data shows that the clutches in the HFSWR system are field and geometrical orientation dependent. Second, supported by this information, we develop a fixed-rate positive frequency index (MDDL-CFAR) based multi-directional dictionary, where the field information and geometric orientation are determined by the multi-dimensional dictionary [9]. Therefore, research on acquisition and tracking techniques becomes more important, and efforts are needed to increase the speed of scientific and practical applications [10].

This paper first briefly discusses several types of moving target detection algorithms, then introduces the theoretical basis of particle filter, and analyzes the particle filter algorithm. A moving object detection method based on visual attention mechanism. The improvement of the model and particle filter algorithm is proposed, and then the video sequence is preprocessed by using the visual attention mechanism with motion features to detect the region of interest. Finally, the improved particle filter algorithm is applied to the region of interest to segment the moving target. Through experiments, the rationality and effectiveness of the improved model are verified.

## **2 Research on Sports Video Moving Target Detection and Tracking Based on Particle Filter and Related Algorithms**

### **2.1 Commonly Used Moving Target Detection Algorithms**

#### **(1) Optical flow method**

When the object is moving, a velocity field will be formed on the surface, and the velocity vector of the moving pixel and the stationary pixel is different, so as to detect the moving target [11]. When the target moves, the corresponding sports field also moves. From the sports field, it can be found that the motion vector of the pixels on the moving target will be significantly different from the motion vector of the background pixels. According to the different characteristics of the motion vectors, the moving target and the background area are divided to complete the dynamic analysis of the image [12].

#### **(2) Difference method between frames**

By setting a specific threshold, if the change of pixel gray value exceeds the preset threshold, it will be marked as a moving target as a foreground pixel, otherwise it will be marked as a background pixel. The calculation of the interframe



difference method is limited to the difference operation of addition and subtraction. For the images captured when the camera position of the monitoring screen is fixed and the background is not complicated, the foreground detection effect is ideal.

### (3) Background difference method

During initialization, the background image is estimated and the background model is established through the prior knowledge images obtained in the first few frames. After the next image in the video is obtained, the background model of the 8-neighborhood in the background image is used to make a difference, so as to extract the foreground target. This effectively separates the background area from the foreground target.

## 2.2 Particle Filtering

Particle filter can deal with the problem that the model equation is nonlinear, that is, the probability of an event is represented by its corresponding frequency. In the process of filtering, when calculating the probability such as  $P(x)$ , the  $x$  in it is sampled, and the distribution of these sampling points is approximately used to represent  $P(x)$ , that is, by using the particle filter method Any form of probability can be processed. Due to the fact that the tracking before detection (TBD) is actually applied, the target observation model and the vector representing the current state of the target often have strong nonlinearity. In addition, in the real scene, the model noise is non-Gaussian, so a nonlinear filtering method needs to be used. However, the more the number of particles used in the particle filter algorithm, the closer the corresponding estimation result is to the real value, which approximates the optimal estimation. It is applied to the pre-detection tracking to detect and track weak and small targets, which has caused widespread concern at home and abroad focus on.

Although the addition of the resampling process solves the degradation of particle weights in the SIS algorithm, the accompanying defect is the sharp reduction of the sample diversity in the particle set, which is called sample depletion. This is because after many iterations, most particles are derived from a small number of particles with larger weights, while most particles with small weights are not copied, so the shield gradually disappears after several generations, and its preserved Location information is also lost. This will lead to the exhaustion of the sample, and it will fall into a local cycle, which cannot express the real trend of particle motion.

## 2.3 Particle Filter Algorithm Based on SIFT Feature and Template Update

When there are objects in the background with a similar color to the target, using the color attribute alone is not enough to describe the target. The SIFT property is extremely unique and can distinguish the target from other objects with similar colors. Integrating the SIFT property in the particle filter based on the color property can complement each other and make the algorithm more powerful. In the particle filter algorithm based on SIFT capability and mode update, the particle weight is determined by color capability and SIFT capability. Not all attribute vectors in the SIFT standard have corresponding attribute vectors, so the meanings of attribute vectors in the SIFT standard are different.

### 3 Investigation and Research on Sports Video Moving Target Detection and Tracking Based on Particle Filter and Related Algorithms

#### 3.1 Experimental Setup

The experimental video is the video of the training and competition of two avenue speed skaters. The video format is AVI, the video resolution is  $720 \times 576$  pixels, the frame rate is 25 frames per second, and the video sampling size is 24 bits. The trajectories of avenue speed skaters can be approximated as local linearity. Relevant parameter settings: the initial number of particles is 60, the size of the target area is  $10 \times 10$  (a certain part of the athlete's body), and the variance when calculating the weight is  $\sigma = 0.20$ . In this paper, the competition video is used to test the algorithm, and the experimental results are compared with the particle filter algorithm and PSO algorithm based on SIFT feature and template update.

#### 3.2 Improvement of Particle Filter Algorithm Based on SIFT Feature and Template Update

The particle filter algorithm has better tracking effect in most cases, so we do not need to use the PSO algorithm to optimize the predicted particles in every frame. When the rapid random motion of the target causes the predicted particles to be located at the tail of the observation model, which leads to the small weights of most predicted particles and the tracking failure of the particle filter algorithm, we only need to use the PSO algorithm to optimize the predicted particles and transfer the predicted particles to In this way, the algorithm can be improved.

Here we can use the average value of the predicted particle weights as a condition for determining whether the predicted particle distribution is good or not. Average the weights of all predicted particles in the  $k$ th frame, that is:

$$\bar{w}_k = \frac{1}{N} \sum_{i=1}^N w_k^i \quad (1)$$

Among them,  $\bar{w}_k$  represents the average probability that all particle states in the  $k$ th frame are the true state of the target, and then average the  $\bar{w}_k$  of the first 10 frames, namely:

$$\bar{w}_{av} = \frac{1}{10} \sum_{i=k+10}^{k-1} \bar{w}_i \quad (2)$$

If  $\bar{w}_k < w_{av}$  means that the average probability of all particles being the true state of the target is small, at this time, the sequence PSO algorithm is used to optimize the predicted particles, and then the global extremum position at the time of convergence is used as the estimation of the target state. If  $\bar{w}_k \geq w_{av}$  means that the average probability of all particles being the true state of the target is large, the weighted average of the predicted particles can be directly used as the estimation of the target state.

## 4 Analysis and Research of Sports Video Moving Target Detection and Tracking Based on Particle Filter and Related Algorithms

### 4.1 Sports Video Moving Target Detection and Tracking Results

This experiment successfully tracked 1000 frames, and the images captured during the tracking process (limited to space and only a few frames) are shown in Fig. 1. The first frame is the manually calibrated initial frame. The white point in the figure is the position of the tracking target generated by the algorithm (the first frame is the manually calibrated position to be tracked). Figure 1 shows the normal tracking of the athlete in the first frame; in the 300th frame, the black-clothed athlete occludes the target athlete. Due to the measures of occlusion and recovery of tracking after occlusion, the tracking results do not have much deviation from the actual position; The 400th frame shows that the target can still be tracked normally after being continuously occluded by the field staff, and the same is true for the 440th and 532nd frames. It can be seen from the screenshots that the size of the athletes shot in different frames changes greatly, and the lighting and color also change to a certain extent. However, because the HSV color histogram is used as the feature and observation model update mechanism, it can also be used for a long time. It tracks the target very well, and the algorithm is robust to tracking after continuous occlusion.



frame 1



Frame 300

**Fig. 1.** Tracking results of quasi-linear moving targets

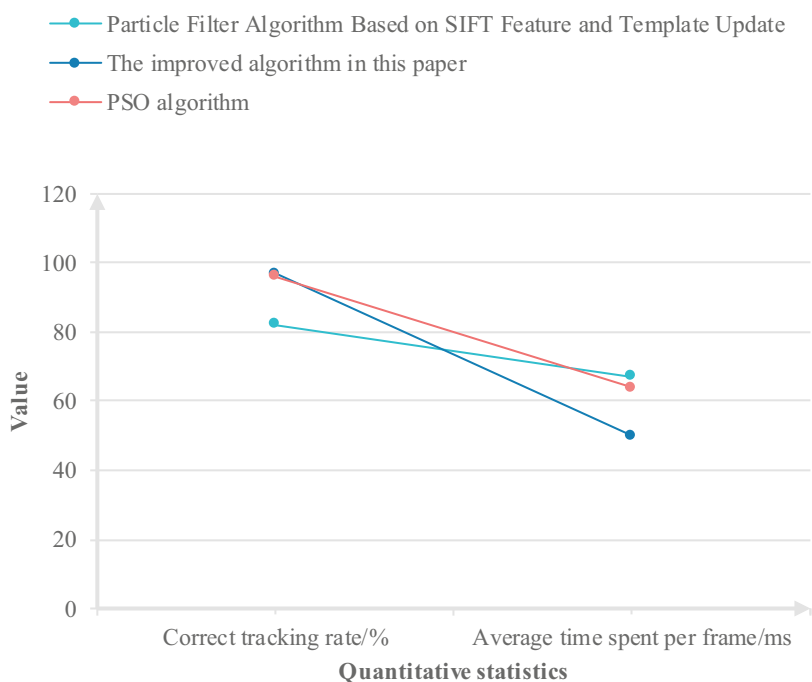
### 4.2 Algorithm Comparison

The PSO algorithm and the improved algorithm can track the target accurately in the whole video sequence. In the iterative process, the PSO algorithm updates the speed and

position of the particle according to the individual extreme value and the global extreme value, and integrates the observation value at the latest moment into the transition model to approximate the optimal importance sampling function. More in line with the true state of the target. After the PSO algorithm iteration, the particles are transferred to the region with higher probability value of the observation model, so as to obtain larger weights, so the tracking effect is improved.

**Table 1.** Quantitative comparison of tracking results in Video

Quantitative statistics	Particle Filter Algorithm Based on SIFT Feature and Template Update	The improved algorithm in this paper	PSO algorithm
Correct tracking rate/%	82	97	96
Average time spent per frame/ms	67	50	64



**Fig. 2.** Quantitative statistics of the tracking results of Video

The quantitative statistics of the tracking results of Video 1 are shown in Fig. 2. It can be seen that the tracking accuracy of the PSO algorithm and the progress of the algorithm in this paper are much higher than that of the particle filter algorithm based on SIFT features and standard updates. Since some particles are used in the PSO

algorithm, the real-time performance of the PSO algorithm is slightly higher than that of the particle filter algorithm. However, the PSO algorithm is a final iterative process. Using the PSO algorithm to process the predicted particles of each frame increases the importance of the algorithm and reduces the real-time performance of the algorithm. The advanced algorithm in this paper uses the PSO algorithm for optimization only when the predicted particle distribution is poor, so the real-time performance is higher than that of the PSO algorithm, as shown in Table 1. Therefore, the algorithm suggestion shows that this document is superior to the particle filter algorithm in terms of accurate tracking and real-time performance, and better than the PSO algorithm while ensuring accurate tracking and real-time performance.

Although this paper has done some research work on the theory, and also provided some improvements and methods, but due to the time relationship, there are still many problems that need to be further solved. The main problems are: in the detection of moving objects, due to the influence of light, background complexity, occlusion, etc., there are large uncertain factors in the actual detection and tracking process. Therefore, there will be errors in the detection of objects if there is an unexpected situation, so this point needs to be further studied. In the process of using particle filter algorithm, because the number and diversity of particles directly affect the tracking effect, when there are too many particles, it will increase the efficiency of computer operation, and even cause the computer to crash. Therefore, this problem can be studied from the aspects of particle utilization efficiency and intelligence.

## References

1. Kennedy, H.L.: Isotropic estimators of local background statistics for target detection in imagery. *IEEE Geoece Remote Sens. Lett.* (7), 1–5 (2018)
2. Lu, W., Lin, Q., Song, N., et al.: Target detection in intelligent reflecting surface aided distributed MIMO radar systems. *IEEE Sens. Lett.* (99), 1–1 (2021)
3. Jiang, X., Chen, W.X., Nie, H.T., et al.: Real-time ship target detection based on aerial remote sensing images. *Guangxue Jingmi Gongcheng/Optics Precis. Eng.* **28**(10), 2360–2369 (2020)
4. Kim, T.H., Shin, J.H., Lee, S.W., et al.: Design of waveform and signal processing of target detection for detecting closely spaced airborne targets in airborne radar. *J. Korean Inst. Electromag. Eng. Sci.* **31**(2), 154–164 (2020)
5. Nam, J.H., Koh, I.S., Park, S.H., et al.: Target detection and tracking simulation of air defense radar in high-density clutter environments using a real-time simulator. *J. Korean Inst. Electromagn. Eng. Sci.* **31**(8), 738–745 (2020)
6. Ruzhentsev, N.V., Volosyuk, V.K., Pavlikov, V.V., et al.: UAV target detection using radiometers of X, K, KA, AND W BANDS. *Telecommun. Radio Eng.* **79**(17), 1489–1512 (2020)
7. Nikaein, H., Sheikhi, A., Gazor, S.: Target detection in passive radar sensors using least angle regression. *IEEE Sens. J.* (99), 1–1 (2020)
8. Shahraki, H., Moradi, S., Aalaei, S.: Infrared target detection based on the single-window average absolute gray difference algorithm. *SIViP* **16**(3), 857–863 (2021). <https://doi.org/10.1007/s11760-021-02027-9>
9. Wang, X., Li, Y., Zhang, N., et al.: An automatic target detection method based on multidirection dictionary learning for HFSWR. *IEEE Geosci. Remote Sens. Lett.* (99), 1–5 (2021)

10. Ranchet, M., Morgan, J.C., Akinwuntan, A.E., et al.: Visual search and target detection during simulated driving in Parkinson's disease. *Acc. Anal. Prevent.* **134**(Jan), 105328.1–105328.6 (2020)
11. Southwell, B.J., Cheong, J.W., Dempster, A.G.: A matched filter for spaceborne GNSS-R based sea-target detection. *IEEE Trans. Geosci. Remote Sens.* (99), 1–10 (2020)
12. Kim, T.H., Ryu, S.H., Shin, J.H., et al.: Design of signal processing for target detection and tracking based on high pulse repetition frequency waveforms in airborne active electronically scanned array radars. *J. Korean Inst. Electromagn. Eng. Sci.* **30**(12), 951–963 (2019)



# Visual Reconstruction Design Based on Image Technology Emotion

Meng Hou<sup>(✉)</sup>

School of Art and Design, Modern College of Northwest University, Xi'an, Shaanxi, China  
qiaoqian19870619@163.com

**Abstract.** The visual representation of emotion in image technology is particularly important in today's society. Visual perception, which is mainly based on the changes in the audience's psychological emotions, has become a thinking derivative of digital image technology, and further technical implementation of image transformation, adjustment, repair, reconstruction, replacement, etc. Ingesting the characteristics of emotional appeal is the main direction of this research. With the continuous development of blockchain, virtual and information technology, the research on the content and emotion of digital images has become a hot spot in the design of visual reconstruction. A multi-dimensional computer-generated language spanning space is proposed to improve the methods of digital image acquisition, recognition, transformation and segmentation. Reconstruction design research on the semantic characteristics of computer-generated digital images and content emotional representations. Based on the multi-features of digital images, the differences in image shape, texture, color, motion and spatial domain features are analyzed, and the content emotion and semantic reconstruction of digital images are distinguished. In information transmission, multi-dimensional space capture, output, and quantitative analysis methods are used to establish a new sample arrangement matrix of digital images, select the best parameters for content expression, and reconstruct and process quasi-vision, combined with parameters to realize emotion recognition of digital images. The detection coefficient and dimension are determined by simulation and virtual technology, and the high recognition rate of digital image is obtained, and a good digital image technology display effect is achieved.

**Keywords:** Digital Image · Visualization · Visual Reconstruction · Content Emotion

## 1 Overview of Digital Image Technology

Image is objective existence, from the visual point of view it is directly or indirectly acting on the human eye image, and further according to the environment to produce visual perception of the entity. In the field of vision, an image is a projection of an objective reflection of the scene. The human perception system is typical for capturing and observing images. Images obtained by the visual perception system form images in people's hearts and minds, and further describe information. The concept of image is very

extensive, and it is very rich in emotional information for various objects in life, including landscape, pictures, life, animation, image, drawing, documents and so on. Image is the main visual source and information source for human to obtain information.

Image technology from a certain point of view, is the use of electronic media, equipment to complete a variety of work technology, which such as image media acquisition, resource acquisition, electronic coding, media storage, data interaction, information extraction, etc.. Through technology, images can be collected, synthesized and produced, images can also be displayed, modified and output, and even the transformation, adjustment, repair, reconstruction, replacement of the image technology implementation. The medium of technology implementation is computer. In order to carry out technical and artistic processing of all kinds of images collected, it is necessary to use computer to convert original images into digital images that can be edited, establish corresponding databases, and classify images, including the functional manifestation of retrieval [1]. The realization of digital image involves two aspects. On the one hand, the spatial sampling of image, the positioning of spatial coordinates, the realization of object image discretization; On the other hand, the amplitude quantization of the image produces quantization and discretization of the amplitude of the image, so as to carry out the analog-digital conversion of the image [2].

## 2 Visual Transformation and Segmentation of Digital Images

### (1) Digital image visual transformation

The transformation form of digital image is relatively abundant, is a digital image space generate visual impression of the pixel at the same time from a visual mapping to another visual location, from the point of view of space displacement, digital images can also be presented in a visual impression, from a visual expression space in a variety of modulus style change to another visual expression space [3]. The most common visual representation of pixel position mapping is mostly coordinate transformation, which includes the translation transformation of image data, the rotation transformation of image position, the scaling transformation of image pixels, the stretching transformation of image coordinates, and the shearing transformation of visual image. The spatial transformation of visual images is a technical means of special processing of images in a more efficient, fast and vivid way, and an effective way of diversified presentation of digital images. Further is to convert digital image through various technical means to a new space, fully implement the new visual expression, with the property of the new space unique to upgrades, more easily and quickly on digital image processing technology, and existing space to achieve consistency, and the technical data to get the results back to the original space needed for visual effects [4].

### (2) Digital image visual segmentation

Digital image segmentation is a computer vision processing technology, which analyzes and processes the collected images and uses various segmentation technologies to achieve visual levels of images [5]. In the design of visual information to the research and application of the image usually rally point of interest of audience, the audience is often strong interest on the part of the image information,



first through some information for other information, such as the main target in the image first image, the prospect of visual image, jumping visual elements, such as interest, in the transformation to the background information. At this time, part of the information corresponds to the areas with strong impact force, special and specific image in the image. In order to distinguish clearly, the designer will use technical means to decompose these elements. The segmentation of digital image is to divide the visual characteristics of the image into their own characteristic performance of the image language, extract the audience interested in the visual points, for technical processing. In addition, the image has the characteristics of color quantization processing, similar to the change of hue, purity, lightness, gray, saturation, texture and visual changes, image transformation can be a single image target, but also can be a plurality of visual areas.

Digital image segmentation on the visual language development always attaches great importance to the designers and the public [6], in various fields is also a variety of segmentation algorithm is proposed and the form method, on the form and the algorithm can be from two aspects, on the one hand, the image pixel segmentation techniques, according to the material has similarity, breakout and compatibility etc., the image to image segmentation. This kind of segmentation technology can use the interval brightness value of the image to achieve the boundary technology, can also use the image hue value to achieve the region technology; On the other hand, there are a variety of different technical processing strategies in the process of image segmentation. Serial strategies can be implemented for image segmentation, and parallel operations can be carried out during the implementation process. Visual judgment and technical decisions can be independently and output at the same time.

### **3 Visual Reconstruction Analysis of Digital Images**

Digital image features are diverse and can be divided into shape feature, volume feature, color feature, texture feature, style feature, situation feature, space feature, interaction feature and motion feature according to the way of technical implementation. Among them, texture is an inherent feature of the image surface of the object, so it is subordinate to an important attribute in the digital image editing area [7]. The digital image data analysis of the texture has the detailed description of the texture characteristics, showing the recognition data of the digital image texture. According to computer technology, texture is described by statistical method, structural method and spectral method. Shape feature is an important branch system of digital image analysis. The key point in display operation is to describe the various shape and shape features of each object in digital image. From the perspective of nature, shape feature description can be realized based on theoretical technology. With the help of theoretical technology, shape features of different digital images can also be obtained, and data module matching can be carried out after determining the shape similarity of images. The matching method is also the shape description symbol. According to the computer technology, the shape has the description feature, the transformation shape, the decomposition primitive, the association algorithm and so on. Movement characteristics according to the maturity of virtual technology in recent years, the movement of digital image data were analyzed, and contains the emotional change of the context of the scene here include quantitative detection

of trajectory data of safety inspection, the digital image of all sorts of changes are considering dynamic description in clear again, will be faster and distinguish rate, slow the movement of the object, Dynamic monitoring is realized by matching technology according to actual situation.

According to the principle of digital image reconstruction, the matching of digital image is analyzed and displayed, the image fusion phenomenon is mastered by the principle of registration, and Mosaic technology is incorporated with panoramic information. Based on advanced technology, the unknown visual image can be collected and input [8], and the existing visual image can be connected organically with its cognitive thinking and cognitive results, and then the unknown information can be interpreted with the known information, and finally the image input and output can be established. The multi-function matching of images can also be expressed at the level of abstract thinking. In the level of image pixels, color patterns are combined with digital templates to carry out irregular matching, and in the level of feature expression, target matching is carried out by using the thinking perception of the public. Image matching is an important technique in image reconstruction analysis and understanding. On the one hand, it can combine different images of the same scene to provide more comprehensive scene information; On the other hand, it can connect the previously unknown visual input with the previously existing cognitive results, so as to explain the unknown with the known and finally establish the interpretation of the input. Image matching can be performed at different levels of abstraction. Template matching can be used in pixel layer, perception matching can be used in feature layer, and visual habit can be used in image target layer to match elements.

## 4 Visual Design of Digital Image Based on Content Emotion

### (1) Content emotional design

Content emotional design is the basis of digital image and dynamic retrieval. Data collection and information growth provide a research hotspot for the public to absorb useful visual language. Visual image performance is gradually paid attention to. To volume features, color features and texture features, style, situational features, spatial characteristics, interaction and movement characteristics of data, to have access to emotional content matching visual characteristics, so as to search for similarities between digital image, digital information sharing, which can use semantic matching content [9]. Digital image through computer calculates registration related to the content of digital image, image matching, the content of the registration scope here is relatively narrow, mainly to digital image is displayed in different time or space data set up a relatively active space corresponding to the vision, make its content emotional echo relationship formation, especially in geometry in the form of digital image acquisition and data correction, The final digital image effect acquired by content emotional language is often reflected in pixels and thinking levels, giving play to the abstract nature and content attributes of digital images [10].

### (2) Form emotion design

Instinct emotion is an unconscious intuition of the audience for digital image recognition. Pleasant, interesting and beautiful emotions are a characteristic direction of visual expression design. Graphical forms can intuitively attract audiences in the field of first vision, especially the When the audience emotionally resonates, it will effectively convey the communication meaning of digital images [11]. Emotional specialization will also be affected by the individual differences of the audience. In order to meet the needs of the audience, the digital image design should be more interesting and human, pay more attention to human culture and humanity, and turn to the visual embodiment of diversification and liberalization. Form follows emotional design, but it cannot deny the primacy of function and content to determine form. Therefore, the emotion of form must always be attached to function and content, and mutual promotion can design digital images that are more in line with the emotional needs of contemporary audiences. Good for design services.

### (3) Functional emotional design

Functional emotional design can evoke a person's behavior or experience memories. In addition, it can also be a thinking expression of the designer's own heart and emotions. Digital images can resonate with the audience, and more will be reflected in the language of function. Strengthen the level of functional emotional design, strengthen in vision, content and creativity, not simply in the straightforward content of copywriting. A good emotional graphic language can express creatively and implicitly, and at the functional level, it will make the audience smile. The usability, ease of learning, ease of use, and ductility of digital images have derived the design of the audience behavior level, focusing on functional utility, which is achieved through interaction and technical means, so that we can establish an emotional demand model corresponding to the behavior layer to efficiently To solve the audience's behavioral level of using digital images, the functional design of using behavioral water for digital images is the level that is more concerned at present.

**Table 1.** Visual object parameter design

Visual Emotional Design	Content Emotional Design	Morphological Emotional Design	Functional Emotional Design
color reconstruction	Content build	morphological language	Interactive function
emotional appeal	Text connotation	emotional memory	use function
situational space	emotional resonance	mass characteristics	motion features
image data	emotional semantics	Texture	behavioral hierarchy
visual element	artistic conception	quality control	Extended function

Multi-dimensional digital image fusion is a new research and application following the development of virtual technology, 5G technology and blockchain technology in

recent years. The visual reconstruction design is obtained after comprehensive processing and analysis of image information obtained by different transmission media from multiple perspectives. In the process of digital image fusion, coordination, integration, diversification and optimization, more effective information content is extracted and new visual semantics are extracted to increase the reliability of technology transmission and the reliability and security of digital system [12]. In the process of information, content, form, feature and semantic fusion of Duofu digital image, it is necessary to register the data parameters of each group of visual objects participating in the fusion (See Table 1: Visual Object Parameter Design). The progress of multi-dimensional fusion registration of digital image has very high requirements. If the conversion spatial error exceeds one value or pixel, it will have a very serious impact. Not only the quality of digital image, the Angle of fusion, the implementation of technology, and the integration of content and emotion will produce deviations, so that the information transmission results are not as expected [13].

## 5 Conclusion and Outlook

Computer technology to realize visual image as the technology innovation constantly emerging a new language, visual perception and digital image technology for radiation across time and space for digital image visual refactoring provides continuously technical support, will be more intuitive, convenience, flexibility, frontier and virtualization features such as [14]. The advantages of real-time convenience, high precision data, semantic repeatability, diversified content and multi-dimensional vision of digital image make it more and more widely used in various fields and more and more life[15]. Virtual vision and digital image are closely linked, also has a large range of overlap on information coverage, while applied to different areas Angle is different, but in most situations and professional, cultural background and behavior using different content to focus on emotional terms and complement, so the technology of progressive for digital image visual reconstruction design laid a foundation of emotion, In the future information transmission, the multiple semantic representation of digital images will provide fresh elements for all walks of life.

## References

1. Chen, Y.: Application of cloud storage image processing technology in seed analysis. *J. Plant Genetic Resour.* **23**(02) (2022)
2. Dong, S.: Research on emotional innovation design of graphical system in public space - taking Beijing children's hospital as an example. North China University of Technology, Beijing (2021)
3. Yang, S., Kan, X.: Design of biological behavior analysis system based on image and IMU sensor. *Electron. Sci. Technol.* **35**(04) (2022)
4. Zhang, C.: Emotional design in digital media interface design. *J. Chifeng Univ. (Nat. Sci. Edn.)*. **31**(14) (2015)
5. Wu, Z.: On the superpixel segmentation method of digital image processing. *Print. Ind.* (01) (2022)

6. Sheng, F.: Research on the emotional communication phenomenon of positive propaganda in the digital media era. *News World* (01) (2022)
7. Wang, S.: Overview of digital image watermarking technology. *J. Hunan Inst. Technol. (Nat. Sci. Edn.)* **35**(01) (2022)
8. Zhang, B.: Emotional knowledge: emotional design in the digital display of Yuanshangdu site, inner Mongolia Normal University. *Inner Mongolia Auton. Reg.* (2021)
9. Chen, L., Yan, H.: In-plane displacement measurement based on digital holography and digital image correlation methods. *Acta Photonica Sinica* (2022)
10. Niu, Y., Zhao, Y., Li, X.: Review of research progress in passive forensics of JPEG images. *Sig. Process.* (2022)
11. Hong, Q., Li, X., Zhao, A.: Research on the emotional design of interactive art installations for young “Ant House Clan”. *Ind. Des.* (04) (2022)
12. Wang, M.: Research on product emotional design based on affordability theory. *Industrial Design* (03) (2022)
13. Li, G., Zhao, H., Lin, M.: Research on computer generated image recognition based on null space analysis. *Comput. Simul.* **39**(01) (2022)
14. Wu, Z.: Design of object-based digital image processing software. *Wirel. Internet Technol.* **19**(01) (2022)
15. Shi, K., Rao, M., Chen, W., Jin, L.: Based on spatial experience - research on the integration of emotional lens value in the classroom of “preliminary architectural design”. *Mod. Educ.* **7**(46) (2020)



# System of Cross-Border E-commerce Network Pattern Evolution on Account of Bayes-BP Algorithm

Lina Zhang<sup>(✉)</sup>

Department of Business Administration, Xi'an Eurasia University, Xi'an 710065, Shaanxi, China  
guomao8311@163.com

**Abstract.** In recent years, The export volume of China's cross-border e-commerce has been rising rapidly. Under the background of the Belt and Road Policy, the development pace of the world economy has been accelerating, and global logistics and other industries have been improving and developing. Cross-border e-commerce will become a key area of China's foreign trade in the future. Cross-border e-commerce networks are changing day by day. One of the most obvious details is that China's e-commerce imports and exports are increasing year by year. This paper studies the evolution analysis system of cross-border e-commerce network pattern based on Bayes-BP algorithm, and explains the working principle of the evolution analysis of cross-border e-commerce network pattern. The data show that the bayes-BP algorithm based on the evolution of cross-border e-commerce network pattern analysis system can reflect the change of cross-border e-commerce network pattern efficiently and accurately.

**Keywords:** Bayes-BP Algorithm · Cross-border E-commerce · Network Pattern · Evolution Analysis System

## 1 Introduction

Internet+ cross-border e-commerce system adopts information calculation and handles transaction volume data. The system analyzes e-commerce transaction data from multiple angles and in an all-round way. Faced with large-scale e-commerce data, the industry urgently needs to put forward a set of data information system with large collection quantity and strong analysis ability to solve the problems of e-commerce data transaction. Through the data system analysis of the data to achieve a scientific and correct analysis of the evolution of e-commerce pattern. The evolution analysis system of cross-border e-commerce network pattern based on Bayes-BP algorithm is conducive to the progress of cross-border e-commerce data analysis.

Bayes-BP algorithm has been studied by many scholars at home and abroad. In foreign studies, MustofaRL proposed to use the classification results based on dictionaries and the naive Bayes classifier algorithm to process the training data in the testing process. Generally speaking, the research stage of sentiment analysis includes the process of data

capture, text preprocessing, feature extraction and classification. The sentiment analysis results show that the proportion of social media users on Twitter about #newnormal is 33.19% [1]. DonnellanE proposes that there is a consensus that there is a difference between a well-trained simple bayesian classification algorithm to distinguish the curiosity and interest of free text definitions ( $n = 396$ ) and the use of cross-validation test classifiers in two sets of data (mainn = 196; Additionaln = 218) [2]. Rhernandez-sanjaime proposes a method based on variable decibel Bayesian learning algorithms that do not need to be performed on different numbers of groups in order to identify a fully fit data. Statistical theory is described, the performance of the proposed algorithm is evaluated using simulated data, and the two-step method is applied to macroeconomic problems [3].

The exploratory data analysis platform has multiple data processing methods. It displays e-commerce trade data through intuitive graphics, and then deeply analyzes the correlation of data and the correlation between data, so as to achieve the ultimate goal of data analysis with the greatest efforts [4, 5]. These data can be provided to customers for further research, data solidification, analysis and sharing. The analysis system of cross-border e-commerce network pattern evolution based on Bayes-BP algorithm improves the professional level of solving the analysis of cross-border e-commerce network pattern evolution [6, 7].

## **2 Design and Exploration of Cross-Border E-commerce Network Pattern Evolution Analysis System on Account of Bayes-BP Algorithm**

### **2.1 The Bayes - BP Algorithm**

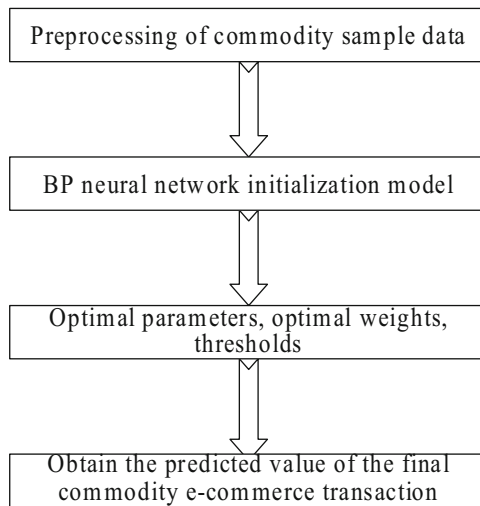
The definition of Bayesian neural network refers to regularization processing by adjusting the weights of neural network, that is, processing countless neural network values [2, 8].

A significant performance of BP neural network is the ability to classify objects, and the nonlinear model can be established. This algorithm is very suitable for many application scenarios [9, 10]. One step of BP neural network is the ability to repair errors in sample data. The error is dealt with by weight adjustment. At the same time, this algorithm has its defects. The convergence speed of the algorithm is very low when processing targets, so the algorithm is easy to fall into local minimization. E-commerce transaction volume data processing, using Bayesian neural network algorithm. The algorithm can classify and optimize the object, and determine the square deviation and mean value of the processed data. At this point, the problem of the algorithm model falling into local minimization needs to be solved [11, 12].

#### **2.1.1 Algorithm Flow**

Bayes-bp algorithm flow is as follows, as shown in Fig. 1:

- 1) Preprocessing of commodity sample data.
- 2) BP neural network algorithm formula for data processing, need to deal with a lot of parameters, in this case, need to select a reasonable algorithm formula.
- 3) Gaussian image was obtained according to Bayes model, and the algorithm was iterated for several times to obtain the best parameters, and finally the optimal weight and threshold value were obtained.
- 4) Use algorithms to train data and process the output results. If errors beyond the range are generated, the errors will be corrected. Finally, the predictive value of e-commerce transaction is obtained after algorithm training.



**Fig. 1.** Bayes-BP algorithm flow

## 2.2 Analysis System of Cross-Border E-commerce Network Pattern Evolution Based on Bayes-BP Algorithm

After the above analysis, the steps to realize the predicted volume of commodity e-commerce transactions include:

Selection of influencing factors [13, 14]. After a large amount of data analysis, the study shows that there are many factors that affect the trading volume of e-commerce commodities: For example, whether the e-commerce seller can return goods in compliance with regulations within a certain period of time, whether the commodity tax can be reduced or exempted, whether the price of goods is reasonable, whether the commodity category is rich, whether the e-commerce transaction speed is fast, It can avoid the interference of seasonal factors, whether the browsing quantity of commodity advertisements can be improved, and the commodity reserves of shopping cart, etc. Factors can be analyzed according to the actual situation of e-commerce transactions.



The concept of transaction rate of e-commerce goods is: if the inventory of e-commerce goods is enough, there must be no need to purchase in advance. In terms of data processing, the rate is assumed to be 0; If the inventory of e-commerce goods is insufficient, they need to purchase in advance. Different types of goods have different rates. For example, the ordering time of food can be set to 5–10 days and the speed is set to 1. The ordering time for women’s makeup products is set to 10–20 days, and the speed is set to 2; The ordering time for clothing category is 20–40 days, and the rate is set at 3.

If  $n = 13$ , the influence set is:  $X = \{x_1, x_2, \dots, x_{13}\}$ .

Elimination of interfering factors [15, 16]. Granger method was used to verify the influencing factors and the sales of e-commerce goods, remove the interfering factors and generate new factors:

$$X_{new} = \{x_1, x_2, \dots, x_i\}, i < n \tag{1}$$

Data preprocessing. Eliminate the interference factors and update the algorithm model.

Outlier handling: Outlier handling is a very common phenomenon. For example, when a shopping mall encounters a large holiday, it is necessary to carry out price reduction promotion activities. The sales of the activities will be different from the usual level, and the data will increase sharply, which generates outliers. For example, Tmall’s Double Eleven, JINGdong’s 618 and so on.

Vacant value processing: in the process of commodity data collection, it is found that some key values will be missing, this value has a certain importance, it is best to obtain this value. In this case, the method of complement can be used for data processing. This approach also has drawbacks that may reduce the accuracy of the data.

First, the missing values are analyzed and na complement values are filled, and then  $N$  data are obtained. Then,  $n$  sets of data are parsed so that  $n$  estimates are obtained. Assuming the estimate is  $X$ , the algorithm involved is:

$$\hat{w} = \frac{1}{N} \sum_i^N w_i \tag{2}$$

Then, the interpolation value can be calculated by:

$$\hat{X} = \sqrt{\left[ \left( \frac{1}{N} \sum_i^N X \right) + \left( 1 + \frac{1}{N} \right) \frac{1}{N-1} \sum_i^N (w_i - \hat{w}) \right]^2} \tag{3}$$

The complement method can accurately reflect its information and reduce the error.

Set reasonable parameters. Train the data and analyze the difference between the algorithm model and the traditional algorithm.

### 3 Research on the Effect of Cross-Border E-commerce Network Pattern Evolution Analysis System on Account of Bayes-BP Algorithm

Bayesian models have a lot of parameters, constants, variables, etc. Where the random variables are unknown, the research scheme is to carry out probability distribution. If sample data is missing, the formula of prior probability is:

$$p(\theta|H) = \frac{p(H|\theta)p(\theta)}{p(H)} \quad (4)$$

Type:  $p(H|\theta)$  is H a posteriori probability,  $p(\theta)$  and  $p(H)$  is the prior probability of the event.

In the algorithm model, if the sample data is given in advance, the posterior probability is calculated through prior probability to obtain the given data D, then the posterior probability formula is as follows:

$$P(\theta|D, H) = \frac{p(D|\theta, H)p(\theta|H)}{p(D|H)} \quad (5)$$

Among them:  $p(D|\theta, H)$  is the likelihood algorithm,  $p(\theta|H)$  is a priori algorithm,  $p(D|H)$  is constant. The posterior probability is modified in many ways. The results of the posterior probability algorithm can be processed more accurately.

The propagation mechanism of BP neural network has two directions: forward and back propagation. If the error is found to be large, reverse processing will be carried out. The difference value is apportioned to the neural layer according to a certain probability, and the nodes of the neural layer are processed at the same time.  $X = (X_1, X_2, \dots, X_n)$ , the resulting term is  $Y = (Y_1, Y_2, \dots, Y_n)$ , the output value is obtained after several iterations, that is, the predicted value of e-commerce trading volume:

$$a_{nj} = f_j\left(\sum_i w_{ji}a_{ni} + b_j\right) \quad (6)$$

where,  $a_{nj}$  is the output of the NTH sample J layer of the neural network;  $w_{ji}$  is the weight of neural layer I and J;  $a_{ni}$  is layer I output item;  $b_j$  is the threshold of layer J.

Based on the known samples, this paper calculates the prior probability of Formula (4), and then calculates the posterior probability according to Formula (5). After the posterior probability calculation, the prediction accuracy of the algorithm model is further improved. Finally, formula (6) is used to predict the trading volume of e-commerce.

#### 3.1 System Module Structure

##### 1) Development of public service functions of the platform

There are many public service functions of the platform, such as news and information service, service function, organization name processing, platform measurement standard detection, article safety processing, scientific and technological solutions, etc.

- 2) Development of platform detection service module  
Platform detection service module refers to customer-oriented detection processing, so as to solve many typical problems of customers. In the process of customs quarantine, the staff need to test a sequence of operations.
- 3) Development of mobile application functions on the platform  
There are many mobile applications used by users, such as mobile network, wireless network and so on. Use the system functions to test the application, provide inquiry, retrieve the inspection results, provide inspection services and other links.
- 4) Development of big data analysis functions of the platform  
The platform adopts big data mining tools to collect data, and the collection method is to collect useful data through relatively innovative big data technology. Then, model analysis is carried out on these data through data algorithm, so as to call up the data analysis results.

#### 4 Investigation and Research Analysis of Cross-Border E-commerce Network Pattern Evolution Analysis System on Account of Bayes-BP Algorithm

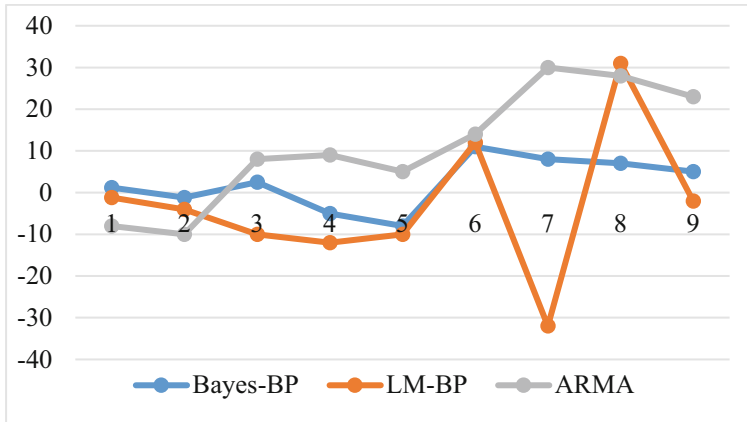
The main components of this test include the server, WEB front-end. The application tool for WEB front-end is Microsoft Visual Studio2008. This application tool is the development tool of a well-known software company and covers all tools of all software development cycles. These tools can be used to develop high-performance, data secure, and robust system architectures. Java language is used in the bottom layer, which can improve the security of database data processing. Meanwhile, mobile development tools such as Android Studio and Xcode are used to process the data. This development architecture can be used for processing software applications on Android and IOS platforms. Software platform data exchange is applicable to the data socket technology, at the same time, SSL data encryption for data, so as to ensure the security of data exchange.

**Table 1.** Influencing factors and ADF test of sales volume

variable name	t statistic
Comment on the amount	- 6.742
unit price	- 5.384
Good rating rating	- 3.734
7 days return	- 4.413
Season	2.961 2

As shown in Table 1, e-commerce influencing factors and ADF test of sales volume are shown. Variable name includes Comment on the amount, unit price, Good rating rating, 7 days return, season, and season. Their T statistics are shown in the table.

Bayes-bp algorithm and other algorithms LM-BP and ARMA are used to calculate the error percentage effect (%), and the effect diagram is shown in Fig. 2.



**Fig. 2.** Bayes-BP algorithm and other algorithms LM-BP, ARMA error percentage effect

As can be seen from the figure, the error of Bayes-BP algorithm is smaller than the other two algorithms, indicating that bayes-BP algorithm model has the best effect in e-commerce data prediction.

The data prove that the bayes-BP algorithm based cross-border e-commerce network pattern evolution analysis system accurately shows the error value of cross-border e-commerce network pattern evolution, thus improving the accuracy of e-commerce network pattern evolution.

## 5 Conclusions

In this paper, the e-commerce transaction volume for sample collection, collection. By processing the sales volume of goods, the time series characteristics of sales are collected. Through the consumption characteristics, the cross-border e-commerce system based on Byes-BP neural network is designed. At the same time, the relevant data of the buyer and consumer are examined. The network pattern evolution analysis system of cross-border e-commerce based on Bayes-BP algorithm effectively improves the proportion change of each economic form of cross-border e-commerce.

**Acknowledgements.** This work was supported by key projects of higher education teaching reform in Xi'an Eurasia University, Item number: 2021ZD005.

## References

1. Mustofa, R.L., Prasetyo, B.: Sentiment analysis using lexicon-based method with naive bayes classifier algorithm on #newnormal hashtag in Twitter. *J. Phys.: Conf. Ser.* **1918**(4), 042155 (2021)
2. Donnellan, E., Aslan, S., Fastrich, G.M., et al.: How are curiosity and interest different? Nave Bayes classification of people's beliefs. *Educ. Psychol. Rev.* **34**(1), 73–105 (2021)

3. Hernández-Sanjaime, R., González, M., Pealver, A., et al.: Estimating simultaneous equation models through an entropy-based incremental variational Bayes learning algorithm. *Entropy* **23**(4), 384–384 (2021)
4. Narmak, T.: Classification of sleep Apnea using EMD-based features and PSO-trained neural networks. *Biomed. Eng./Biomedizinische Technik* **66**(5), 459–472 (2021)
5. Gill, N., Bayes, A., Parker, G.: A review of antidepressant-associated hypomania in those diagnosed with unipolar depression—risk factors, conceptual models, and management. *Curr. Psychiatry Rep.* **22**(4), 1–8 (2020). <https://doi.org/10.1007/s11920-020-01143-6>
6. Huda, N., Hasbi, M., Susyanto, T.: Seleksi Penerima Bantuan Pangan Non Tunai di Desa Menggunakan Metode Nave Bayes dan Simple Additive Weighting. *Jurnal Ilmiah SINUS* **19**(1), 39 (2021)
7. Rahmadi, M., Kaurie, F., Susanti, T.: Uji Akurasi Dataset Pasien Pasca Operasi Menggunakan Algoritma Nave Bayes Menggunakan Weka Tools. *JURIKOM (Jurnal Riset Komputer)* **7**(1), 134 (2020)
8. Sombra, T., Santini, R., Morais, E., et al.: Quantitative analysis powered by Nave Bayes classifier algorithm to data-related publications social-scientific network. *Int. J. Innov. Educ. Res.* **8**(6), 205–217 (2020)
9. Tesoro, J.C.: A semantic approach of the nave bayes classification algorithm. *Int. J. Adv. Trends Comput. Sci. Eng.* **9**(3), 3287–3294 (2020)
10. Utami, D.Y., Nurlalah, E., Hikmah, N.: Attribute selection in Naive Bayes algorithm using genetic algorithms and bagging for prediction of Liver disease. *J. Inform. Telecommun. Eng.* **4**(1), 76–85 (2020)
11. Rino, R.: The comparison of data mining methods using C4.5 algorithm and Naive Bayes in predicting heart disease. *Tech-E* **4**(2), 44 (2021)
12. Jaiswal, M., Das, S., Khushboo, K.: Detecting spam e-mails using stop word TF-IDF and stemming algorithm with Nave Bayes classifier on the multicore GPU. *Int. J. Electr. Comput. Eng.* **11**(4), 3168 (2021)
13. Li, S.: Structure optimization of e-Commerce platform based on artificial intelligence and blockchain technology. *Wirel. Commun. Mob. Comput.* **2020**(12), 1–8 (2020)
14. Ko, B.M.: Major digital trade agreements and the implications for the Korea-Singapore Digital Partnership Agreement (DPA). *J. Int. Trade Comm.* **16**(6), 215–233 (2020)
15. Lipysh, L., Mokhniuk, A.: E-business in Ukraine: peculiarities, tendencies, prospects. *Econ. J. Lesia Ukrainka East. Eur. Natl. Univ.* **1**(21), 74–79 (2020)
16. Sosnov, T., Pasko, A.: Online platforms in the system of modern international digital trade. *E-Management* **3**(2), 63–69 (2020)



# Construction of Internet Financial Risk Early Warning Model Based on Data Mining Algorithm

Jingshuang Liu<sup>(✉)</sup> and He Yu

Changchun College of Electronic Technology, Changchun, China  
aiueo1233@163.com

**Abstract.** Due to the continuous progress of network technology, traditional financial service products are gradually transitioning to digitization, which not only completes the development of traditional financial service products and reasonable risk management, but also helps to improve big data mining technology and relevant technology on the basis of strengthening management, provide innovative business model, and realize the reasonable direction of HP and technology upgrading of financial service products. At the same time, it is fully found that online financial services can give full play to the innovation ability of the platform, reasonably avoid business risks, develop effective investment and financing channels for small and medium-sized enterprises, bring support and protection to the short-term growth process of small and medium-sized enterprises, and enable investors to reduce investment difficulties. In addition, under the influence of the network, the financial service model has distinct characteristics, and the data algorithm continues to promote its stable and sustainable growth.

**Keywords:** Data Mining · Online Finance · Risk Warning

## 1 Introduction

The establishment of online financial service platform provides a great risk to the development of China's market economy, and there is a huge problem of platform bankruptcy in the early stage of the development of China's financial network platform. In addition, criminals use China's financial network platform for illegal investment activities. These factors are not conducive to the healthy development of China's market economy. In July 2015, the people's Bank of China and other ten departments issued the guiding opinions on promoting the development of online financial services, which clearly expounded the basic meaning of online financial services: "Internet finance is an emerging financial business model in which Chinese traditional financial institutions and Internet enterprises use Internet technology and information and communication technology to carry out capital financing, trading, fund management and information intermediary services". However, network financial service does not mean the simple combination of network technology and financial industry, but an effective way to achieve financial security.

## 2 Internet Finance Concept

As a subset of the integrity system of the whole society, the integrity system in network financial institutions is a new force that can not be ignored. Only by making full use of the booming network information technology and cutting-edge data mining technology can we promote the high-speed development of the integrity system of network financial institutions and judge, warn and avoid the risks of network financial institutions [1].

The innovation of Network Finance on financial risk structure not only creates a competitive capital trading market centered on financial users, but also makes financial users bear the brunt in the process of financial risk diffusion and utilization. The reform of the transaction structure that follows the rationality of the financial market and restores the nature of the financial market has triggered a legal regulation path of financial risk to realize the risk adsorption ability of financial consumers and match the risk of financial assets. The impact mechanism of the consumer protection law of financial institutions on the risk exposure and risk diffusion of China's online financial market and the new risk legal regulation paradigm that encourages fair competition in the market also reflects and tests whether the consumer risk adsorption capacity of financial institutions that abide by the rule of law is in line with the market rational basis composed of market investment and financing methods and fair value, This will also become the logical center of the risk legal regulation approach of the consumer protection law of China's online financial institutions, which will help to achieve the goal of the new financial law of establishing a fair value market [2, 3].

## 3 Preprocessing of Platform Data

In the process of analyzing the data and information preprocessing of the platform, it is also necessary to comprehensively analyze the possible illegal fields in the loan data of the network financing platform, including the characteristic fields related to creditors, such as age, income amount, major, family situation, professional title, work field, company attribute and other related aspects. In addition, it also includes the customer loan information content in relevant fields such as the proportion of arrears to collection, the amount of monthly loan repayment, the loan principal and the number of repayment periods. Because these financial data contents come from various online financial platforms, and the financial data contents have the characteristics of authenticity or richness, the financial data contents must be cleaned and generalized before data mining. In order to clearly grasp the key financial data content fields, it has laid a favorable foundation for the orderly progress of data information preprocessing, so that the data mining information content can be gradually transformed into the direction of regularity and mining [4].

### (1) Data generalization

Through the comprehensive classification of the data information segments in the online financial platform, it can be found that the data involved in these fields generally presents a continuous situation and exists in the form of data at the same time, while the data in non data form will also produce continuous field

value changes, and there will be many different types of work when processing these data. According to the decision tree algorithm, the final ideal processing effect will be more inclined to obtain the data value based on discrete types, and therefore it can be easier to become a tree structure processing method based on generality, so as to facilitate the classification and comprehensive management of relevant data values. Therefore, the continuous data value can be discretized and classified through different standards and specifications, as well as the combination of quantitative and qualitative management ideas, so that the overall granularity of data can be gradually increased. When dealing with different text fields, we must first process the data according to the specific characteristics of the real world, fixed length box or tag number processing, and then generalize the data according to the overall characteristics of the text field, so that all different types of text fields can have the characteristics of generality and mining at the same time [5].

## (2) Data cleaning

By analyzing the data quality and mining characteristics of network financial platform based on data mining technology, it can be found that the research on digital information cleaning can be divided into two aspects: useless information elimination and missing data supplement. On the other hand, because there will be important fields or information value gaps in information collection, when the gap information is numerical, the relevant data information can be supplemented by means of average or average of the same type; When the information of the gap is non digital, it can be supplemented by decision tree or Bayesian regression, so as to supplement the most likely digital information to the place of the gap. In addition, useless digital information must be eliminated according to the regular design principle, and the focus is to manage it according to the filling status of the borrower's virtual information. Since most online financial service platforms do not strictly follow the market supervision and evaluation system, they have made detailed management on the filling of false data [6].

## 4 Application of Big Data Algorithm in Financial Industry

In today's Internet information age, Internet data information discovery technology, because of its great data analysis and information extraction power, has led to great transformation and transformation in many industries, promoted the rapid growth of various industries and market economic systems, and constantly developed new science and technology businesses in the development of new fields. Data mining technology has been widely used in the financial field. From macro to micro economy, from the overall development trend to local economic development, from current data analysis to future prediction, the specific application of data mining technology in finance has achieved very good results. At present, the specific application of data mining technology in is mainly reflected in the following aspects [7].

First, risk management. The financial service industry will come to different conclusions under the influence of various factors when measuring the credit efficiency and customer credit rating. The methods of data mining, including feature selection and attribute association analysis, can determine the related factors and non related factors.



For example, there are many reasons related to the problem of credit repayment, but the core reason is the ratio of interest to income. Commercial banks may adjust their credit supply strategies. At the same time, through the analysis of historical data in this regard, it is helpful to manage business problems.

Second, industry correlation analysis. By using data mining technology to analyze a large number of transaction data information of the company and establish the corresponding mathematical model, we can find the market trend, customer demand change and other trend signals behind many transaction information. The company can use these information systems to gain insight into the relevant change trends of the financial service industry and timely adjust the company's operation strategy to support the company to make reasonable evaluation and decisions. For example, commercial banks keep a wealth of customer transaction information. Using association analysis, they can find the hidden relationship network in the database, which helps to tap and analyze the guests' usual income, purchasing power, purchasing habits and other aspects, and find the potential needs of guests.

Third, customer relationship management. This refers to the integrated marketing implemented by enterprises to customers through the use of data mining technology. It is the realization of technical implementation and management of enterprise sales with customers as the core. It is mainly used to attract and leave useful managers of the enterprise. Customer relationship management technology mainly includes the following four aspects: customer collection, customer classification, customer profit potential research and customer service quality maintenance. Data mining technology can obtain the conventional signals formed in the communication between customers and financial companies, and study customer behavior. From this technology, we can further explore and find the law of customers' psychological activities, so as to further improve the industry share of products and the comprehensive competitiveness of small and medium-sized enterprises.

Fourth, strictly suppress and prevent online financial fraud, money laundering and market economy crimes. There is usually huge money circulation among customers in the field of financial services, resulting in many economic crimes such as financial fraud and money laundering, including stealing user secrets through card theft, malicious overdraft, counterfeiting bank cards, and money laundering. Criminal activities in financial services have become one of the thorny problems encountered by the industry today, which also requires strict supervision in the field of financial services to prevent financial risks and identify fraud in financial institutions.. To detect bank crime, the most important thing is to combine a large amount of relevant information, such as industrial economy, personal credit and other risk data, and then use a variety of different data mining means and methods to find out abnormal behavior, and use the frequent large cash flow turnover in the short term or a few people to estimate its harm degree, so as to provide reference data for decision-makers to avoid risks. In addition, data mining technology can dig out their hidden commonalities by studying these fraud activities, and can timely issue warnings to enterprises and urge companies to do a good job in management [8].

### 5 Construction of Risk Early Warning Model

Firstly, the risk warning index reflecting the characteristics of the online financial industry is constructed, and the significance of the index is detected. If the indicator is not obvious, the indicator will be eliminated; If the index is obvious, the principal component analysis is used to reduce the dimension and refine the main components. Then, some information of the extracted principal components is input into BP neural network as a training set, so as to form a crisis early warning model of Internet financial enterprises using BP neural network. Each principal component can be regarded as a linear combination of original variables, and there is no correlation between them, so as to reduce the dimension of variables. The specific algorithm steps are as follows:

$$Y_{ij} = \frac{X_{IJ} - EX_J}{\sqrt{DX_J}}$$

For the calculation of the correlation coefficient matrix of matrix R, the formula is:

$$r_{ij} = \frac{\sqrt{\sum_{k=1}^n (X_{ki} - X_i)^2 (X_{kj} - X_j)^2}}{\sqrt{\sum_{k=1}^n (X_{ki} - X_i)^2 \sum_{k=1}^n (X_{kj} - X_j)^2}}$$

Then, the test set is injected into the trained model to detect the stability of the model. Finally, input the overall information, and then draw the final early warning conclusion. The specific process is shown in Fig. 1.[9]The basic principle of principal component analysis is to use the method of mathematical transformation to transform the original multiple variables into several variables that can represent most of the information of the original variables, which is called principal component analysis [9].

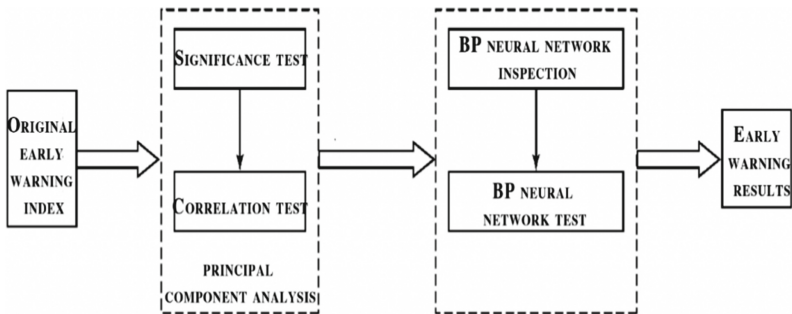


Fig. 1. Risk early warning portfolio model

Early warning target is also the core of risk early warning model. The main reason for the mutual influence of China’s financial services enterprises is the complex risks of the Internet. Therefore, after establishing the early warning model, the enterprise should, according to the enterprise characteristics and business characteristics, select 18 main indicators from the five aspects of solvency, profitability and cash flow as the benchmark

PRIMARY INDEX	SECONDARY INDEX	PRIMARY INDEX	SECONDARY INDEX
SOLVENCY	CURRENT RATIO V1	CASH FLOW	CASH RECOVERY RATE OF TOTAL ASSETS V10
	QUICK RATIO V2		OPERATING CASH FLOW PER SHARE V11
	ASSET LIABILITY RATIO V3		TOTAL ASSET GROWTH RATE V12
PROFITABILITY	INTEREST COVER V4	BUSINESS DEVELOPMENT ABILITY	NET ASSET GROWTH RATE V13
	COST PROFIT MARGIN V5		OPERATING PROFIT GROWTH RATE V14
	RETURN ON NET ASSETS V6		INVENTORY TURNOVER V15
	NET PROFIT MARGIN ON SALES V7		NET ASSETS PER SHARE V16
CASH FLOW	WORKING CAPITAL V8	INDEX PER SHARE	PROFIT DISTRIBUTION AT THE END OF EACH SHARE V17
	CASH SALES RATIO V9		EARNINGS PER SHARE V18

Fig. 2. Risk early warning indicators

of enterprise risk assessment, and make a comprehensive evaluation from the five aspects of operation and development ability and stock indicators, as shown in Fig. 2 [10].

At present, as far as online financial institutions are concerned, the scope of risks they can identify is not comprehensive, and they face different business risks under different business modes. But in general, this paper mainly gives the following opinions on the risk management of Internet Finance: first, improve the risk mechanism of enterprises, Improve the transparency of integrity “To minimize the harm caused by trust bird information asymmetry to enterprises. At the same time, online financial institutions should pay attention to the identification and integrity evaluation of basic investors by enterprises, and strengthen efforts to eliminate the disadvantages caused by the virtuality of the Internet, so as to reduce such losses. Second, further improve the industry control system, and improve the early warning level of online financial risks by improving the openness to the whole field and clarifying the legal bottom line. Online financial institutions Institutions cannot legally raise funds, do not occupy enterprise funds and other illegal acts. We will improve the self-discipline system of the Internet financial industry and the supervision of local government departments over the Internet financial industry, so as to promote the sound development of the Internet financial industry. In recent years, there have been some deviations from the norms in the field of Internet finance. After that, there have been many phenomena, which have done harm to all stakeholders. Therefore, we must effectively control the risk before it occurs, so as to minimize the loss. Therefore, we must build a more sound Internet financial risk prevention model [11–15].

## 6 Conclusions

The sustainable development of Internet financial platform has promoted the financial industry to gradually improve the traditional management mode. Under the far-reaching influence of the Internet era, we are timely aware of the great significance of the establishment of online financial service platform, and can carry out corresponding standardized management of market competition. At this stage, we must fully consider the important factors of big data mining technology for the development of online financial service platform, and use big data analysis technology to promote the long-term development

of finance. The financial industry can actively adapt to the new situation of the market under the new financial market management mode, establish scientific and reasonable risk control measures, and fully grasp the reality and application value of big data mining technology in the financial field and financial insurance field under the guidance of the principle of step-by-step, so as to minimize the probability of major risk problems and effectively manage the business risks in the platform.

## References

1. Liu, C.: Authorized public auditing of dynamic big data storage on cloud with efficient verifiable fine-grained updates. *IEEE Trans. Parallel Distrib. Syst.* **25**(9), 2–7 (2014)
2. Chen, C.L.P., Zhang, C.-Y.: Data-intensive applications, challenges, techniques and technologies: a survey on Big Data. *Inf. Sci.: Int. J.* **275**, 314–347 (2014)
3. Zhang, D.: Granularities and inconsistencies in big data analysis. *Int. J. Softw. Eng. Knowl. Eng.* **23**(6), 887–893 (2013)
4. Kambatla, K.: Trends in big data analytics. *J. Parallel Distrib. Comput.* **74**(7), 3–9 (2014)
5. Dobre, C., Xhafa, F.: Intelligent services for Big Data science. *Future Gener. Comput. Syst.: FGCS* **37**, 267–281 (2014)
6. Lesk, M., Data, B., Brother, B., Money, B.: *IEEE Secur. Priv.* **11**(4), 85–89 (2013)
7. Im, J.-K.: A time-dependent proportional hazards survival model for credit risk analysis. *J. Oper. Res. Soc.* **63**(3), 306–321 (2012)
8. Chen, T., He, J., Wang, J.: Bifurcation and chaotic behavior of credit risk contagion based on FitzHugh–Nagumo system. *Int. J. Bifurcat. Chaos* **23**(07), 1350117 (2013)
9. Allen, D.E.: Modelling tail credit risk using transition matrices. *Math. Comput. Simul.* **93**, 3–10 (2013)
10. Ding, S.: An optimizing BP neural network algorithm based on genetic algorithm. *Artif. Intell. Rev.: Int. Sci. Eng. J.* **36**(2), 153–162 (2011)
11. Weiss, G.N., Pelger, K., Horsch, A.: Mitigating adverse selection in P2P lending: empirical evidence from Prosper.com. Working paper, TU Dortmund, Germany (2010)
12. Collier, B.C., Hampshire, R.: Sending mixed signals: multilevel reputation effects in peer-to-peer lending markets. In: *Proceeding of the 2010 ACM Conference on Computer Supported Cooperative Work*, no. 2, pp. 197–206 (2010)
13. GAO-11–613: Person-to-person lending: New Regulatory Challenges Could Emerge as the Industry Grows (Report to Congressional Committees), 7 July 2011. <http://www.gao.gov/>
14. Lin, M., Prabhala, H.R., Viswanathan, S.: Judging borrowers by the company They keep: Social networks and adverse selection in online Peer-to-Peer lending. *Manag. Sci.* **59**(1), 17–35 (2012)
15. Person-to-Person Lending, New Regulatory Challenges Could Emerge as the Industry Grows. The Report of Peer-to-Peer Network Lending’s Development and Supervision from Government Accountability Office to Congress (2012)



# Research on an Intelligent Logistics Blockchain Consensus Algorithm

He Yu<sup>(✉)</sup> and Jingshuang Liu

Changchun College of Electronic Technology, Changchun, China  
yuhe18845616325@163.com

**Abstract.** Aiming at the increasingly urgent problems to be solved in the traceability of big data and goods such as logistics, capital flow, information flow and business flow of logistics enterprises, this paper provides blockchain intelligent logistics model and blockchain consensus algorithm for the non repudiation and capital security of supply chain authentication and anti-counterfeiting technology of logistics enterprises. This paper mainly analyzes the technical problems and transaction modes encountered in the process of traditional logistics transactions, and puts forward the basic idea of using cloud computing and blockchain technology to overcome their shortcomings. According to the technical characteristics and characteristics of cloud computing and blockchain, combined with their respective advantages, this paper introduces the technical advantages of intelligent logistics blockchain. The definition and concept of intelligent logistics blockchain are proposed, and the calculation model based on the consistency algorithm of intelligent logistics blockchain is given.

**Keywords:** Intelligent Logistics · Blockchain · Consensus Algorithm

## 1 Introduction

With the development of digital currency, people pay more and more attention to blockchain technology, and the core of blockchain technology is paid more and more attention. It is particularly important to deeply study the consensus understanding mechanism of the core technology of blockchain technology. The application of blockchain technology in the Internet of things is one of the research hotspots. Consensus mechanism is also one of the core technologies of blockchain, which is in the stage of decentralization. It has an important impact on the key technologies of the Internet of things in terms of value anchoring, transaction processing speed, transaction confirmation delay, security and efficiency, payment verification delay, information security and scalability. The audit core transaction mode of the independent e-commerce certification center will face information security problems: first, it is vulnerable to security challenges, resulting in the disclosure and disclosure of a large number of personal confidential information; Second, the authenticity and traceability of the user's identity cannot be verified, and the unique identity of both parties to the transaction cannot be guaranteed; Third, it shall be determined by both parties of the transaction; Fourth, due to a large number of

unreliable business data, it is difficult to obtain complete and systematic logistics transaction information. The distribution of business data will be difficult. At the same time, the information security of logistics transaction, the security of distribution trade and the trust of upstream and downstream customers need to be solved urgently. Therefore, using blockchain consensus algorithm to solve the decentralization of logistics transactions and trust between users has important scientific commonality, great scientific and technological significance and economic value.

## 2 Definition of Consensus Algorithm

Consensus question is a classic question in the research fields of social science, computer science and other computer science and technology. It has a long research history and has a long research history. At present, the research papers recorded in the literature can be traced back to 1959 at least. In 1959, it was jointly carried out by the RAND Corporation and with Edmund Eisenberg and David Gale of Brown University College [1]. The main topic is to study how to form a group when a group of individuals have their own subjective probability distribution in a specific probability space, How to establish a consensus probability probability distribution. Then, ask questions. Since then, the consensus question has gradually aroused questions and gradually attracted extensive and general research interest in sociology, management, economics, especially management theory, market economy, especially computer science and other disciplines [2]. The generation block of the consensus mechanism selects the node with the greatest contribution through the contribution algorithm, and gives the calculation formula of the contribution algorithm:

$$MC = \sum_{n=1}^N \omega_2 * \frac{\Delta T^2}{\omega_1} + (KC)^3 + \sum_{M=1}^M \omega_3 + \omega_4 * (T_1 - T_2 - T_3)$$

The consensus algorithm of the blockchain system will have to be applied in a more complex, open and distrustful Internet environment with complexity, openness and lack of confidence, because there are more nodes and malicious Byzantine nodes may exist and may appear. Therefore, although some distributed consistency algorithms such as VR and Paxos have been clearly put forward in the early 1980s, how to cross the Byzantine fault tolerance gap and design and establish a simple distributed consensus algorithm is still arithmetic, but it has always been one of the difficult problems in the field of distributed computing.

Generally speaking, because the nodes of the blockchain system have the characteristics of distributed, autonomous, open and free access, most of them adopt the characteristics of P2P random access. Therefore, people mainly organize and distribute the global system through p-2-p network to establish nodes around the world participating in data verification, information authentication and bookkeeping. P2P node. Because each node in the p-2-p network system has equal and relative status, and connects and interacts with each other in a flat topology structure, there is no topology structure that connects and interacts with each other, and does not produce a centralized special node and hierarchical structure. Each node will assume the network architecture. Therefore, each node will perform online routing, verification and authentication block data, dissemination of

information, transmission of block data Basic functions such as discovering new nodes and other information, and appearing new nodes [3]. The blockchain system adopts a specific economic incentive mechanism to ensure that all nodes in the distributed system have the motivation to participate in the data capacity, participate in the generation, formation and verification process of information blocks, and allocate the digital cryptocurrency formed in the process generated by the consensus process according to and according to the work volume actually completed by the nodes, The consensus algorithm is used to select specific nodes, calculate and filter special links, and add new blocks to the blockchain Medium The vigorous development of a series of blockchain technology applications represented by bitcoin not only highlights the rise, but also reflects the importance, application necessity and use value of blockchain technology. Therefore, the consensus of blockchain system has also become a new research hotspot [4]. According to the probability shown in the figure, conduct discrete-time random walk to tips and find the established goal:

$$P_{IJ} = \text{EXP}(-\alpha(K_I - K_J)) \left( \sum_{Z \rightarrow I} \text{EXP}(-\alpha(K_I - K_Z)) \right)^{-1}$$

### 3 Intelligent Blockchain Technology

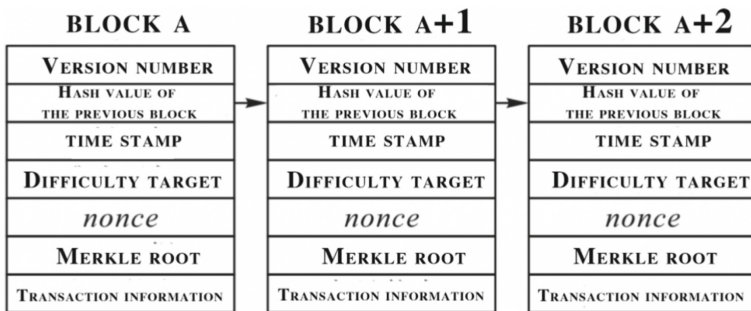
#### (1) Blockchain

At present, there are several blockchain alliances in the world. Each alliance will define blockchain according to its own research focus. In Wikipedia, blockchain is defined as a distributed database that can be used to manage data in time order and ensure that data cannot be tampered with. The concept definition of blockchain was first mentioned in the paper “bitcoin: a peer-to-peer electronic cash system” written by the researcher of “Nakamoto” published by a Japanese scholar with the pseudonym of “Nakamoto” in 2008. Because the blockchain stores a large amount of transaction information and a huge amount of exchange data, it is equivalent to a database. Literature presentation system. Therefore, this paper believes that blockchain is a data structure that can link blocks from back to front and connect them from back to front according to the time sequence. Asymmetric encryption technology is used to ensure that it can not be protected, modified, tampered with and forged. Table 1 is the account book. For example, Table 1 shows the block chain structure diagram, where  $a \geq 0$  (when  $a = 0$ ,  $a \geq 0$  (when  $a = 0$ , it represents the creation block), and the block is composed of a block header and a block body. The block header includes the hash value, version number, time date stamp, difficulty target, nonce value and Merkle root of the parent block. The block body stores a large amount of transaction information, while the blocks in the region store a large amount of transaction information [5].

Blockchain is decentralized and open. It also has the characteristics of DE value anchor, openness, consensus mechanism and non tampering, so as to ensure the modification of content, so as to ensure the correctness and security of transaction information and data. Firstly, the decentralization of blockchain makes each value anchor have the

right to share transaction information data among nodes and verify the correctness of the information; Verify the accuracy; Secondly, the openness of blockchain is characterized by that each open blockchain is characterized by that each node can join or exit at any time; Be able to enter and withdraw at any time; Thirdly, the characteristics of the blockchain consensus mechanism can enable the nodes without contact to reach agreement directly through the consensus mechanism; Consultation; Finally, the tamper proof feature of blockchain is that based on the modified feature, the hash algorithm is used to ensure that the transaction information will not be changed [5]. In addition, it cannot be modified according to the application scenario and openness. In addition, blockchains can be divided into public chains, alliance chains and private chains according to the application occasions and the scope of information disclosure.

**Table 1.** Blockchain structure



(2) Consensus mechanism

Decentralized value anchor is a core and key element of blockchain. Because the power of each node is the same, in order to ensure the mutual cooperation between each node, a set of algorithm is required, and the value is the same. Therefore, if you want to determine the mutual cooperation between different nodes, there must be a calculation. This set of algorithm is called calculation, which is called consensus mechanism. Mechanism [6]. The research mechanism of consensus mechanism originated relatively early. In 1982, it was also quite early. In 1982, Lamport, Shostak and pease proposed that Shostak and pease jointly studied how to reach the Byzantine general problem in the consensus agreement under the condition of Byzantine nodes, which promoted and promoted the development mechanism of consensus mechanism. The blockchain core technology consensus mechanism was first applied to bitcoin. Later, with the continuous development and further development and improvement of blockchain technology, the consensus mechanism became more and more mature. The common mechanism is also becoming more and more perfect. Common consensus mechanisms include proof of workload, proof of stake, proof of authorized equity, proof of authority, proof of combustion, proof of contribution, proof of existence, proof of data recoverability, proof of storage, Byzantine fault tolerance, etc. [7].



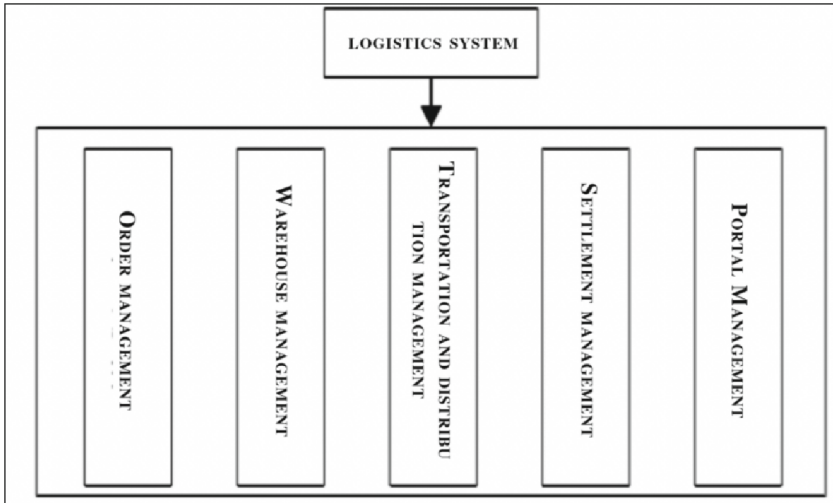
## 4 Logistics Transaction Process and Model

### (1) Logistics transaction process

At present, in addition to using the relevant modules of the logistics information management network system to customize the transportation plan and the warehousing module to simply manage the warehousing, the corresponding template is provided to formulate the logistics distribution scheme, and the warehousing template is used to keep and track the goods easily. Most operations still remain in manual management and custody or the preservation of paper documents. The logistics transaction process includes the order placing process, including customer information management, warehousing, inventory counting, transportation tracking, transaction distribution tracking, trade feedback, etc. Because the detailed sensitive information of the circulation logistics links such as transportation related distribution expenses, document review, transaction document verification, trade docking and goods integrity generated in this process is not unified, the whole system centralized processing cannot be realized [8]. Therefore, the whole logistics transaction process is not visible in the process, and it is unable to connect all customers in the supply chain to provide efficient logistics services, Nor can it be connected to every user in the enterprise to realize effective distribution business, nor can it meet the requirements of enterprise logistics decentralization and non repudiation.

### (2) Traditional logistics transaction mode

The traditional logistics distribution center transaction model has a certain model and has reached the corresponding degree of modernization. Through, we can use the operation specifications and rules of the logistics distribution center, the technology and means of intelligent identification of the Internet of things, and use the network station and database of the logistics distribution center to establish an information network platform, as shown in Fig. 1. Users can provide information on their network platform to realize data query and logistics transaction. Once such a centralized commodity transaction information data platform is attacked by hackers, it will steal and steal sensitive information in the database system, resulting in data leakage; Disclosure of customer information; The imperfect user review mechanism leads to the imperfect network review system, resulting in the proliferation of online registered users and customers, which can not eliminate false and prevent unfair competition such as virtual commodity trading; Improper activities of logistics transportation; In case of loss of goods, loss of goods and dishonest transactions during the transportation of goods, it is impossible to trace the origin of goods, collect trade, obtain the origin of goods, and obtain comprehensive, complete and systematic logistics transactions and effective goods trade information. The above problems and situations can not be properly solved under the premise of the traditional logistics goods centralization model. Therefore, it is proposed that cloud computing technology and blockchain technology should be applied to the logistics transaction model to achieve decentralization and apply to the goods trade model, so as to realize the purpose of DE value anchor, honest transaction, traceable trade and traceability [9].



**Fig. 1.** Traditional logistics management mode

### (3) Logistics blockchain Technology

Liang Bin understood the consensus mechanism and mechanism of blockchain through “mining with bitcoin”, described in detail the advantages and disadvantages of the four basic consensus mechanisms under the office chain, pointed out the characteristics of the mechanism, and put forward a consensus mechanism that does not require complete consensus under the alliance chain. Wang miaojuan pointed out the mechanism [10]. Wang miaojuan believes that due to the potential security risks caused by the information caused by the complete and full transparency of blockchain data, as well as the increase of data due to the increase of the number of statistical information, it is more and more difficult to save data analysis between node storage, as well as the information game caused by the competition caused by node competition. Finally, she envisages the issuance of tokens based on the design, development and adoption of blockchain, In order to realize the token system of information technology, in order to carry out the digital management of logistics express business and distribution service and the storage of package data and quantity information [11]. The scheme is not divorced from preservation in essence. Although this method can not fundamentally get rid of the traditional token mode, it is applied in logistics business, but it also provides ideas and Thoughts on the application of blockchain information technology in express industry [12]. Logistics chain technology integrates information flow, transportation and storage, and spans multiple logistics steps, many processes, hundreds of locations and destinations, which makes event tracking, historical data tracking, verification and verification, and unexpected response speed all face challenges. Moreover, it is difficult to investigate and study illegal activities in the logistics chain due to the lack of transparency. In fact, blockchain information technology is a safe and efficient multi-party and effective distributed ledger system with multiple information exchange and interaction modes. Its decentralization, tamperability,

forgery and high transparency make it transparent, making it an excellent tool to change the best chain of the logistics chain. Each participant in all chains of the logistics chain can be in a consensus network, which can provide simple and effective tracking tools for customers and auditors, and more simple and efficient tracing work [13].

## 5 Design of Intelligent Logistics Blockchain Consensus Algorithm Model

In order to realize resources, in order to meet the user needs for information flexibility, rapid adjustment between nodes, low scheduling, cost reduction and high robustness, as well as the application requirements of various transactions and high robustness, no trading entity needs to deploy large-scale computing clusters in local organizations. All large computing groups, because any consensus authentication process has verification, and the calculation is completed through blockchain or cloud network platform in the process. Therefore, the operation is carried out. Therefore, it further proposes to build a logistics distribution blockchain model based on cloud computing technology [14]. Pattern based. The logistics distribution blockchain model model based on cloud computing technology truly reflects the core requirements of decentralization. Multiple authentication, that is, multiple application nodes cooperate to verify any transaction activities in logistics activities. In the long run, considering the upstream suppliers and downstream sellers, the upstream providers and downstream distributors are dynamic, which can be guaranteed. In this way, we can improve the number of nodes, prevent the total amount of malicious nodes and avoid large-scale destruction and massive attacks of malicious nodes. Each authentication node is the beneficiary of logistics transaction service trade behavior, will actively abide by the main promoter, and will independently implement the consensus authentication mechanism. Therefore, it is consistent with the verification system. Therefore, the logistics service blockchain model of base cloud computing has a high transaction mode and strong trade stability and fault tolerance. As can be seen from Fig. 2, based on the second middle school, according to the logistics blockchain model of cloud computing, all transaction authentication behaviors are in the same mode, and all trade verification actions are performed on the Hadoop blockchain cloud platform. Firstly, the map function is used to dynamically allocate  $n$  nodes to each transaction information to different trading entities, then simulate its transaction authentication process and the process of trade verification, and then transfer the logistics transaction trade information through the hash encryption algorithm. Then, the Byzantine consensus ((pbft)) algorithm is used to complete the authentication, and the reduce function is used to realize the protocol processing in the reuse of the authentication process. The overall authentication process has, The whole verification process has high fault tolerance, which is convenient for internal transactions, which is also conducive to the cooperation of trade subjects, suppliers and sellers in the company to maintain the logistics transaction and distribution trade information system. Blockchain uses consensus mechanism and Hadoop and Hadoop distributed storage information technology to achieve logistics decentralization, which can realize the decentralization and tamperability of goods, customization, and users can trace the source of all goods, the source of the whole goods and the process of logistics and transportation [15].

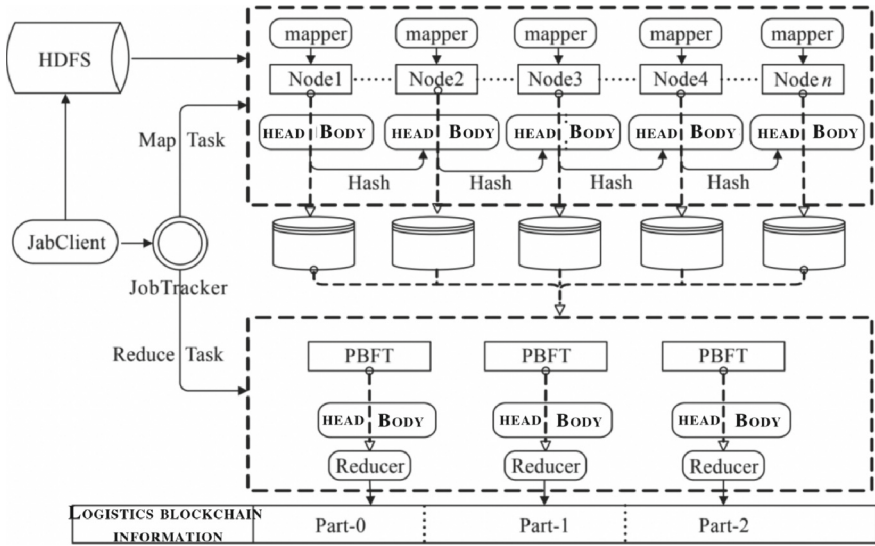


Fig. 2. Logistics blockchain model based on Cloud Computing

## 6 Conclusions

Using blockchain technology and consensus algorithm, this paper designs an intelligent consensus algorithm calculation research model on smart logistics, ensures the technical requirements of decentralization and non tampering, solves and solves the computational power problem of large-scale consensus operation, and provides a basis for a series of problems faced by the logistics industry, such as opaque transactions and non disclosure of information [16], The basic idea of blockchain solution is put forward. The simulation results show that the module has high performance in security, stability and throughput. However, the embodiment. However, due to the openness and transparency of the blockchain, it also means that the logistics personal information is published in the logistics network and user websites. While tracing the origin of the information flow, some personal information that is not suitable to be disclosed and some personal information that is not suitable to be disclosed will also be exposed. The next research will be exposed. Therefore, future research and development work will focus on the structural optimization of the construction mode of logistics decentralization model, And consider the application of digital signature technology to do a good job in order to keep users' privacy information safe and confidential.

## References

1. Tosh, D.K., Shetty, S., Liang, X., et al.: Security implications of blockchain cloud with analysis of block withholding attack. In: 2017 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID) 2017. IEEE (2017)
2. Zhu, Y., Chen, Z.: RealID: building a secure anonymous yet transparent immutable ID service. In: 2017 IEEE 3rd International Conference on Big Data Security on Cloud (Bigdatasecurity),

- IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS) 2017. IEEE (2017)
3. Xie, C., Sun, Y., Luo, H.: Secured data storage scheme based on block chain for agricultural products tracking. In: 2017 3rd International Conference on Big Data Computing and Communications (BIGCOM), 2017. IEEE (2017)
  4. Gramoli, V.: From blockchain consensus back to Byzantine consensus. *Future Generation Computer Systems* (2017)
  5. Toyoda, K., Mathiopoulou, P.T., Sasase, I., et al.: A novel blockchain-based product ownership management system (POMS) for anti-counterfeits in the post supply chain. *IEEE Access* **5**, 17465–17477 (2017)
  6. Petersen, O., Fredrik, J.: Blockchain technology in supply chain traceability systems. Lund University (2017)
  7. Dujak, D., Sajter, D.: Blockchain applications in logistics and supply chain. *Contemporary Logistics* (2017)
  8. Francisconi, M.: An explorative study on blockchain technology in application to port logistics. Delft University of Technology (2017)
  9. Wenjing, L., Jie, Z., Zhongming, L., et al.: Hadoop-based dynamic load balance scheduling algorithm of logistics inventory. In: 2016 12th International Conference on Computational Intelligence and Security (CIS) (2016)
  10. Nakamoto, S.: Bitcoin: a peer-to-peer electronic cash system. Satoshi Nakamoto Institute (2008)
  11. Lamport, L.: The Byzantine generals problem. *ACM Trans. Program. Lang. Syst.* **4**(3), 382–401 (2002)
  12. Kosba, A., Miller, A., Shi, E., et al.: Hawk: the blockchain model of cryptography and privacy-preserving smart contracts. In: *Proceedings of 2016 Symposium on Security and Privacy*, pp. 839–858 (2016)
  13. Grinberg, R.: Bitcoin: An Innovative Alternative Digital Currency. Social Science Electronic Publishing (2011)
  14. Bowers, K.D., Juels, A., Oprea, A.: Proofs of retrievability: theory and implementation. In: *Proceedings of the 2009 ACM Workshop on Cloud Computing Security (ACM)*, pp. 43–54 (2009)
  15. Dahlin, M., Clement, A., Wong, E.: Zyzzyva: speculative Byzantine fault tolerance. *Ramarishna Kotla Microsoft Res.* **4**(7), 29–35 (2009)
  16. Castro, M.: Practical byzantine fault tolerance and proactive recovery. *ACM Trans. Comput. Syst. (TOCS)* **20**(4), 398–461 (2002)



# Improved ID3 Algorithm in Sports Tourism Service System

Quanan Gui<sup>(✉)</sup>

The International School of Equestrian, The School of Physical Education,  
Wuhan Business University, Wuhan 430056, China  
542740013@qq.com

**Abstract.** In recent years, my country has made every effort to promote the investment and development of the sports tourism industry; adhere to market leadership, government support, standardize the development of sports tourism activities, and accelerate the formation of a sports tourism industry system and product system with a reasonable system structure and complete functions. The purpose of this paper is to study the application of the improved ID3 algorithm in the sports tourism service system. The concept of sports tourism service system is studied and divided into four systems for exposition. The relevant knowledge in the field of data mining is introduced, the idea of ID3 algorithm and the criteria for ID3 algorithm to select splitting attributes are given, and then the research problem is introduced; Based on the concept of misclassification ratio, a first pruning strategy based on misclassification ratio is proposed, and the BAID3 algorithm is applied to the weather analysis of the service system of sports tourism service system. A decision tree about whether it is suitable for sports tourism is constructed, and compared with the traditional ID3 algorithm and the C4.5 algorithm, the BAID3 algorithm is better than the ID3 algorithm and the C4.5 algorithm in the number of internal nodes and the number of leaf nodes.

**Keywords:** Improving ID3 Algorithm · Sports Tourism · Tourism Service · Service System

## 1 Introduction

After the material needs of the masses are met, the demand for services and products at the spiritual level begins to increase. At the same time, leisure time is also increasing. The tourism industry has developed rapidly around the world. Tourists no longer stop to participate in traditional single Sightseeing tourism activities, but to the pursuit of experience, recreation and fitness leisure sports activities [1, 2]. The integration of sports and tourism has given birth to sports tourism, a new type of tourism that integrates leisure, fitness and entertainment. Therefore, sports tourism has become a new development trend of modern tourism and is favored by tourists from all over the world [3, 4].

With the increasing development of sports + tourism, policy documents related to the development of sports tourism have been issued one after another, and the development

of sports tourism in various provinces and cities is in full swing [5]. As a sporting activity with a long tradition, existing products and future prospects, golf is a first-class tourism resource, for which Daries N developed an integrated model for analyzing golf course websites: an analysis based on web content with four dimensions (i.e. information, communication, e-commerce and additional functions), and the stages of extending the model [6]. Putra F analyzes potential interest in PERSIS Solo sports tourism in Suragada following Manahan Solo Stadium renovation plan. This type of research is descriptive qualitative research. The analysis method uses SWOT, and the results show that PERSIS Solo is in the first quadrant, so club management can adopt policies that use the power of the internal environment to maximize the existing market opportunities [7]. Therefore, it is necessary to study the data analysis in the sports tourism service system [8].

This paper introduces the basic idea, basic process and production indicators of the decision tree classification algorithm, and introduces several common decision tree classification algorithms. Then it introduces the principle, description, advantages and disadvantages of ID3 algorithm and the key content of the sports tourism service system. The specific steps of improving the algorithm, the main data structure of the algorithm application and the flow of the algorithm are also given. Finally, the improved ID3 algorithm is implemented on the development platform whose processing environment is Eclipse. The improved ID3 algorithm is applied to the analysis example of sports tourism service system. This example implements the comparison between the improved ID3 algorithm and the original ID3 algorithm.

## 2 Research on Application of Improved ID3 Algorithm in Sports Tourism Service System

### 2.1 Data Mining

#### (1) Decision tree

A decision tree is a tree structure similar to a flowchart. It takes a retrospective top-down approach. Each node in a decision tree compares attribute values. The downstream branches of a node are determined by the results of determining various eigenvalues [9, 10]. Leaf nodes reflect the conclusions of the decision tree, and by analogy, a decision tree is such an iterative tree structure. Among them, each node without sheet represents the input attribute of the dataset, the corresponding attribute value is defined as the attribute value, and the leaf node represents the final output attribute value [11, 12].

The basic algorithm of decision tree is the greedy algorithm. The existing commonly used decision tree learning algorithms are based on this algorithm. The algorithm uses retrospective search and top-down scrolling. This is the most basic bootstrap algorithm. Among decision tree-based classification algorithms, ID3 is the most classic bootstrapping algorithm [13, 14]. CART and C4.5 are extensions of ID3. The main advantage of a decision tree is that it saves a lot of data preprocessing tasks because its structure is simple and easy to understand and it is very good at dealing with non-numeric data. The model has high efficiency and fast sorting speed, and is especially suitable for large-scale data processing, does not

require knowledge other than training data, and has high accuracy. It is currently the leading classification technology and has been successfully applied to data analysis in multiple industries [15, 16].

## (2) ID3 algorithm

The core of the ID3 algorithm is: if the attributes at all levels of the decision tree are selected, the information gain is used as the attribute selection criterion, so if each node without leaves is checked, the maximum category information of the test record is obtained. The method is as follows: discover all attributes, select the attribute with the highest information gain to create a decision tree node, create branches with different attribute values, then call the method retroactively on a subset of each branch, creating a branch for node determination. Trees up to all subsets contain only the same class of data, resulting in decision trees that can be used to classify new samples [17].

From the basic principle of the ID3 algorithm, it can be seen that the ID3 algorithm uses the information entropy value of each attribute to determine the separation attributes in the data set, and the selection tends to favor attributes with more values. In response to this problem, many methods have been proposed, such as: profit rate method, Gini index method, G-statistics method and so on. From the principles of many existing improved algorithms and the basic types of information gain, it can be seen that the size of the information gain determines the information entropy, and the information entropy is used to reflect the uncertainty of each attribute in the entire dataset [18].

## 2.2 Sports Tourism Service System

Sports tourism has become a rising star in my country's tourism industry, and it is a new form of sports that combines sports and tourism. There are many definitions of sports tourism, and there is a lot of controversy about the specific content and form of distinguishing tourism and sports tourism. But no matter how the definition is studied and analyzed, the basic definition of sports tourism contains an important key point, that is, the sum of the social relations associated with the project of tourists who participate in or watch various sports content. The sports tourism service system should include management system, supply system, product system and service system.

The management system of the sports tourism service system is the institution and department that manages the ice and snow sports tourism service activities.

The supply system of the sports tourism service system refers to the collection of enterprises, organizations and departments that interact with tourists with the goal of improving tourist satisfaction, including sports and leisure business venues, tourism companies, public welfare clubs and for-profit sports companies, etc.

The product system of the sports tourism service system refers to all sports activities that are finally provided to tourists as products through resource development, including sports event tourism products, sports leisure tourism products, festival tourism products and national traditional sports tourism products.

The service system of the sports tourism service system refers to the service content and specific measures and behaviors provided by the service provider to improve the



pleasure level of tourists in the process of tourism experience, including supporting equipment and facilities, auxiliary items and services.

### 3 Investigation and Research on Application of Improved ID3 Algorithm in Sports Tourism Service System

#### 3.1 BAID3 Algorithm

The BAID3 algorithm uses the improved information gain as its attribute selection criteria. Based on the information gain adopted by the original ID3 algorithm, the corresponding relationship between attributes and class tags, the distribution of attribute values, and the corresponding relationship between attribute values and class tags are calculated. Taking this into account, that is, using the influence factor of the attribute and the influence factor of the attribute value to modify the original calculation of the information gain.

After adding the influence factor of the attribute, the improved expected information amount  $C\_Info_A(Set)$  required to classify any tuple in the training dataset  $Set$  according to the attribute  $A$  is defined as:

$$\begin{aligned}
 C\_Info_A(Set) &= IFAs(A, Set) \times Info(Set) \\
 &= \frac{Diff(Set)}{Diff(S)} \times \sum_{i=1}^m p_i \log_2(p_i)
 \end{aligned}
 \tag{1}$$

Among them,  $p_i$  represents the non-zero probability that any tuple in the training dataset  $Set$  belongs to the class  $C_i$ . After adding the influence factor of the attribute value, the dataset  $Set$  needs to be split according to the discrete attribute  $A$  to achieve an accurate classification. The total improved information amount  $W\_Info_A(Set)$  is defined as:

$$W\_Info_A(Set) = \sum_{j=1}^v \left( \sum_{i=1}^t \left( \frac{IFAVs(a_j, C_i)}{\times (-p_i \log 2p_i)} \right) \times \frac{|Set_j|}{|Set|} \right)
 \tag{2}$$

Among them,  $|Set_j|/|Set|$  represents the weight of the  $j$ th subset,  $p_i$  represents the non-zero probability that any tuple in the training data set  $Set_j$  belongs to class  $C_i$ ,  $p_i = |C_i, Set_j|/|Set_j|$ ,  $S$  is the original sample set. Like the ID3 algorithm, the smaller the amount of information required, the higher the purity of the division.

#### 3.2 Example Application of Sports Tourism Service Algorithm

In order to illustrate the application method of the BAID3 algorithm in the sports tourism service system, we use the following example to illustrate it. In order to improve the pleasure level of tourists in the process of sports tourism experience, the service system of the sports tourism service system in this paper provides weather analysis services.

Select the randomly selected training sample set about “climate” in February 2022. In the sample set, each attribute is discrete, among which “weather”, “temperature”, “humidity” and “wind” are general attributes, “Sports” is a class label attribute, which has two different attribute values {Suitable for sports tourism, not suitable for sports tourism}; therefore, there are two different classes here, let class C1 represent “Sports tourism”, and class C2 represent “Not suitable for sports tourism”. There are 9 tuples belonging to class “fit” and 5 tuples belonging to class “not fit”.

Experimental running environment: Intel(R) Core(TM) 2 Quad CPU Q8400 @ 2.66 GHz 2.67 GHz, 2.00 GB RAM, Windows 7, Eclipse Java EE IDE for Web Developers.

## 4 Analysis and Research on Application of Improved ID3 Algorithm in Sports Tourism Service System

### 4.1 Algorithm Application Process

Among the four attributes, the attribute “weather” has the highest information gain, so “weather” is selected as the division attribute. Label the node N as the attribute “weather” and grow branches with each attribute value of “weather”, the tuples in the training dataset Set are divided into three subsets as shown in Fig. 1. We notice that the tuples that fall into the “weather = cloudy” branch all belong to the same class label “suitable”, so a leaf node should be created at the end of this branch and label this leaf node as “suitable”. Finally, the decision tree structure established by the BAID3 algorithm is shown in Fig. 1, which is used to predict whether a certain day is suitable for sports tourism, each internal node represents a test attribute, and each leaf node represents a class label (“Suitable for sports tourism”) or “Not suitable for sports tourism”).

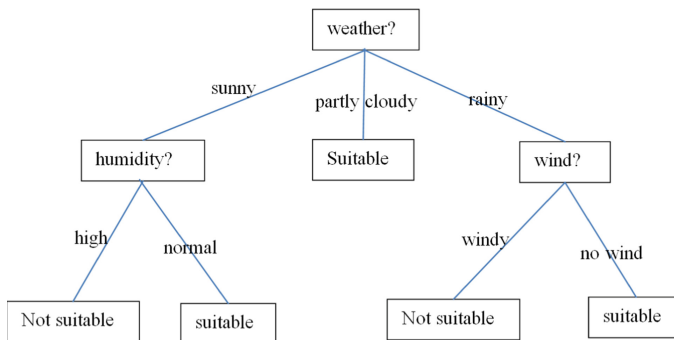


Fig. 1. The decision tree generated by the BAID3 algorithm based on the dataset

### 4.2 Algorithm Comparison

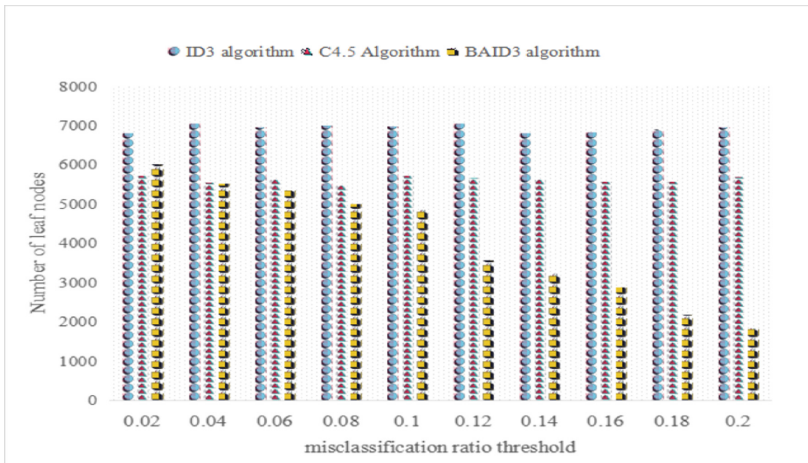
The decision tree generated by the BAID3 algorithm under different misclassification ratio thresholds  $\delta$  is superior to the ID3 algorithm and the C4.5 algorithm in both the

number of internal nodes and the number of leaf nodes, as shown in Fig. 2. Therefore, the decision tree structure constructed by the BAID3 algorithm is simpler and more reasonable and easier to analyze and understand, that is, the decision tree structure generated by the BAID3 algorithm is more reasonable and accurate; and the C4.5 algorithm, as an improved algorithm of ID3, the structure of the decision tree constructed is also excellent. Based on the ID3 algorithm, as shown in Table 1.

When the misclassification ratio threshold  $\delta$  is 0 (at this time, the BAID3 algorithm has not undergone any pre-pruning operation in the entire decision tree generation process), the structure of the decision tree generated by the BAID3 algorithm is better than the ID3 algorithm, because the BAID3 algorithm has The improved information gain is used to select the attributes for dividing the data set, which overcomes the problem of multi-value bias in the traditional ID3 algorithm to a certain extent, so the generated decision tree is more reasonable and closer to the ideal. Decision tree structure. When the misclassification ratio threshold  $\delta$  is not 0 (at this time, the BAID3 algorithm will take a certain first pruning operation in the decision tree generation process), the reason why the decision tree structure generated by the BAID3 algorithm is better than the ID3 algorithm and C4.5 The algorithm is not only because it overcomes the problem of multi-value bias, but also because of the adoption of the first pruning strategy, which removes the decision tree branches caused by noise or outliers, and avoids the occurrence of overfitting. This makes the decision tree structure more reasonable.

**Table 1.** Algorithm Comparison Results

Misclassification ratio threshold $\delta$	ID3 algorithm	C4.5 Algorithm	BAID3 algorithm
0.02	6820	5721	6032
0.04	7052	5551	5529
0.06	6950	5632	5367
0.08	7012	5488	5018
0.10	6990	5716	4872
0.12	7054	5688	3567
0.14	6811	5628	3261
0.16	6850	5569	2888
0.18	6915	5578	2197
0.20	6973	5690	1847



**Fig. 2.** The number of leaf nodes of the decision tree generated under different misclassification ratio thresholds  $\delta$

## 5 Conclusions

The sports industry is becoming more and more popular, and sports tourism, as a part of it, plays a vital role. In the context of the continuous improvement of China's economic development level, people's living needs have shifted from basic physical activities to outdoor leisure activities and tourism activities far away from the place of residence, and the pursuit of a spiritual level centered on health, family affection and life experience is increasingly escalating. Therefore, sports tourism has become a research hotspot. Starting from a large amount of accumulated weather raw data, this paper mainly uses the decision tree classification algorithm of data classification, combined with the improved decision tree classification ID3 algorithm to analyze the specific trends of customer groups and formulate relevant strategies. It focuses on the analysis of sports tourism travel plans of various types of tourists, provides scientific knowledge support for sports tourism service management and sports tourism marketing strategies, and has important practical significance in the application of sports tourism informatization.

**Acknowledgements.** This work was supported by the Science Research Project of Wuhan Business University 'Study on sustainable development of Jingchu folk sports tourism under the background of sports industry', (No.2021KY020).

## References

1. Praisra, H., Endyana, C., Khan, A., et al.: Potensi wisata olahraga air sungai Cihérang: Kayak X; Water sport tourism potential Cihérang river: Kayaking X. *Altius Jurnal Ilmu Olahraga dan Kesehatan* **10**(1), 19–30 (2021)
2. Weed, M.: Maximising the use of existing knowledge and evidence: showcasing systematic reviews in sports tourism. *J. Sport Tour.* **25**(1), 1–2 (2021)

3. Putra, M.J., Hermanto, B., Sukoco, I., et al.: The effect of tourist objective photos of tourist festival on sport tourism (Asian Games) In Palembang, Indonesia. *Halaman Olahraga Nusantara (Jurnal Ilmu Keolahragaan)* **4**(2), 350–368 (2021)
4. Cahyadi, H.S., Suherlan, H., Dewantara, M.A.: Cost and benefit analysis of sports tourism in Downhill Trek, Cikole Lembang, Bandung Barat District, West Java Province. *Jurnal Kepariwisata Destinasi Hospitalitas dan Perjalanan* **5**(1), 1–9 (2021)
5. Terzi, A., Demirovi, D., Petrevska, B., et al.: Active sport tourism in Europe: applying market segmentation model based on human values. *J. Hosp. Tour. Res.* **45**(7), 1214–1236 (2021)
6. Daries, N., Cristobal-Fransi, E., Ferrer-Rosell, B.: implementation of website marketing strategies in sports tourism analysis of the on line presence and e commerce of golf courses. *J. Theor. Appl. Electron. Commer. Res.* **16**(3), 542–561 (2021)
7. Putra, F., Aqmala, D., Praswati, A.N.: The impact of gelora manahan solo renovation on persis solo potential sport tourism. *Jurnal Penelitian Ekonomi dan Bisnis* **6**(1), 45–57 (2021)
8. Jaelani, E., Erdinaya, L.K., Dr. Rohanda, et al.: The religious meaning of equestrian and archery sport tourism: a phenomenological analysis. *Geoj. Tour. Geosites* **28**(1), 246–256 (2020)
9. Park, S.H., Cho, K.B., Roh, D.Y.: The relationship among sports tourism SNS information characteristics, satisfaction and intention to reuse. *Korean J. Sports Sci.* **29**(6), 447–459 (2020)
10. Kim, M., Park, S.: A study on priority assessment of policy factors for revitalization of sport tourism. *J. Cult. Exch.* **9**(4), 29–45 (2020)
11. Hasibi, S., Shojaei, V.: Strategic analysis of sports tourism marketing mix in Mazandaran with 7P's approach. *J. Appl. Res. Geograph. Sci.* **20**(57), 169–186 (2020)
12. Traverso, L.D., Oliveira, A., Maciel, J., et al.: Sports tourism and the 9th international marathon of Punta Del Este, Uruguay. *Revista Rosa dos Ventos - Turismo e Hospitalidade* **12**(3), 616–634 (2020)
13. Omorczyk, A.: A tourist in pursuit of the world of sport review of the book by Ewa Malchrowicz-Moko entitled sport tourism, academy of physical education in Poznań, Poznań 2018. *Acta Universitatis Lodziensis Folia Sociologica* **75**(208–600X), 133–139 (2020)
14. Tayebisani, S., Rouhani, A.: Shahrood sport tourism planning using SWOT technique with emphasis on natural attractions. *Int. J. Cult. Tour. Hosp. Res.* **7**(1.2), 65–82 (2020)
15. Choi, Y.J., Kim, H.Y., Hur, S.E., et al.: A study on exploring of moving sports tourism: case of bicycle tourism. *J. Sports Appl. Sci.* **4**(1), 14–27 (2020)
16. Harahap, Z., Kartika, T.: Community development in sports tourism development as economic driver inclusive in south sumatera. *J. Indones. Tour. Hosp. Recreat.* **3**(2), 197–206 (2020)
17. González-García, R.J., Escamilla-Fajardo, P., López-Carril, S., et al.: Residents' perceptions of sports tourism: impacts, quality of life and support for the industry. *Cuadernos de Psicología del Deporte* **20**(2), 174–188 (2020)
18. Petrovi, M.D.: Editorial on sports tourism issue. *J. Hosp. Tour. Res.* **45**(1), 1–3 (2020)



# Design of Public Sports Service System in the Era of 5G + Artificial Intelligence

Xinying Liu<sup>(✉)</sup>

The School of Physical Education, Wuhan Business University, Wuhan, Hubei, China  
15239059@qq.com

**Abstract.** With the deepening of digital construction in the field of public sports, the volume of public sports resources has also become very large. Some public cultural and sports service units have built digital service network sites, using the Internet to allow users to obtain resources and share in a timely manner. The purpose of this paper is to study the design of public sports service system based on the era of 5G + artificial intelligence. Combined with the development level at home and abroad, give a suitable solution for the current situation in China, and design and implement a public cultural and sports resource service system; sort out the business requirements of the system to ensure that the system can solve the research problems; use the sports industry public service cloud platform As the carrier, the instant messaging system of the platform can provide sports business consulting services and platform consulting services for the platform audience, and the implementation of the system data statistics module is analyzed in detail, combined with the actual application scenarios, so that the system can meet the overall design and expectations.

**Keywords:** Artificial Intelligence · 5G Technology · Public Sports · Service System

## 1 Introduction

Urban community public sports service is the focus and part of national public service construction, and the construction of healthy urban community cannot be separated from the rational distribution of public sports service resources [1, 2]. Achieving the balanced development of the supply of public sports services in urban communities is an important way to achieve social equity and a social redistribution of public sports services [3, 4]. Integrating health into every link of urban construction fully reflects the importance and necessity of urban community public sports services in the process of promoting the “Healthy China” strategy. Improving the supply of public sports services in urban communities is an effective means to improve the health level of community residents and promote the construction of “Healthy China” [5].

Public sports services are a hot topic at the moment, and Oh H’s research focuses on extracting the factors that influence user satisfaction who regularly use large public sports centers, and using them in the planning and operation of future facilities. In terms

of theoretical background, the definitions, samples and facility satisfaction factors of large public sports complexes are reviewed. Through the survey, the factors affecting user satisfaction are analyzed in the order of program level and type, sports equipment, convenience of visiting facilities and program price [6]. Elebi E's study of 176 sports fans found that public relations strategies (controlling reciprocity, trust, satisfaction) were precursors to fan loyalty. In addition, public relations perceptions were found to be a determinant factor in supporting sports club behavior. Sports fans define their relationship with sports organizations as a public relationship based on one-sided support rather than an exchange based on mutual benefit. As a result of the study, several recommendations were made for sports clubs to improve the quality of their relationships with supporter groups [7]. Therefore, it is feasible to build a public sports service system under the background of 5G + artificial intelligence era.

The innovation of this paper: In terms of research content, combined with the current national vigorous promotion of the development of intelligent public sports services and the implementation of the "5G + artificial intelligence" action plan, the community is the starting point to study the development of public sports service systems. For other public sports The subject research of the service is relatively new, and it is also the expansion of the content. Using basic analysis to understand the public's needs and reasons for intelligent public sports service equipment, platforms, and content, and then analyze the influencing factors of demand, and innovate the research perspective of intelligent public sports service system.

## **2 Research on the Design of Public Sports Service System in the Era of 5G + Artificial Intelligence**

### **2.1 5G Technology**

The high speed and convenience of 5G technology make the application of digital sports more extensive. Digital sports is the conversion of traditional sports into the application of digital technology under the promotion of digital technology to meet the user's purpose of physical exercise [8, 9]. The mobile public sports service system itself is a digital sports activity. The development of the mobile public sports service system should seize the development opportunity of 5G, change the way of thinking, re-cognize mobile e-sports, and create a new market orientation. To meet the needs of sports, develop traditional sports services [10]. The breakthrough of 5G technology has promoted the widespread popularity of instant messaging, which enables users to experience public sports services without leaving their homes, and is no longer limited by the limitations of traditional venues.

### **2.2 Artificial Intelligence**

With the development of science and technology, artificial intelligence technology is becoming more and more mature, and it is more and more applied to traditional services, enabling it to actively and automatically obtain information, conduct analysis, and match needs. Therefore, while applying artificial intelligence, it is necessary to regulate artificial

intelligence and other related technologies [11, 12]. The problems brought about by the development of artificial intelligence are mainly reflected in the legal and ethical aspects, which are also the major disputes in the current development of artificial intelligence. When developing artificial intelligence systems, it is necessary to pay attention to the security of the system, to improve the security of the system at the professional and technical level, and to strictly test the artificial intelligence technology and system, and to improve the security of the artificial intelligence system under various guarantees [13, 14].

### **2.3 Sports Public Service**

There are many classifications of public services, and according to different standards, there are different classifications. According to the characteristics of public services, public services can be divided into pure public services and quasi-public services (or mixed public services); according to the functions of public services, public services can be divided into maintenance public services, economic public services and social public services. Services [15, 16].

The academic circle uses two terms for this proposition: “Sports Public Service” and “Public Sports Service”. The objects denoted by the two concepts are identical, but there has been debate as to which term is the most normative [17]. Whether it is “sports public service” or “public sports service”, the conceptual cognitions expressed by the two tend to be the same, that is, they both have public attributes, are dominated by government departments, and participate in the society and individuals. The related sports products and services are a combination of sports and public services [18].

### **2.4 Analysis of System Business Requirements**

The system services of the sports venue service platform are mainly extracted from actual consumption scenarios, and these services are established based on user needs. The service platform system provides a new type of consumption service based on the O2O model for both sports players and venue operators. Athletes do not need to blindly search for suitable sports, and venue operators do not have to invest huge costs. In terms of marketing, sports venue operators add their own venue information on the platform, and add the sports services that venues can provide in their own venues. Project, and then the venue information and sports project information will be submitted to the system administrator for approval, and after completion, the information will be displayed on the sports venue service platform. Sports participants only need to open the sports venue service platform to browse all the surrounding sports venue resources in one stop, and then they can view the detailed information and sports items of a venue on the platform, and can also view the detailed introduction of each sports item. After understanding the information of these sports resources, sports players can go offline to exercise or consume.



## 2.5 Cloud Computing Technology

The cloud computing model is actually to store a large amount of application software and data in a cloud with super storage and computing functions built through virtualization technology for users who access the cloud. Cloud computing is often used in data computing and analysis. The distributed computing method divides the program that needs to be calculated into thousands of subprograms, and hands these subprograms to each IP network unit for processing. Finally, each processor aggregates the processing results into the cloud platform to form the results. This computing mode can greatly improve computing efficiency and reduce computing time, enabling users to enjoy super network computer-like services with a computing power of more than 10 trillion operations per second, and the storage and computing power of cloud computing are based on demand Scale by adding remote data center servers.

## 3 Investigation and Research on the Design of Public Sports Service System in the Era of 5G + Artificial Intelligence

### 3.1 Architecture of Public Sports Service System

The instant messaging system in the public sports service cloud platform has a wide range of uses. For example, when inquirers make inquiries to the information center, they need to use the instant messaging system; when consumers purchase sports services, they need to use the instant messaging system to learn more about the service information; investors need to use the instant messaging system. An instant messaging system is required for project information. The instant messaging system based on the C/S architecture is the basis for the cloud application in the cloud platform to play its functions. The operation mechanism of the instant messaging system service of the public sports service cloud platform is as follows: users access the cloud platform, choose to access the corresponding sub-platform according to their own needs, and enter the information consultation questions in the sub-platform. After the cloud platform system performs simple information processing, the user's consultation information is sent to the corresponding service window through the platform instant messaging system, and the user can communicate with the corresponding staff. After the exchange, the user will give feedback and evaluation of the service according to the service situation, and the cloud platform system will analyze the feedback information and store it in the cloud database for reference by relevant subjects.

### 3.2 System Data Statistical Model

The conversion rate of venue facilities booking is the ratio of the actual number of registrations for the venue facilities to the total number of independent users visiting the venue facilities introduction page of the mobile client within a period of time. The calculation formula is formula (1).  $k$  represents the time period. If the calculation period is one month, then  $k = 30$ , which is the conversion rate of venue facilities within 30 days.

$UV_k$  represents the number of unique users visited by the mobile client on the  $k$ th day, and  $DU_k$  represents the number of subscribers on the  $k$ th day.

$$ConversionRate(k) = \frac{\sum_{n=1}^k DR_k}{\sum_{n=1}^k UV_k} \bullet 100\% \quad (1)$$

The formula for calculating the utilization rate of venue facilities is formula (2).  $UsageRate(k)$  represents the utilization rate of venue facilities, where  $k$  represents the time period. If the utilization rate of a venue is calculated for one month,  $k = 30$ .  $DU_k$  represents the number of sessions actually used on the  $k$ th day, and  $DS_k$  represents the total number of sessions opened on the  $k$ th day.

$$UsageRate(k) = \frac{\sum_{n=1}^k DU_k}{\sum_{n=1}^k DS_k} \bullet 100\% \quad (2)$$

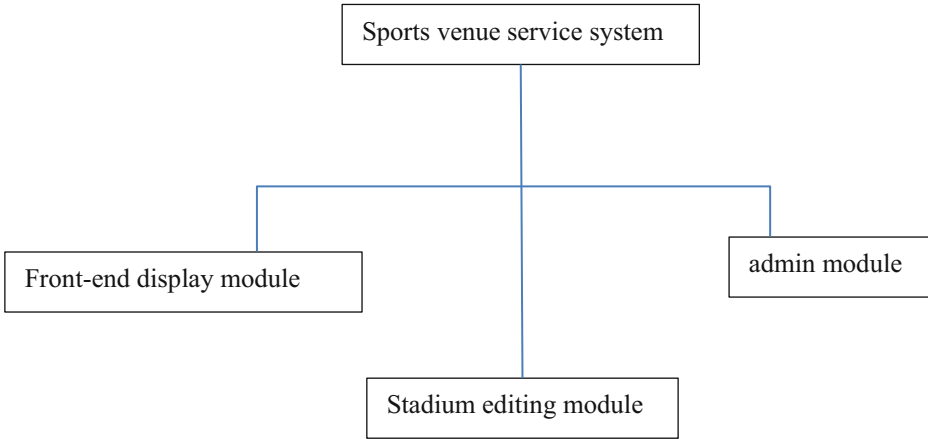
## 4 Analysis and Research on the Design of Public Sports Service System in the Era of 5G + Artificial Intelligence

### 4.1 System Function Module Design

The sports venue service system is divided into three modules, namely the front-end display module, the sports venue editing module and the administrator module, as shown in Fig. 1. The front-end display module includes the display of the home page of the service platform, and the information required by the home page mainly includes the advertisement information of the recommended position, the list of recommended venues, the classification of commonly used venues, and the statistics of user behavior data. The venue display function fulfills the requirements of displaying surrounding venues in a list, displaying detailed information of each venue, and searching for corresponding venues. The sports item display function fulfills the requirements of displaying various sports items in a list, displaying the detailed information of a sport item and searching for a specific sport item. The stadium editing module is mainly for stadium operators, and performs account information management, stadium information editing, sports item editing and authority verification operations. The administrator module has the highest authority of the platform, the user management adds and bans accounts, the sports venue management audits the venue information, and the authority verification ensures the security of important operations.

### 4.2 Implementation of System Data Statistics Module

The management background data statistics page has statistics on the usage data of client users, such as the number of visits, readings, usage of venue facilities, etc., and

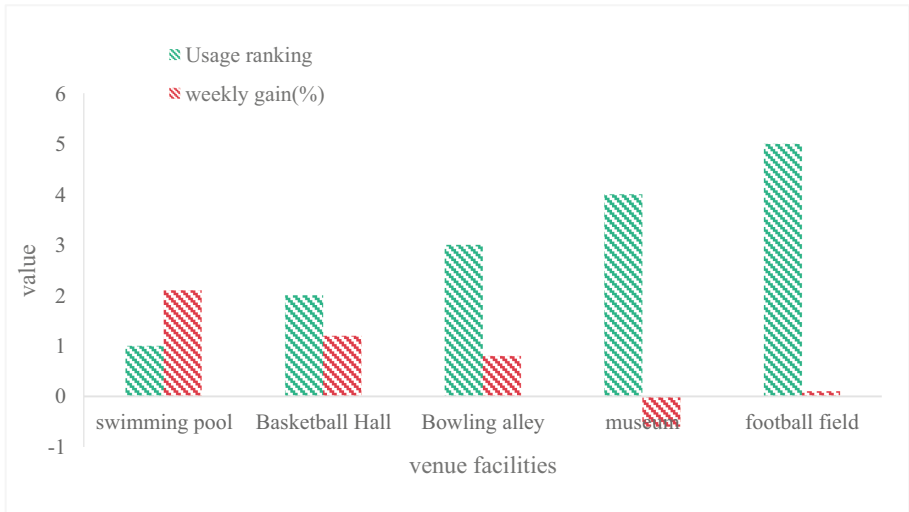


**Fig. 1.** Three modules of the system

the user’s behavior data is counted and displayed in the form of graphs, which can help relevant agencies and units analyze public affairs. The use of cultural and sports resources services, better allocation of resources, improve the effective supply of public cultural and sports resources and services, rational allocation of common cultural and sports resources, and data-driven service upgrades. The data returned by the client is calculated by the server and stored in the database, and then the output front-end is displayed by JS\_Charts. The statistical data mainly include the number of client visits, the reading volume of news information, the reading volume of event information, the statistics of event participation data, the reading volume of venue facilities information, the number of venue facilities reserved, the ranking of venue facilities, and the utilization rate of venue facilities as shown in the Fig. 2, the age distribution of event participants, etc., according to this information, the conversion rate of the event activities, type preference analysis, age distribution, etc., the usage distribution and usage rate of venue facilities, etc. are shown in Table 1.

**Table 1.** Age distribution

Age distribution	usage(%)	user count
10–20	31	1996
21–30	38	2447
31–40	20	1288
other	11	708



**Fig. 2.** Venue Facilities Utilization Rate

## 5 Conclusions

At present, with the strong support of national policies and the application of 5G + artificial intelligence, many problems faced by the supply of public sports services in urban communities will be effectively solved. Based on the theoretical perspective of 5G + artificial intelligence, this paper conducts an in-depth study on the supply of public sports services in urban communities, in order to explore the internal connection of the supply of public sports services in urban communities and the existing problems at this stage, and try to propose the network governance of the public sports service system. Action. It is hoped that it will be beneficial to the research work on urban community public service governance in the new era, and make contributions and efforts to strengthen and innovate the construction of the social system and promote the modernization of the national sports service system and governance capacity in the new era.

**Acknowledgements.** This work was supported by Humanities and Social Science Research Project of Education Department of Hubei Province in 2021, Research on the Construction of Youth Sports Public Service System in Hubei Province with the Background of Healthy China, No.21G108.

## References

1. Pagani, M.M., Shimoda, E., Matta, L., et al.: A Produo Do Conhecimento Em Poltica Pblica De Esporte No Brasil E No Mundo/Knowledge production in public sport policy in Brazil and the world. *Braz. J. Dev.* **7**(2), 11744–11756 (2021)
2. Kim, I.G., Park, S.T., Kim, M., et al.: Improving the operation of public sports facilities supported by the National Sports Promotion Fund: Focused on the National Sports Center. *Korean J. Sport Manag.* **25**(6), 18–33 (2020)

3. Rozanski, S.L., Kwasowski, W., Castejon, J.P., et al.: Heavy metal content and mobility in urban soils of public playgrounds and sport facility areas, Poland. *Chemosphere* **212**(Dec.), 456–466 (2018)
4. Kalman, K.: Health awareness and study performance in public educational type sports schools. *Central Eur. J. Educ. Res.* **3**(1), 134–136 (2021)
5. Prikhodko, V., Tomenko, O., Matrosova, S., et al.: strategic issues of public governance in sports development in Ukraine. *Sport Sci. Hum. Health* **5**(1), 73–83 (2021)
6. Oh, H., Kim, Y.K.: A study on analysis of factors affecting user satisfaction of large-scale public sports complex. *J. Eng. Appl. Sci.* **13**(18), 7487–7491 (2018)
7. Elebi, E.: The effect of organization-public relations on sports fan loyalty. *J. Hum. Sci.* **18**(1), 94–103 (2021)
8. Silva, D.S., Silvestre, B.M., Cristina, S., et al.: Assessing the Timemania Lottery as a sports public policy. *J. Phys. Educ.* **31**(1), 1–10 (2020)
9. Kleinmann, C.M.: Do we really want sports public relations to return to normal? *Int. J. Sport Commun.* **13**(3), 1–7 (2020)
10. Moriconi, M., Cima, C.D.: To report, or not to report? From code of silence suppositions within sport to public secrecy realities. *Crime Law Soc. Chang.* **74**(1), 55–76 (2020)
11. Hull, K., Wilson, B.: Journalism or public relations? Coverage of sports teams in high school journalism programs. *J. Sports Med.* **15**(1), 29–50 (2020)
12. Staack, M., Haut, J.: Sport im öffentlichen RaumSport in public space. *Sport und Gesellschaft* **16**(3), 370–374 (2019)
13. Deemua, G.A.: Managing male and female athletes for increased sports performance in public universities in rivers state. *Adv. Soc. Sci. Res. J.* **7**(6), 633–640 (2020)
14. Kim, B.G.: The effects of sport brand corporate public relationship on identification and consumers' behavior intention. *J. Korean Assoc. Phys. Educ. Sport Girls Women* **34**(3), 35–48 (2020)
15. Park, E.S., Lee, J.E., Lee, T.H.: A research on concernment, participation, and perceptual change concerning sports of the public, according to sports star image. *Korean J. Sports Sci.* **27**(4), 307–320 (2018)
16. Cooper, A.F.U.S.: public diplomacy and sports stars: mobilizing African-American athletes as goodwill ambassadors from the cold war to an uncertain future. *Place Brand. Public Dipl.* **15**(3), 165–172 (2019)
17. Dubinsky, Y., Dzikus, L.: Analyzing Israel's use of sports for public diplomacy through international relations perspectives. *Int. J. Civic* **17**(1), 15–25 (2019)
18. Zsari, A.: Sport diplomacy as public diplomacy element. *Int. J. Sci. Cult. Sport* **6**(28), 339–349 (2018)



# Automatic Design System of Product Shape Based on Intelligent Algorithm

Ren Li<sup>(✉)</sup>

Academy of Art and Design, Shaoyang University, Shaoyang, Hunan, China  
1r465823778@163.com

**Abstract.** The appearance design of the product affects the overall design of the product. Without a good product appearance design, there is no perfect product design. In today's knowledge economy, the shape design of products has irreplaceable significance in the whole process of product manufacturing. This paper starts with the characteristics of product appearance design, expounds the development significance of the product appearance design system, and then describes the goals that the automatic product appearance design system needs to achieve and the characteristics of the work. Finally, the process of system design is expounded on the basis of intelligent algorithm.

**Keywords:** Intelligent Algorithm · Product Appearance · Appearance Design · System Design

## 1 Introduction

Product design refers to the decorative or aesthetic qualities of an item. Product design may contain a large number of 3D elements, such as the shape and feel of the product, or flat elements such as shape, curve and color [1]. The product range is also very broad, from technology and medical equipment to watches, jewelry and other luxury goods, from home appliances to cars and buildings. The product design is exquisite, which directly reflects the overall appearance of the enterprise. The product shape design is to serve the product and create a good environment for the product to enter the international market.

## 2 Aesthetic Features of Product Design

With the development of modern economy and society, people's consumption thinking and aesthetic concepts have undergone tremendous changes, and the appearance of modern commodities has become increasingly personalized [2]. Looking at the previous products, there are relatively few style choices, mainly to achieve those main functions. In addition to feature selection, the design also has many personality preferences, including handsome, elegant, noble, cartoon and other colorful appearance changes to meet the unique taste of people with various positions and identities. The main connotation

of design is ergonomic factors, as well as the artistic significance of product design. Pure plastic art, aiming to seek the sensual beauty produced in nature or influenced by the artist. Compared with the shape design of the commodity, it must meet the basic requirements and represent a technical solution. The design of a commodity needs to use rational and logical thinking to guide perceptual and imaginative thinking, and take problem-solving as the standard. It can be said that it is impossible to play freely [3]. Product design is not simply seeking personal aesthetic value, but a universal and diverse life value orientation. It involves intellectual property, scientific and technological value, economic benefits, social and psychological significance, etc. Based on the above objective value orientation, the design of goods must adhere to the following three main criteria: practicality, economy and aesthetics. The so-called practicality means that when using a commodity, the practicality of the commodity must be fully considered in order to achieve the requirements of comfort, speed and safety. Moreover, people's aesthetics and consumption patterns should also be considered. Therefore, the design of the product should not be guided by obvious formalism, but should only focus on the practicality of the product, and cannot ignore other factors [4]. The so-called practical means that the product needs to be adapted to the most advanced production technology in modern times, in order to obtain the highest benefit with the smallest financial resources, material resources, labor and time. The so-called aesthetics means that the design of the product shows a complete, healthy and harmonious new environment for production art under the condition that it meets the purpose and technological requirements, decorates people's life and health as much as possible, and develops noble and happy aesthetics.. Practicality, economy and aesthetics are inextricably linked and cannot be ignored. Only by organically combining and coordinating practicability, economy and aesthetics in the design and production process, the product can fully display the creative design concept and provide people with better services [5].

### **3 Development Significance of Product Appearance Design System**

#### **3.1 Research Background of Product Design Process**

The product design process is a complex thinking process, with a wide range of high uncertainties, covering most areas. In recent years, people's requirements for product design appearance design are also increasing. In order to efficiently carry out the customized design of the product design shape, through the design program of the functional components that can change or share the product design shape, and use CAD technology to integrate each functional unit to achieve the diversity of product design appearance design [6], Comprehensive evaluation of different appearance features, so as to get the design scheme that best meets the designer's requirements. Although the design process is very simple, this scheme cannot be realized; another idea is to decompose the knowledge about the design process into reuse knowledge, and obtain the knowledge about the designer's design process and the knowledge about the user's needs. On this basis, a knowledge base about the appearance of product design is established, and then the prototype system is used to get the design scheme. Although this method has high design efficiency, it relies too much on prior knowledge, and the obtained results cannot meet the needs of users. Another aspect is to propose an artificial intelligence-based

product design automation system, which uses the appearance to synthesize the product design system and introduces the possibility of changing the skin, so that the product design can be completed without programming [7]. According to the deficiencies of these schemes, this paper presents a new product visual collaborative design method based on intelligent algorithm based on human-machine collaborative interaction, and introduces the general composition of the design environment.

### 3.2 Elements and Features of Automatic Product Appearance Design System

#### (1) Beauty and uniqueness

The beauty of design is something that people cannot control. But the more things that can't be mastered, the more worthy designers are eager to achieve perfection. However, uniqueness is one of the important factors in the formation of style, and it is the key to product design. Product design attracts attention, makes consumers interested, and quickly generates attractive purchasing motives [8]. Product appearance design must pay attention to color and shape, that is, personality and style. Unique design that conveys the messages "I am different" and "I represent an unforgettable atmosphere".

#### (2) Cater to the preferences of the market and consumer groups

The market competition is becoming more and more fierce, and the products are updated more and more, and the appearance also needs to be constantly updated to adapt to the market demand and price trends. Therefore, enterprises must promote product innovation to adapt to the ever-changing needs of customers, so as to achieve the goal of further enhancing product competitiveness [9]. In the design, it is also very beneficial to inject a pleasing exterior design. For example, the popular bionic design can make people resonate with the aesthetic taste of the present. For another example, popular bionic designs can easily guide consumers to agree on aesthetic preferences. For example, the exterior design of the Geely Panda car is a very mature case of bionic product design [10]. As the world's second bioengineered car after the Volkswagen Beetle, the Geely Panda is also becoming a classic. The innovative design of the product is mainly based on the current popular information and conforms to the development of the times. It can also bring a larger market to the company.

#### (3) A large number of high-tech applications

The quality of design reflects to a considerable extent the degree of material development, scientific and technological achievements in our country, and is related to the level of scientific and technological development in our country. With age, design is more strongly influenced by technical forms and types of knowledge [11]. First, China's booming high-tech industry has introduced a lot of new technologies from traditional production. The second is the diffusion of technology in traditional industries, which has greatly changed the production mode of traditional Chinese industries, popularized and applied in traditional Chinese production, and integrated many advanced new technologies. Third, the application of various 3D digital software technologies in the entire product design process greatly improves the quality and speed of design.



## (4) Multifunction

The rapid progress and wide application of microelectronics, new materials, new energy and high nanotechnology have greatly reduced the quantity, weight and cost of materials required to perform the function of a single product, thus realizing the integration of product functions. The multi-functional integrated product is the inevitable result of people's needs, technological development and changes in market rules, and is more reflected in the product's shape and product design. In order to express the appearance of function, its life span is longer, but for the appearance of style requirements, its life span is shorter. The integrated design is more competitive in the market economy. So now, one or more mobile phones in people's hands is a good reflection of the power of multifunctional integration [12].

## (5) Highlight people-oriented

The ergonomic appearance in the design is designed to make the product more convenient and smooth to use. In order to meet the simple aesthetics, the shape of traditional industrial product design began to emphasize practicality, making the product suitable for human use. In our information age, the meaning of humanization is becoming more and more abundant [13]. In the past, the traditional definition of personalization often involved more intuitive spiritual elements, including personal safety, comfort, harmony with the surrounding environment, etc., but now humanization refers to the spiritual elements of a consumer. Consumers not only able to use electronic products safely and happily, and at the same time get spiritual experience in the practical application of electronic products [14]. The design, development, production, marketing and utilization of commodities all revolve around the existence of "people", while the social significance of commodity existence and development all depend on human emotional factors. In this way, the reason for judging people becomes particularly important.

## 4 Ways to Implement Intelligent Algorithms

### 4.1 Description Transformation of Designer Knowledge

Assuming that  $e_{ij}$  ( $1 < i < n$ ,  $1 < j < k$ ) is used to describe the  $j$ th technique, there are  $k$  techniques that can convert the  $i$ th innovation element into functional parameters, then the set of conversion techniques can be constructed:

$$E = [E_1, E_2, \dots, E_n]^T$$

In the formula  $E_i = [E_{i1}, E_{i2}, \dots, E_{ij}]^T$ . According to the designer's cognitive type, different description conversion methods and rules are used, and data mining method is used to convert the designer's cognitive description.

### 4.2 Determination of Design Parameters

The parameters involved in the product design process can be transformed into the basic parameter description problem through reconversion. The formula can be described as:  $X = AX_0$

In the formula,  $X_0 = (x_{10}, x_{20}, \dots, x_{j0} \dots x_{n0})^T$ ,  $x_{j0}$  describes the functional parameters in the product design;  $A$  describes the parameter transformation matrix, and its calculation formula is as follows:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

Because there is an inevitable relationship between various product appearance and design parameters, which leads to their contradictions and conflicts, the process of product design and appearance design is the process of resolving parameter contradictions. The entire problem solving process can be realized by using the solution state space of the product design problem:

$$F(X) = [-x, x_j] = - \begin{bmatrix} x_1 & \cdots & x_1 \\ \vdots & \ddots & \vdots \\ x_m & \cdots & x_m \end{bmatrix}$$

In the formula, it is used to describe  $[-x, x_j]$  the conflict resolution matrix between the  $F(X)$  parameters  $x_i$  and;  $x_j$  it is used to describe the operations that generate conflict resolution and problem state transition. Through the above analysis, the following solution state space for innovative design problems can be obtained, which can be described by the following two-tuples  $X_s, X_r$ , which  $X_s$  describes the initial problem state vector:

$$X_s = [x_1, x_2, \dots, x_{ms}]^T$$

It is exactly what the designers expected from the design solution.  $X_r$  Represents all the questions that can be solved using the TRIZ contradiction matrix, the equation states:

$$X_r = [x_1, x_2, \dots, x_{mr}]^T$$

According to the above analysis method, the binary array ( $X_s, X_r$ ) can be used to evaluate the designer's cognitive semantic mode, and then the product appearance evaluation result obtained by the evaluator can be used to obtain the final design scheme.

In the design process, the color quantification data of the preliminary color scheme completed by the designer is substituted into the color image prediction model, the color multi-objective image value of each preliminary scheme is calculated, and the optimization direction of the preliminary scheme is clarified by comparing with the color design objectives. In the process of clarifying the optimization direction, different color scheme design ideas are formed according to whether there is a consistent correlation between key color variables and multi-objective imagery. If there is a consistent correlation, use the correlation to adjust the key color variables to generate an optimized color scheme [15]. On the contrary, the color multi-image optimization model is used to complete the evaluation and optimization of multiple preliminary product color schemes. After completing the optimization, judge whether the optimized color scheme meets the color

design goals. The process of selecting the best color implementation plan is jointly judged by all the subjects, and see whether the results are consistent with the overall goal of the engineering design. If it is not completely consistent, continue to iterate until the final result is the final color implementation that is consistent with the overall goal of the engineering design. Program.

### **4.3 Automatic Design System for Optimizing Product Shape**

Computer-aided product appearance design technology is generally composed of design modules such as design analysis, appearance prediction, appearance multi-objective optimization and optimal appearance plan decision-making. The technologies used in the two design modules of design analysis and appearance prediction in the previous article process and analyze data in an objective way, which can realize computer-aided design. However, in the appearance generation and evaluation links corresponding to the two design modules of appearance multi-objective optimization and optimal appearance plan decision, because the subjective experience of designers is required for manual optimization and decision-making, the goal of full-process computer-aided design has not yet been achieved.. Since the multi-objective optimization method in the intelligent algorithm is the most core technology that determines the realization degree of the optimization goal of the electronic product design technology, and the multi-objective optimization method in the previous research has the defects of algorithm performance and application scope, Therefore, a new ISPEA II improved multi-objective optimization method is provided.

In this method, the improved crossover operator and adaptive mutation operator are used to enhance the search ability of the computational solution space, and the self-correction operator is creatively introduced to further improve the computational convergence ability. Through the design of these operators, ISPEA2 makes up for the performance defects of the original algorithm, and has the applicability and effectiveness for computer-aided product design. After completing the automatic calculation of the product appearance multi-objective optimization design module, the generated Pareto appearance scheme set is a set of individual appearance schemes containing multi-objective image adjective calculation scores. In order to simulate the design process through the computer, a multi-attribute decision-making method is needed to complete the optimal decision of the appearance scheme. The technologies used in the previous design modules have accumulated the advantages of product design innovation and multi-image matching accuracy. In order to consolidate the advantages and obtain accurate decision-making results, the optimal solution decision-making design module needs a method that can objectively utilize multi-target images. The adjective calculation score data is automatically calculated, and it is a multi-attribute decision-making method with strong operability and multi-technology fusion.

## **5 Conclusions**

The appearance design of the product appears in front of human beings in a physical state, and is created by various means and artistic methods according to the laws of

function and aesthetics. In the process of today's society entering the information age, it has penetrated into all areas of human daily life. Product design concerns have also turned to finding connections between individuals and objects. The satisfaction of the design object is essentially the satisfaction of the design concept and the aesthetic feeling of the public. In fact, the word "design" already contains the word "aesthetics". The special expression of product design penetrates into the daily life of human beings more extensively and deeply. Product design is the result of the endless expansion of the field of applied art and industrial development, the result of the invasion of aesthetics to the technical field and the invasion of art to production. Product design is a comprehensive aesthetic form, including material culture, spiritual culture and art culture.

At present, although the intelligent algorithm can realize the establishment of the product shape automatic design system, the research on the use of the multi-attribute decision-making method in the field of product appearance design is still very limited. However, this technique has the shortcomings of objectivity and convenience of operation, and the accuracy of the obtained results also needs to be improved. Since the calculation scores of multi-target image adjectives are completely automatically generated by the computer and do not involve subjective reasoning, the TOPSIS method is more objective and accurate compared with the subjective evaluation methods such as the fuzzy analytic hierarchy process. It is suitable for multi-attribute decision-making and other features, and can be well integrated with the main technologies of other design modules. Therefore, it is selected as the multi-attribute decision-making method of the optimal appearance design decision-making module to form a complete product appearance design technology system and application performance. The article is expected to pave the way for subsequent research.

**Acknowledgements.** This work was financially supported by Hunan Provincial Department of Education "Research on Tantou New Year Pictures under the Vision of Intangible Cultural Heritage Protection" (21C0594); Shaoyang Science and Technology Bureau "Cultural and Creative Product Design based on Shaoyang Blue calico Technology Characteristics" (2021120ZD).

## References

1. Zhu, X., Lv, W.: Intelligent analysis of core identification based on intelligent algorithm of core identification. *Discrete Dynamics in Nature and Society* 2021 (2021)
2. Li, F.: Network collaborative modeling design of intelligent products in industrial design system CNC machine tools. In: 2015 International Conference on Automation, Mechanical Control and Computational Engineering. Atlantis Press (2015)
3. Ding, M., Bai, Z.: Product color emotional design adaptive to product shape feature variation. *Color. Res. Appl.* **44**(5), 811–823 (2019)
4. Chandrasegaran, S.K., et al.: The evolution, challenges, and future of knowledge representation in product design systems. *Computer-Aided Design* **45**(2), 204–228 (2013)
5. Chan, W.M., et al.: A 3D CAD knowledge-based assisted injection mould design system. *The Int. J. Adv. Manufacturing Technol.* **22**(5), 387–395 (2003)
6. Guo, S., et al.: Vinci: an intelligent graphic design system for generating advertising posters. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (2021)

7. Satam, D., Liu, Y., Lee, H.J.: Intelligent design systems for apparel mass customization. *The Journal of The Textile Institute* **102**(4), 353–365 (2011)
8. Mao, D., et al.: Visual and user-defined smart contract designing system based on automatic coding. *Ieee Access* **7**, 73131–73143 (2019)
9. Granadeiro, V., et al.: Building envelope shape design in early stages of the design process: Integrating architectural design systems and energy simulation. *Automation in Construction* **32**, 196–209 (2013)
10. Weingaertner, T., et al: Smart contracts using blockly: Representing a purchase agreement using a graphical programming language. 2018 Crypto Valley Conference on Blockchain Technology (CVCBT). IEEE, pp. 5564 (2018)
11. Skotnica, M., Pergl, R.: Das contract-a visual domain specific language for modeling blockchain smart contracts. *Enterprise Engineering Working Conference*. Springer, Cham (2019). [https://doi.org/10.1007/978-3-030-37933-9\\_10](https://doi.org/10.1007/978-3-030-37933-9_10)
12. He, X., Chen, X., Li, K.: A decentralized and non-reversible traceability system for storing commodity data. *KSII Trans. Internet Inf. Syst. (TIIS)* **13**(2), 619–634 (2019)
13. Franz, F., Fertig, T., Schütz, A.E.: Democratization of smart contracts: a prototype for automated contract generation. In: 2020 IEEE International Conference on Blockchain and Cryptocurrency (ICBC). IEEE (2020)
14. Bessa, E.E., Martins, J.S.B.: A blockchain-based educational record repository. *arXiv preprint arXiv:1904.00315* (2019)
15. Hunter, E.: Graph-manipulation based domain-specific execution environment. *U.S. Patent No. 11,132,403* (2021)



# Sports Video Tracking Technology Based on Optimized Decision Tree Algorithm(DTA)

Zhong Wu<sup>(✉)</sup>

Physical Education School, Wuhan Business University, Wuhan, China  
7849800@qq.com

**Abstract.** With the continuous progress of society, video surveillance system has been more and more used in various occasions, and is gradually developing in the direction of intelligence. Moving target detection and tracking has always been the key to the intelligence of video surveillance. This paper analyzes the research of optimal DTA in sports video tracking technology(VTT); In terms of moving target tracking, firstly, the existing classification of moving target tracking and common moving target tracking methods are summarized, and then the improved epanchnikov kernel function target tracking algorithm is introduced, and the optimized DTA is introduced into the tracking technology. Through the experimental data, it is found that among the six groups of data tested, the accuracy of the optimized DTA has reached more than 75%, and the lowest recall has reached 77.45%. It is proved that the optimized DTA has good comprehensive performance and high precision in sports VTT.

**Keywords:** Optimized Decision Tree Algorithm · Sports · Video Tracking · Tracking Technology

## 1 Introduction

Moving target tracking technology has indispensable and important applications in many fields. So far, tracking moving targets is still a difficult task: the moving target object itself has a complex shape. In terms of speed, most systems have real-time requirements for the tracking process. At present, the tracking algorithm based on sports video moving target still can only deal with the visual information of single angle, and can not obtain the omni-directional information of the whole three-dimensional moving target. Therefore, it can achieve good tracking effect when the requirements for foreground details are not high. Once the requirements for foreground details are improved, it is often difficult to achieve good results. Based on this, this paper studies and analyzes the application of optimized DTA in sports VTT.

Many scholars at home and abroad have analyzed the research of optimal DTA in sports VTT. Mulimani n proposes a new framework to implement the sports extension of athlete recognition and tracking technology to the implementation of event processing process. In this method, Kalman filter technology is used in video preprocessing; By using contour point model and morphological operation, more real player detection

results are given in spatial domain. The algorithm is tested on self-developed video data and real-time video under dynamic background, and higher tracking accuracy and performance index are obtained under different video sample states [1]. Ortega J P in view of the accuracy of data collection, radar based local positioning system is a promising technology for monitoring team sports training load. The effectiveness and reliability of semi-automatic video technology and global positioning system in team sports are evaluated [2].

The complexity of sports brings many difficulties to the actual detection and tracking of moving targets. In order to effectively detect and track athletes, this paper improves the commonly used single tracking algorithm and proposes the optimized DTA, which further improves the effect of sports target detection and tracking and greatly improves the robustness and accuracy of sports target tracking [3].

## **2 Research on Optimal DTA in Sports VTT**

### **2.1 Motion VTT**

How to select the characteristics of the target is a very important step in the process of target tracking. The representation methods of targets mainly include point, contour, shape, histogram and so on. Usually, the feature selection of target is closely related to the representation of target. For example, when tracking the moving target represented by histogram, the color of the target is usually selected as the feature; For tracking the object represented by contour, the edge feature of the target is usually selected. In order to effectively improve the robustness of target tracking, most systems often select multiple features of the target for tracking in order to achieve satisfactory results.

Generally, it is considered that a suitable feature should have the following characteristics. Targets of the same type should have the same or similar eigenvalues; The quantity is appropriate. Selecting too many features will greatly increase the amount of calculation of the system, while insufficient features will lead to the inability to distinguish the target, so it is necessary to have an appropriate number of feature sets; Good discrimination. The eigenvalues of the target are significantly different from other non tracking objects in the scene;

### **2.2 Analysis of Mean Shift Tracking Algorithm under Occlusion**

Occlusion is a major interference that needs to be solved in the process of target tracking. In the process of tracking, the probability of target occlusion is very high, and occlusion is often encountered in the process of tracking. Most importantly, when the target is occluded, the target no longer has any information in the whole image. At this time, target tracking mainly depends on the prior information of the video [4, 5]. The main function of prior information is to enable the tracking algorithm to predict the location of the target, so as to lock the tracking target again when the target reappears. When the target enters the occlusion area, the color information of the target will be affected, the color feature distribution of the selected target and candidate area will change, and the tracking algorithm will track the wrong target or even lose the target; When the target leaves the occlusion area, the color information of the target is restored.

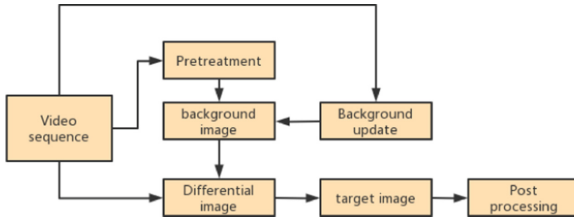
The target tracking algorithm must have the following four functions to track the target effectively under occlusion: it can track the target stably under normal conditions; When occlusion occurs, it can judge whether the target is occluded; In the process of occlusion, the motion of the occluded target can be predicted; After occlusion, the target can be locked again. Therefore, this paper proposes an improved kernel function to enable the mean shift algorithm to deal with the tracking problem under partial occlusion [6, 7]. For the case that the tracking target is seriously occluded, this paper adds the linear recursive Kalman filter to the mean shift algorithm framework, so that the algorithm can predict the motion trajectory of the moving target in the case of occlusion, so that the algorithm can deal with the target tracking problem in the case of severe occlusion.

In the process of target tracking, partial occlusion is often encountered. When the target is partially occluded, the information of the target is partially missing, and the similarity between the selected target and the candidate area will be reduced, which will lead to the loss of the target in the tracking algorithm. In most tracking scenarios, for epanechnikov kernel function, the weight of the center pixel of the selected tracking target is 1, but the weight value of its edge pixel can reach more than 0.5 [8, 9]. Because the weight of edge pixels is too high and the weight proportion of edge pixels to the whole feature distribution is too high, when partial occlusion occurs, the loss of these edge pixel feature information will lead to great changes in the overall feature distribution of the candidate region, so the similarity between the candidate region and the selected target will be sharply reduced, which will lead to the loss of the tracking target. Therefore, a kernel function is needed to make the weight of the edge pixel lower and the weight of the center pixel higher. Through the analysis of epanechnikov kernel function and Gaussian kernel function, an improved kernel function is proposed in this paper, as shown in formula (1):

$$l(a, b) = \exp \left[ \sigma \times \left( -\frac{(a - \hat{a})^2}{r_b} - \frac{(b - \hat{b})^2}{r_a} \right) \right] \quad (1)$$

where  $a$  and  $B$  represent the coordinate values of any pixel in the target area. Its weight is 0.01, and its value range is  $0.01 < \sigma < 0.5$ ,  $\sigma$  The function of parameter is to control the weight of edge pixels so that they will not change dramatically.  $(R_A, R_B)$  is the size parameter of the tracking frame and  $(\hat{a}, \hat{b})$  represents the coordinate of the center point of the tracking frame. The kernel function shown in formula (1) can make the weight of the central pixel higher and the weight of the edge pixel smaller. Therefore, in the actual tracking process, the kernel function gives more reasonable weights to the target pixels. In case of partial occlusion, the mean shift tracking algorithm can still track the target correctly [10, 11].





**Fig. 1.** Principle block diagram of background difference

Background difference method is the most commonly used moving target detection algorithm at present. The principle block diagram is shown in Fig. 1:

### 2.3 Optimization of DTA

In the process of pruning the decision tree, we need to follow some optimization principles: the principle of minimum expected error rate: use pruning method to repair the interior of the tree, compare the expected error rate if it is not repaired, and discard the situation of high error rate.

For the evaluation of a decision tree, there are some quantitative evaluation criteria, and the correct classification is the first consideration. If the data of a decision tree is quite complex, it is difficult for users to understand, so the significance of this decision tree is not so important. Therefore, if the branches and leaves constituting the decision tree are relatively few, the classification process is relatively simple, and the space occupied by storing data is relatively small, so when constructing the tree, pay attention to be as concise as possible, and do not build the branches and leaves of the tree too large.

Target tracking algorithm integrating feature point detection and spatiotemporal context information: spatiotemporal context tracking algorithm has good robustness and fast running speed. It is a good target tracking algorithm. However, when the target is blocked and moves rapidly in the process of movement, the spatiotemporal context information of the target cannot be updated in time, which is easy to cause tracking drift or even failure, However, the algorithm can not correct the errors in the tracking process by itself, and still uses the wrong information to track the next frame. In order to solve these problems, a target detection algorithm integrating feature point detection and spatio-temporal context information is proposed. The algorithm not only learns the spatiotemporal context information of each frame, but also detects the feature points of the target, and then judges whether the tracking is effective according to the feature point matching. When the tracking is effective, it continues to use the spatiotemporal context algorithm for tracking. When the tracking is invalid, it needs to judge whether the target is blocked or the tracking is offset, and then process it according to the judgment results. The algorithm improves the tracking robustness of fast moving targets, Moreover, it solves the problem of tracking failure caused by occlusion and overcomes the shortcomings of spatio-temporal context tracking algorithm [12].

### 3 Optimal DTA

Decision tree algorithm shows some limitations, but it also shows high accuracy. For pruning, each point on the decision tree may be processed. Assuming that there are  $K$  ( $s$ ) leaves in this tree, the classification error of  $H$  ( $H$ ) is:

$$r'(H_1) = \frac{\sum_c [e(c) + 1/2]}{\sum_c m(c)} = \frac{\sum e(c) + \frac{K(c)}{2}}{\sum m(c)} \quad (2)$$

where,  $T$  represents the decision tree,  $s$  represents the subtree,  $H$  represents the original decision tree, and  $n$  ( $H$ ) is the branch tree;  $E$  ( $H$ ) represents the number of misclassifications and  $j$  represents the error rate of the data set.

Generally speaking, the condition that an intermediate node  $h$  is replaced by a leaf node is that the error rate of the replaced child number  $H_1$  is less than that of node  $H$ . If the classification method is applied to classify the data of all sample sets, then

$$e'(H_1) + CE[e'(H_1)] = \frac{1}{2} [K(H_1) + \sqrt{K(H_1)}] \quad (3)$$

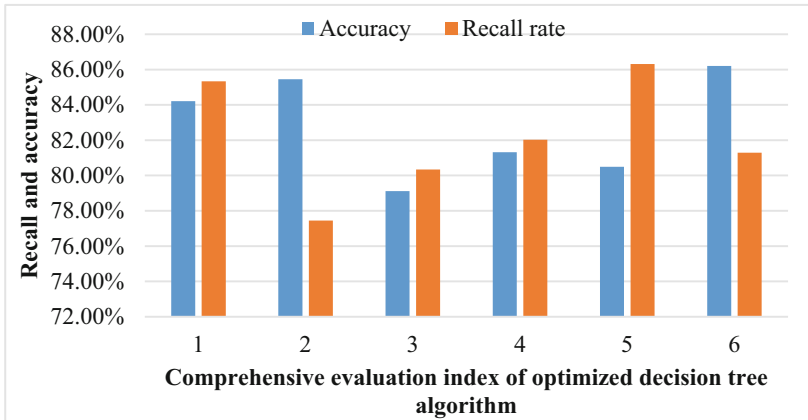
### 4 Experimental Test and Analysis

In order to verify the feasibility and effect of applying the optimized DTA to sports VTT, this paper tests the comprehensive evaluation indexes (accuracy and recall) of the optimized DTA in sports VTT. The test results are shown in Table 1 and Fig. 2

**Table 1.** Data table of comprehensive evaluation index of optimized DTA

experience group	1	2	3	4	5	6
Accuracy	84.21%	85.45%	79.12%	81.32%	80.49%	86.21%
Recall rate	85.34%	77.45%	80.34%	82.03%	86.32%	81.29%

It can be seen from the above chart data that among the six groups of data tested, the accuracy of the optimized DTA has reached more than 75%, the lowest is 79.12%, and the highest is 86.21%; The recall rate was the lowest, reaching 77.45%. It is proved that the optimized DTA has good comprehensive performance and high precision in sports VTT; When the target stays for a short time and moves rapidly, the integrity of the detected target is high, and when there is more interference in the scene where the target is located, the removal effect of the target interference points is also good. At the same time, the accuracy, recall and comprehensive evaluation index of the algorithm are high.



**Fig. 2.** Comprehensive evaluation index of optimized DTA

## 5 Conclusions

On the whole, this algorithm improves the existing target detection and tracking algorithms, and solves some problems in target detection and tracking, but there are still some deficiencies that need to be further improved: in terms of moving target tracking, this algorithm has a good effect on target occlusion and fast-moving target tracking, but the focus of these two aspects is the tracked target, In the future research, we should also comprehensively consider the difficulties brought by the environmental factors around the target to target tracking, and expand the application scope of the algorithm. Although the target detection and tracking algorithm in this paper has achieved satisfactory results in performance, it is still in the experimental stage, and there is still a certain distance from the practical application, which needs further research.

**Acknowledgements.** This work was supported by the Application Foundation Frontier Project (2020010601012294), the Wuhan University of Business Doctoral Fund projects (2021KB002), the Provincial University Teaching Research Project of Hubei Province (2018466), and the Teaching and Research Program at City-owned Universities in Wuhan (2019102).

## References

1. Mulimani, N., Makandar, A.: Sports video annotation and multi- target tracking using extended gaussian mixture model. *Int. J. Recent Technol. Eng.* **10**(1), 1–6 (2021)
2. Ortega, J.P., Lozano, J., Gantois, P., et al.: Comparison of the validity and reliability of local positioning systems against other tracking technologies in team sport: a systematic review. *Proceedings of the Institution of Mechanical Engineers, Part P: J. Sports Eng. Technol.* **236**(2), 73–82 (2021)
3. Athanasiadis, P.J.: On the behavior of slackline webbings under dynamic loads and the simulation of leash falls. *Proceedings of the Institution of Mechanical Engineers, Part P: J. Sports Eng. Technol.* **233**(1), 75–85 (2019)

4. Balasundaram, R., Devi, S.S.: Genetic algorithm-based hybrid approach for optimal instance selection of minimising makespan in permutation flowshop scheduling. *Int. J. Business Intelligence Syst. Eng.* **1**(3), 197–225 (2019)
5. Rana, M., Mittal, V.: Wearable sensors for real-time kinematics analysis in sports: a review. *IEEE Sensors J.* **21**(2), 1187–1207 (2020)
6. Nancy, P., Muthurajkumar, S., Ganapathy, S., et al.: Intrusion detection using dynamic feature selection and fuzzy temporal DT classification for wireless sensor networks. *IET Commun.* **14**(5), 888–895 (2020)
7. Sanz-Corretge J.: A multi-stage algorithm for the FEM design of composite sandwich panels subjected to multiple manufacturing rules. *Applied Composite Mat.* **29**(2), 423–449 (2021)
8. Regensburg, K.W., Meyer, T., Barbara, G.: Kontaktzeiten im Profifuball vor der Covid-19-Pandemie – Eine retrospektive analyse der trackingdaten von fuball-bundesliga-spielen. *Sports Orthopaedics and Traumatology* **37**(2), 189–190 (2021)
9. Secular, S.: The numbers game: the NBA v. motorola, real-time statistics, and the rise of online fantasy sport. *International Journal of the History of Sport* **38**(1), 79–94 (2021)
10. Jooste, F.J., Costello, S.B.: An efficient and stable road maintenance strategy evaluation algorithm utilizing a trigger horizon. *Transp. Res. Rec.* **2674**(11), 424–439 (2020)
11. Wilk, M.P., Walsh, M., O’Flynn, B.: Multimodal sensor fusion for low-power wearable human motion tracking systems in sports applications. *IEEE Sensors J.* **21**(4), 5195–5212 (2020)
12. Mallepalle, S., Yurko, R., Pelechrinis, K., et al.: Extracting NFL tracking data from images to evaluate quarterbacks and pass defenses. *J. Quantitative Analysis Sports* **16**(2), 95–120 (2020)



# Entrepreneurship Platform Under Virtual Reality Technology (VRT)

Lili Wang<sup>(✉)</sup>

School of Economics and Management,  
Mudanjiang Normal University, Mudanjiang 157012, Heilongjiang, China  
liliwang@mdjnu.edu.cn

**Abstract.** With the economic development and the enrollment expansion of colleges and universities, the function of universities has changed from “elite education” to “mass education”, and improving the innovation and entrepreneurship service of colleges and universities has become an important task. At present, the entrepreneurship support platform has the following problems: the single support mode is difficult to make up for the short board of College Students’ entrepreneurship; The platform technology lags behind and cannot well adapt to the mobile Internet; The lack of support services can not meet the diversified needs of entrepreneurial college students. Therefore, it is a good choice to apply VRT to the entrepreneurial platform. This paper studies the application of VRT in the entrepreneurial platform, discusses the VRT and its characteristics, the sub-system design of the entrepreneurial platform and the application of VRT in the entrepreneurial platform, tests and analyzes the experience of students and trainers in the application of VRT in entrepreneurship through the index evaluation of the application of virtual technology in the entrepreneurial platform, And test the entrepreneurship under virtual technology and the entrepreneurship effect under the traditional mode. Under different modes, the proportion of entrepreneurship “always adhere to”, “stop halfway”, “entrepreneurial success”, “entrepreneurial failure” and “always develop” is. Teachers and students in the entrepreneurship platform under VRT said that they had a good experience, which significantly improved their interest in learning and success rate. Under the VRT, 86% of them can always insist on entrepreneurship, of which 55% are successful, and 43% of them have been developing and operating; In contrast, only 31% of entrepreneurs under the traditional mode can always adhere to entrepreneurship, only 23% succeed in entrepreneurship, and 21% of companies have been operating and developing. It can be seen that the application of VRT in Entrepreneurship improves students’ learning initiative, enhances real experience, and plays an important role in the success of entrepreneurship and the operation and development of the company.

**Keywords:** Virtual Reality Technology · Entrepreneurship Platform · Virtual Technology · Application Research

## 1 Introduction

VRT is mainly an emerging science and technology under the prospect of the rapid development of electronic computers, the gradual maturity of simulation technology and the wide application of artificial intelligence. Through the imaging of simulation technology in computers, consumers can feel the real situation only through the relevant VR equipment, giving users a sense of immersion and having the characteristics of various realistic sensory experiences. At present, the application of VRT in the construction industry is mainly for virtual building roaming. After the construction of virtual environment with software, the interaction of buildings in the virtual environment can be realized by wearing VR helmets or other external equipment. This paper applies VRT to entrepreneurship platform, and discusses it.

Many scholars at home and abroad have studied the application of VRT in Entrepreneurship platform. Aliyu f discussed the latest application of VRT (VRT) in the educational environment and the effective strategies that Nigerian pre service chemistry teachers need to adopt in learning chemical concepts. By consulting the existing literature on VRT in chemistry, this paper emphasizes some benefits of VRT for pre service chemistry teachers who have difficulties in teaching the content knowledge of abstract chemical concepts. Some advantages of VRT include interactivity, immersion and visualization [1]. AKB a developed computer software using virtual environment technology as a tool to develop new educational methods for these courses. Use VRT to provide 4D model for specific construction stage, and provide users with immersive and non immersive virtual reality experience. The results show that using BC \ VR software as a tool for building construction course is very useful and effective for students [2].

This paper improves the traditional entrepreneurial model, introduces VRT into the entrepreneurial platform, and discusses the real virtual technology, the matters to be faced in Entrepreneurship and the entrepreneurial design process. Experiments have verified the advantages of the application of virtual technology to the entrepreneurial platform, and improved the enthusiasm and success rate of entrepreneurs. The application of VRT to the entrepreneurial platform is the general trend [3].

## 2 Application of VRT in Entrepreneurship Platform

### 2.1 VRT and its Characteristics

#### (1) Virtual Reality Technology

The reality felt by the human body is actually formed by the transmission of various sensory stimuli of the human body to the brain through nerves. Its essence is the real image formed in the brain through the stimulation of the outside world to the human body. It can be virtual. When the five senses provided to users and any interaction generated by users are real-time and natural, users can not distinguish between virtual and reality. In academia, it is called “brain in the tank”, just like what human beings sleeping in the matrix feel. Therefore, virtual reality is actually an advanced computer user interface. It imitates the external stimuli perceived by the human body and belongs to a kind of computer simulation technology. Virtual reality can create and experience the virtual world, which is the goal that human

beings have been pursuing. In this virtual space, people's perceptual and rational cognitive abilities can be brought into full play [4, 5].

(2) Characteristics of VRT

Interactivity refers to the active participation of users and the real-time and naturalness of interaction between users and the environment, emphasizing the naturalness in the process of human-computer interaction

Sex. At present, the interactivity in virtual reality system is mainly realized through nine aspects. Conceivability is the rationality of the existence of all kinds of things (real or imaginary) in the virtual environment. Most of other digital media also have this feature, such as animation, film and so on. Therefore, immersion and interactivity are the essential characteristics of virtual reality different from other media, emphasizing autonomous human-computer interaction. The virtual reality environment or completely simulates the reality or the reality that can not be realized by adding the idea, closes people's senses. Only when the feedback of the virtual environment conforms to the interaction between people and the natural environment, and when the 3I characteristics of virtual reality are fully reflected, can people experience the perfect virtual reality like a dream.

(3) Virtual reality classification

From the content of virtual reality, at present, this type of virtual reality is most widely used, such as panoramic scenic spot experience, simulating natural disasters, simulating flight driving and so on. Surreal type is to add things that the human body cannot perceive at ordinary times on the basis of the real environment, which can give full play to people's cognitive ability and exploration ability. For example, the virtual spaceport project designed by aerospace (Beijing) technology and Culture Development Co., Ltd. Allows users to experience virtual launch, space walk and other space activities. The pure virtual type makes full use of the conceptual characteristics of VRT, gives full play to human imagination and creativity, and creates situations that do not exist in the real world in the virtual space, such as the reproduction of fairy tales and fairy tales. Such applications are mostly used for leisure and entertainment projects to meet people's curiosity and beautiful fantasy of life [6]. The goal of different kinds of virtual reality systems is to enable users to experience all-round and intuitively perceive the methods of interaction. A typical virtual reality system consists of user, user interface, sensor module, input and output equipment, professional graphics processing computer and professional software.

## 2.2 Subsystem Design of Entrepreneurship Platform

According to the platform design objectives and the concept of mobile Internet, the entrepreneurship support system of the three innovation platform is designed into five subsystems, including store management client, market client, team management system, incubator management system and platform approval management system. Each subsystem not only realizes its own business functions, but also integrates and works together. By integrating and modeling the data in each subsystem, it can assist the decision-making

of various users on the platform, such as innovation activity guidance, professional information, value evaluation and reputation evaluation of people and enterprises, work and service recommendation, etc.

The main functions of each subsystem are described as follows:

- (1) Store management client: mainly responsible for the business function of applying for opening a store and the supply and demand management of work services. After the store is approved and opened successfully, users can put on the shelf and manage their works and services through the store, and participate in project bidding as the store. When there is a work or service order, users can jump to the store management end through the message organization to track the order processing process and service process [7, 8].
- (2) Bazaar client: mainly responsible for the classified display, keyword retrieval and detail viewing of works and services that have been put on the shelves. Through the market management end, users can view the spatial home page of the store, manage their own works and service needs on the palm, track the order process of works services at any time, and publish bidding information for specific tasks or projects. At the same time, based on the saving and analysis of user behavior data, the market management end will also promote works and services that users may be interested in.
- (3) Team management system: users can apply to join the project team through the team management system, apply for and manage their own project tasks, publish their own project work log, and team members can also view and exchange the information of other members and project progress through the management system [9].
- (4) Incubator management system: it is a handheld information management system of entity innovation and entrepreneurship incubation base based on mobile Internet technology.
- (5) Platform approval management system: it is the content management center of the entrepreneurship support platform, which is mainly responsible for the information review and approval process tracking of the whole platform, including all user registration, store application, work service on the shelf, personal professional functions, enterprise qualification, etc. the addition and editing of each information need to be approved, with corresponding permission control and log records, Only after being approved by the platform approval management system can it be displayed in each subsystem and further business operation permissions corresponding to information items be obtained.

### **2.3 Application of VRT in Entrepreneurship Platform**

In the design of the entrepreneurship platform, all subsystems rely on the same service background, save the data of each subsystem through the relational database, have permission control in each subsystem, and the data in the platform database can only be accessed by business operations with specified permissions. The entrepreneurship support function is mainly to provide information exchange and platform support for college students' innovation and entrepreneurship activities, and is committed to solving



the problem of information asymmetry and lack of knowledge guidance in the process of College Students' innovation and entrepreneurship under the background of mobile Internet. VRT provides a new experience, innovative thinking mode and technical means for entrepreneurship training [10, 11].

It has great advantages in improving learners' interest. When entrepreneurs interact with objects in the virtual environment, users can be fully immersed through its more perception and autonomy. In the design and R & D process of entrepreneurship simulation software system, through this technical means, the trained students can be immersive, more directly and truly feel the whole process of business environment and business operation, and then enhance the students' interest and initiative in participating in innovation [12].

### 3 Index Evaluation of Virtual Technology Applied to Entrepreneurship Platform

Virtual technology is applied to determine the weight of risk evaluation index of entrepreneurship platform. In the process of determining the weight of risk evaluation indicators of entrepreneurship platform, the expert consultation method can also be used. The specific steps to determine the index weight can be summarized as follows: determine the number of consulting experts, who must have an in-depth understanding of the entrepreneurial business model and project investment operation, such as operators with venture capital experience. Design the questionnaire, submit the questionnaire and data related to entrepreneurial organization management to experts for questionnaire survey, the degree of dispersion and coordination, and obtain the initial weight vector of each risk evaluation index according to the risk evaluation requirements of the entrepreneurial platform, as shown in formula (1)

$$K_{ab} = \frac{1}{v} \sum_{b=1}^n k_{ab}, (a = 1, 2, \dots, n) \tag{2}$$

where is the mean value of the weight value of  $K_a$  the  $a$ -th index;  $K_{ab}$  represents the weighted value of the  $b$ -th expert on the index, and  $V$  is the number of experts. In order to facilitate the comprehensive evaluation, the weight of risk index is normalized, as shown in formula (2)

$$k = \left( \frac{k_1}{\sum_{a=1}^n k_a}, \frac{k_2}{\sum_{a=1}^n k_a}, \dots, \frac{k_n}{\sum_{a=1}^n k_a} \right) \tag{2}$$

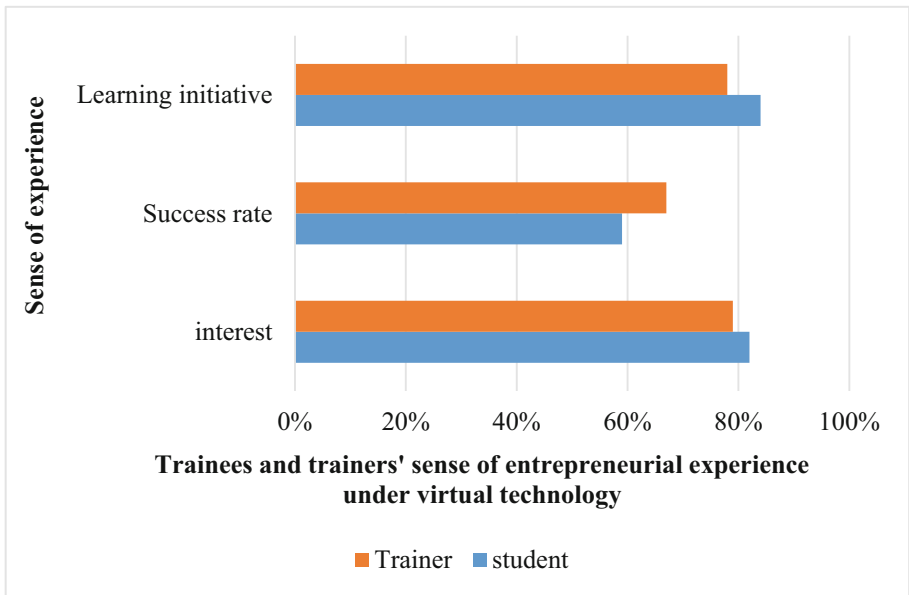
After completing the construction and weight determination of the risk management index system of the entrepreneurial platform, the fuzzy comprehensive evaluation is used to evaluate the single risk factor, the overall risk of the entrepreneurial platform and each investment project respectively, so as to provide support for the risk early warning and control of the entrepreneurial platform.

### 4 Experimental Test and Analysis

This paper tests the students’ and trainers’ experience of the application of VRT in Entrepreneurship in the entrepreneurship platform, including their interest in entrepreneurship, the success rate of entrepreneurship and their learning initiative. The results are shown in Table 1 and Fig. 1.

**Table 1.** Trainees and trainers’ sense of entrepreneurial experience under virtual technology

	interest	Success rate	Learning initiative
student	82%	59%	84%
Trainer	79%	67%	78%



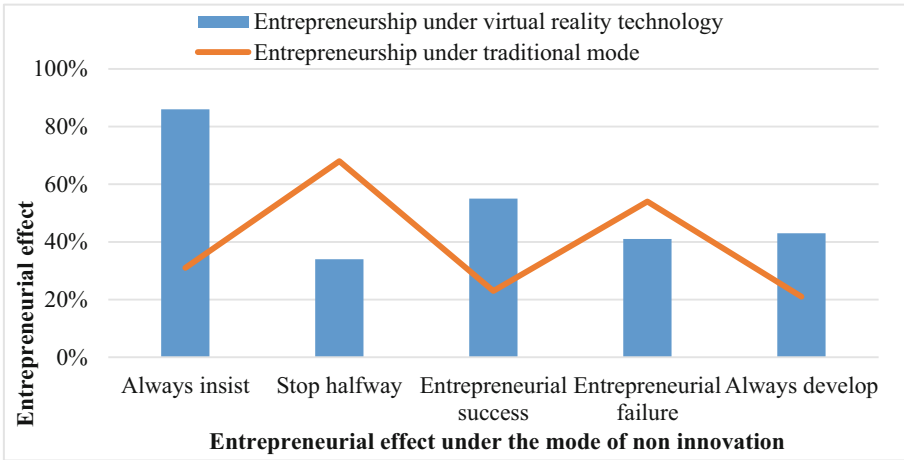
**Fig. 1.** Trainees and trainers’ sense of entrepreneurial experience under virtual technology

The results show that teachers and students have a good experience in the entrepreneurship platform under VRT, which has a significant improvement in improving learning interest and success rate.

Next, the entrepreneurship under virtual technology and the entrepreneurship effect under the traditional mode are tested. Under different modes, the proportion of entrepreneurship “always adhere to”, “stop halfway”, “entrepreneurial success”, “entrepreneurial failure” and “always develop” respectively. The test results are shown in Table 2 and Fig. 2.

**Table 2.** Entrepreneurial effect under the mode of non innovation

	Always insist	Stop halfway	Entrepreneurial success	Entrepreneurial failure	Always develop
Entrepreneurship under VRT	86%	34%	55%	41%	43%
Entrepreneurship under traditional mode	31%	68%	23%	54%	21%



**Fig. 2.** Entrepreneurial effect under the mode of non innovation

The test results show that under the VRT, 86% of them can always insist on entrepreneurship, 55% of them are successful, and 43% of the company has been developing and operating; In contrast, only 31% of entrepreneurs under the traditional mode can always adhere to entrepreneurship, and 68% stop entrepreneurship halfway, only 23% succeed in entrepreneurship, and 21% of companies have been operating and developing. It can be seen that the application of VRT in Entrepreneurship improves students' learning initiative, enhances real experience, and plays an important role in the success of entrepreneurship and the operation and development of the company.

### 5 Conclusions

With the development of the times, as an emerging technology, VRT has great potential and broad prospects and application fields. With the initial maturity of this technology, the development and research of VRT applied to entrepreneurship platform will be more in-depth. Virtual roaming with immersion will replace the current rendering and become the main method and source of entrepreneurial information. In the future, the virtual reality

entrepreneurship system can even rely on the online virtual reality roaming system to further develop virtual classrooms, virtual conferences and so on, so that campus teachers and students can obtain entrepreneurship information in a timely, real-time and real way without leaving home. Virtual reality entrepreneurship system will not only be satisfied with providing entrepreneurs with access to visit and browse, but also develop more diverse applications to adapt to the development of digitization. With the progress and development of science, the application of VRT to the planning and construction of entrepreneurship platform will be the main trend in the field of entrepreneurship in the future. Because the current conditions of this paper are limited and the experimental test data are not accurate enough, the real effect of VRT in entrepreneurship is expected to be further studied and discussed.

**Acknowledgements.** General Project of Higher Education Teaching Reform in Heilongjiang Province in 2021: Research on the training mode system of “new business” innovative and entrepreneurial talents based on Virtual Simulation Technology (Project No: SJGY20210895).

## References

1. Aliyu, F., Talib, C.A.: VRT. *Asia Proceedings of Social Sci.s* **4**(3), 66-68 (2019)
2. Maples-Keller, J.L., Bunnell, B.E., Kim, S.J., et al.: The use of VRT in the treatment of anxiety and other psychiatric disorders. *Harv. Rev. Psychiatry* **25**(3), 103–113 (2017)
3. Cao, D., Li, G., Zhu, W., et al.: VRT applied in digitalization of cultural heritage. *Clust. Comput.* **22**(4), 1–12 (2017)
4. Hughes, S., Warren-Norton, K., et al.: Supporting optimal aging through the innovative use of VRT. *Multimodal Technol. Interaction* **1**(4), 23 (2017)
5. Atsz, O.: VRT and physical distancing: a review on limiting human interaction in tourism. *J. Multidisciplinary Academic Tourism* **6**(1), 27–35 (2021)
6. Akb, A., Hha, B., Mza, C.: The application of VRT in architectural pedagogy for building constructions. *Alex. Eng. J.* **58**(2), 713–723 (2019)
7. Beetson, S.J., Pradhan, S., Gordon, G., et al.: Building a digital entrepreneurial platform through local community activity and digital skills with ngemba first nation, Australia. *Int. Indigenous Policy J.* **11**(1), 1–19 (2020)
8. Iscaro, V., Castaldi, L., Sepe, E.: ExperimentaLab: a virtual platform to enhance entrepreneurial education through training. *Ind. High. Educ.* **31**(1), 13–22 (2017)
9. Akinbola, O., Ogunnaike, O., Tijani, A.: Micro financing and entrepreneurial development in Nigeria; the mediating role of marketing. *Social Sci. Electronic Publishing* **1**(6), 7–19 (2017)
10. Jifeng, H.: Construction and implementation of a college student entrepreneurial platform in Hebei Province under the integration of Beijing, Tianjin, and Hebei. *Agro Food Industry Hi Tech* **28**(1), 366–370 (2017)
11. Kwon, S.J.: Big hit entertainment’s contents innovation : focusing on entrepreneurial leadership and platform strategy. *J. Korea Entertainment Industry Association* **15**(1), 1–12 (2021)
12. Garais, E.G., Cruau, G.: Entrepreneurial game simulation e-platform for subpioent erasmus plus project. *J. Inf. Syst. Manag.* **13**(1), 179–187 (2019)



# A Model of Diversified Fiscal and Taxation System Based on Computer Network Technology

Jianwei Zhang<sup>(✉)</sup>

Heilongjiang Tourism Vocational and Technical College, Harbin 150006, Heilongjiang, China  
ZJW13313657867@163.com

**Abstract.** With the development of my country's national economy and the increase of fiscal and taxation, the contradiction between traditional business and the rapid growth of paper and business scale as the source of information is growing. The purpose of this paper is to study the diversified fiscal and taxation system model based on computer network technology. Undertake the design and implementation tasks of a diversified fiscal and taxation system, including the analysis of some functional requirements, the design of the various levels of the system, and the implementation and final testing of all functions. The system requirements analysis link starts from mastering the system as a whole, and specifies basic factors such as system roles and business processes; the system outline design link faces the overall design task, and reflects functions, front-end, business logic, technology and other elements in the form of an overall architecture, and selects the Java language. MVC design pattern, Hiberante + SpringBoot framework, Oracle database, etc. as development technologies, and build system deployment environment. In the system test, the overall process of the test is described, and the test of system defects is mainly carried out. The experimental results show that the number of system defects shows a decreasing trend.

**Keywords:** Computer Network · Diversified Fiscal and Taxation · Fiscal and Taxation System · System Model

## 1 Introduction

In order to coordinate and manage the economic market, various government departments communicate and coordinate with each other more and more, which requires a platform to build an efficient, accurate, fast, safe, interactive and information-sharing bridge between departments, and to share their own The information is accurately and quickly released to the sharing platform and awaits the corresponding processing of other departments. At the same time, it also strengthens the supervision of government departments, improves their coordination efficiency and law enforcement capabilities, and provides an effective basis for the law enforcement process [1, 2]. At this stage, the websites of various government departments in my country have been widely opened, but there are many problems in the actual use process [3, 4].

The fiscal and taxation system is such a platform. It has a central server and shares the data of finance and other departments on one platform through the communication

lines of Netcom [5, 6]. Vershinin Y V is designed for the task of analyzing systems for processing financial data. From a business perspective, such a system should solve the problem of analyzing the market basket, i.e. finding the most typical buying patterns. From a data mining point of view, the task of searching for association rules is solved, which consists of two stages: searching all frequent sets with their supporting values and obtaining association rules based on the found sets. The result of this work is the description of a fault-tolerant and scalable model for analytical systems [7]. Gorbatov V S describes a protocol for generating a master key for systems exchanging financial characteristics, the method for generating a financial characteristic key is the authentication of financial characteristic generation and verification devices installed in financial drives and financial data operator equipment, and an authority. The protocol is based on the use of known domestic cryptographic transformations and aims to ensure the integrity and authenticity of data transmitted through the communication channel between the means of formation and the means of verification of financial characteristics [8]. In order to give full play to the basic and guarantee role of fiscal and taxation data in improving government governance capabilities, build a fiscal and taxation data brand, further improve the level of “data auxiliary”, use data to improve management, use data to promote innovation, and carry out the construction of fiscal and taxation data management systems [9, 10].

According to the characteristics of the diversified tax system, this project adopts the standard MVC function in the system architecture, and uses the spring technology + Hibernate + WebLogic to execute the functions of the whole system. The currently popular SpringWeb framework is used as a web-based platform application. The construction of the diversified fiscal and taxation system model in this paper is realized on the basis of these computer network technologies.

## **2 Research on the Model of Diversified Fiscal and Taxation System Based on Computer Network Technology**

### **2.1 Design Requirements for Diversified Fiscal and Taxation Systems**

The diversified fiscal and taxation system system is planned to be deployed in the big data resource management center of M province as a whole. According to the different access objects of the system, access channels are provided through the electronic fiscal and taxation extranet and the Internet [11]. The design requirements are as follows:

- (1) Users such as the public and service institutions served by the system are connected through the Internet [12].
- (2) Ministry and provincial application resource exchange, municipal finance and taxation data resource exchange, data resource exchange of various commissions, offices and bureaus, urban finance and taxation departments, street community management departments, and some directly affiliated institutions are connected through the electronic finance and taxation extranet [13].
- (3) The system accepts the user’s access request, and needs to use the interface to call unified identity authentication, the user center and the API gateway to complete the login operation to access the system [14-15].

## 2.2 Overall Framework

### (1) Model side

In MVC, the model is the functional code, there is no concept in this part of the code that decides how it is presented to the user. A model is a transparent view of the workflow, a set of open paths through which all the functionality of the model can be accessed. In the process of struts, the model is implemented with Action and EJB technology [16].

### (2) View side

In the MVC process, the Model may have multiple views, but in practice most views have the original motivation to use MVC. Using the MVC framework allows multiple Views to exist, and the ability to record the desired view when needed.

### (3) Controller side

MVC views are used with MVC controllers. When the user group is integrated with the corresponding view, the user can update the template status through the viewing window, and this update is made by the control group. The control group changes the position value in a perfect way in the model group. At the same time, the control panel informs all registered views to report the agent to the user [17].

## 2.3 COMMAND Mode

The COMMAND strategy is to make a request from the client to the project without knowing the action resulting from the request or the details of processing it by accepting the request. This is a means of communication between two devices, similar to the CallBack function of traditional procedural languages. This function disconnects the sender and receiver. The sender invokes a service, the receiver receives the request and executes the corresponding service, the sender does not need to know any view of the receiver. Complies with specific packages and minimizes system coupling.

## 2.4 XML

The main purpose of Extensible Markup Language is to overcome the weaknesses of HTML. The W3C is responsible for the development and maintenance of many web applications, especially the hyperlinking language [18].

To accommodate web applications, people extended HTML and introduced more tags. Over time, HTML has become a rich language with nearly 100 tags, and the combinations of these tags are endless, and the same icon combination will produce different results in different browsers.

While many indicators are already in place, one will find that more indicators are needed. For example, e-commerce applications require labeling techniques for product descriptions, pricing, names, addresses, and more. Image tags that control image and audio traffic - search engine keywords and descriptions require attribute tags - security systems, and also information about digital signatures, which is why XML was developed.

### 3 Investigation and Research on Diversified Fiscal and Taxation System Model Based on Computer Network Technology

#### 3.1 Use Case Model

(1) Taxation subsystem

The taxation subsystem collects and manages data from diverse fiscal and taxation systems, generates XML documents, and sends them to the financial system via FTP. The data is stored in the XML file upload management system.

(2) Treasury subsystem

The Treasury subsystem lets business owners perform tasks such as sales, public certification, and end-to-end transactions. The bank payment information of the sales number is obtained from the same-city withholding system, and the taxation information of the sales number is collected from the tax subsystem.

#### 3.2 File Transfer

After the XML is generated, the diversified financial and taxation system uses FTP to transfer files; the information exchange processing system acts as an FTP server.

(1) Download the diversified taxation system

Notify network devices with a warning message when files are ready to be downloaded from each network device. Once each network driver is notified, it sends a feedback message that it has been notified, and then logs the XML file itself as an FTP client.

(2) Upload the networking unit system

Download the XML file to be exported to the FTP server of the Diversified Finance and Taxation System, and then send the specified file transfer notification group to the Diversified Finance and Taxation System. Immediately after receiving the tax, it publishes a feedback report.

#### 3.3 Algorithm Description

The system data processing mainly includes offline calculation and real-time calculation, among which this program firstly adopts offline performance model and real-time performance model. Offline working model with cell-based statistics and map statistics distribution and related data mining services. The task manager invokes corresponding services according to different workloads to perform consistent calculations on the specified data sources.

In text, the cosine measurement method is more efficient. For implementation in Mahout, the calculation types are:

$$\cos(X, Y) = \frac{(X \cdot Y)}{\|X\| \|Y\|} \quad (1)$$



In the recommendation,  $r_{ui}$  is used to represent the User's score on Item  $i$ , and the  $u$  vector score on each Object can be used as User's hobby. Next, the formula for calculating the cosine similarity between User  $u$  and User  $v$  is:

$$\cos(u, v) = \frac{\sum_{i \in I_{uv}} r_{ui} r_{vi}}{\sqrt{\sum_{i \in I_u} r_{ui}^2 \sum_{i \in I_v} r_{vi}^2}} \quad (2)$$

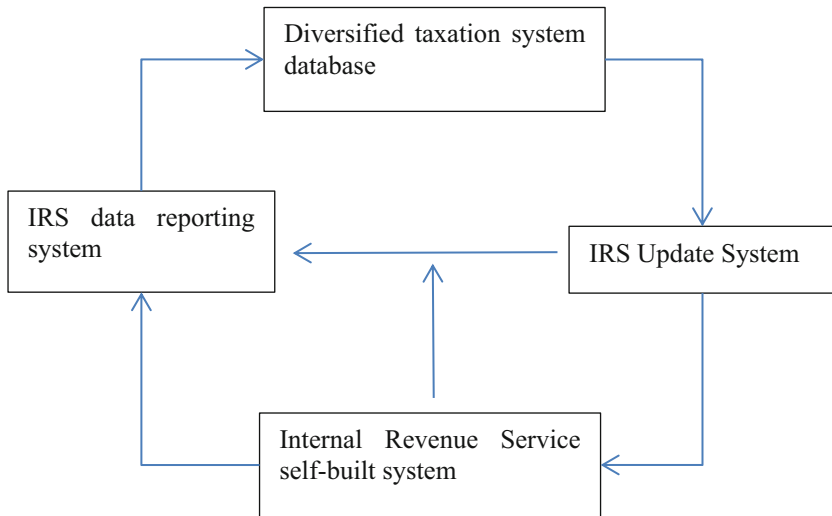
### 3.4 Test Plan

The system uses SVN to manage system performance during development and IBM Rational ClearQuest to ensure the quality of testing and results. The system test system is divided into performance test, final test and entry test. Tests include visible tests, test results, test results, standardized tests, and screening tests. At the same time, the system undergoes development-based testing, focusing on the results of the system, modeling, testing, and the use of verification testing; in addition, group comparison testing and combination testing are established to see if there are any changes that will not affect other areas.

## 4 Investigation and Analysis of Diversified Fiscal and Taxation System Model Based on Computer Network Technology

### 4.1 Functional Structure Design

The main function of the diversified financial and taxation system is to publish the basic data provided by the industrial and commercial enterprises under the jurisdiction of the foreign economic and trade bureau, economic and trade bureau, project office, development zone and other departments to the sharing platform of the diversified financial and taxation system. Some basic attributes of enterprise data will be On the sharing platform of the diversified fiscal and taxation system, the review and approval information based on the data provided by the Bureau of Industry and Commerce, Foreign Trade and Economic Cooperation, the Economic and Trade Bureau, the Project Office, and the Development Zone will be published on the diversified fiscal and taxation system. For review and approval, the foreign economic and trade bureau, economic and trade bureau, project office, development zone and other departments can modify some daily conditions of the enterprises under their jurisdiction. According to the data on the Finance and Taxation Online, the finance department provides the core financial statements based on the corresponding data for the superior leaders to query and make decisions. The functional modules of the national tax department are shown in Fig. 1:



**Fig. 1.** Functional modules of the national tax department

On this platform, different roles have different permissions and functions. It clarifies the responsibilities and scope of rights of each department, ensures the security and accuracy of data, and realizes interaction and sharing of data.

**4.2 System Test Results and Analysis**

Test reports show that the diversified fiscal system has met the proposed requirements and is functioning properly. Manage defects with IBM Rational ClearQuest. Therefore, the emergence and repair of defects can be well applied. The statistics of the identified deficiencies of the diversified fiscal and taxation systems are shown in Table 1.

**Table 1.** Statistics of flaws found in fiscal and taxation systems

System Model Function Module	Day one	Day two	Day three
Batch processing	7	6	10
Report generation function	5	3	6
Other functions	2	0	1

As testing progresses, the number of defects found in system testing decreases, as shown in Fig. 2. Dev-Test can control risks from changing system requirements, ensure compliance including design requirements, and reduce post-test entry costs.

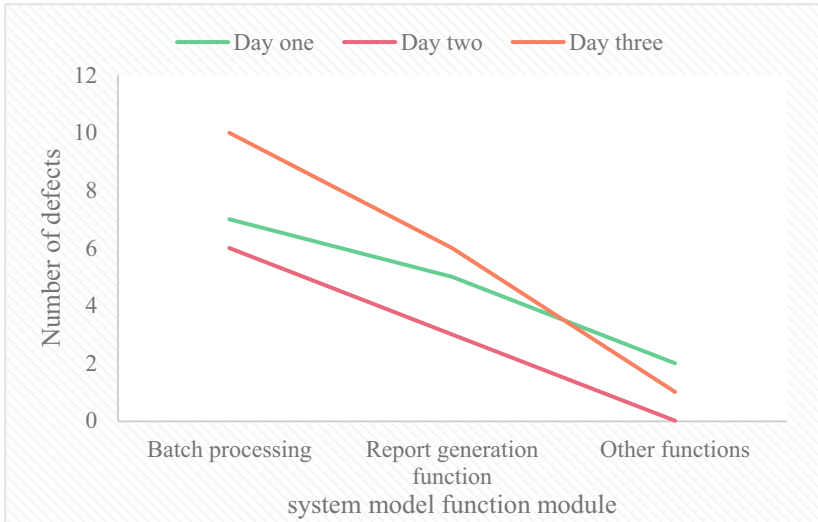


Fig. 2. System Test Results

For system model applications such as batch processing and reporting, in the case of a large amount of data, the proposed performance requirements may not be fully met. Therefore, it is necessary to further improve the code and make full use of the characteristics of commercial databases to improve performance. Improved index and layout settings to avoid table views and views, reducing data collection costs and improving overall performance. Optimization algorithms, such as optimizing state control, reducing loop levels, reducing calls to data services, and reducing execution costs. SQL details, try to complete the required business functions in one statement to reduce the data execution time.

## 5 Conclusions

With the development of my country's e-government level, the demand for information exchange of national financial services such as national taxation is becoming more and more intense. However, due to some technical reasons and previous policy issues, the compatibility and integration of the software are also weak. How to see the distribution of software information on the basis of existing software storage has become a real problem faced by the national tax department. The system designed in this paper aims to improve the level of informatization of relevant units. Through various computer network technologies, information resources can be shared among the departments of finance, industry and commerce, national taxation, local taxation, foreign economic and trade bureau, economic and trade bureau, project office, development zone, etc. It also strengthens the supervision of government departments, improves their coordination efficiency and law enforcement capabilities, provides an effective basis for the law enforcement process, and makes management more efficient, standardized and scientific.

## References

1. Ermekbayeva, B.G., Alimov, N.D.: The role of tax potential in improving the quality of the fiscal system. *REPORTS* **2**(330), 125–130 (2020)
2. Golubovi, S.: The role of the european fiscal board in the EU fiscal system. *Zbornik radova Pravnog fakulteta Nis* **58**(84), 15–30 (2019)
3. Bovsh, L., Okhrimenko, A., Boiko, M., et al.: Tourist tax administration in the fiscal target system for hospitality businesses. *Public and Municipal Finance* **10**(1), 1–11 (2021)
4. Chugunov, I., Makohon, V., Vatulov, A., et al.: General government revenue in the system of fiscal regulation. *Investment Manage.Financial Innovations* **17**(1), 134–142 (2020)
5. Kostenkova, V.G.: Formation of an effective system of fiscal federalism is a strategic factor in the development of the russian economy. *Bulletin of Udmurt University Series Economics and Law* **30**(5), 639–646 (2020)
6. Rao, M.G.: Redesigning the fiscal transfer system in India. *Economic and Political Weekly* **54**(31), 52–61 (2019)
7. Vershinin, Y.V., Prokofyev, M.L., Afanasyev, V.R.: Developing fiscal data processing analytical system. *Issues of Radio Elect.* **3**, 78–82 (2019)
8. Gorbatov, V.S., Zhukov, I.Y., Murashov, O.N.: Security of a key system of a fiscal feature. *Autom. Control. Comput. Sci.* **52**(8), 1065–1070 (2018)
9. Nasrullah, F., Amin, R.U., Soomro, K.A.: Decentralization of education: the affective role of fiscal decentralization in Pakistani education system(s). *J. Social Sci.* **15**(1), 41–51 (2021)
10. Ivkovi, A., Panti, N., Rosi, M.: Fiscal sustainability of the macroeconomic system of European Union members. *Oditor - casopis za Menadzmt finansije i pravo* **5**(2), 32–41 (2019)
11. Kozakevych, O.: Features of fiscal policy in the system of levers of stimulation of the national economy of Ukraine. *Path of Science* **4**(7), 1001–1006 (2018)
12. Kooij, J., Groot, T.: Towards a comprehensive assessment system of local government fiscal health. *Maandblad Voor Accountancy en Bedrijfseconomie* **95**(7/8), 233–244 (2021)
13. Muinelo-Gallo, L., Bordon, J.U., Scavone, P.C.: Fiscal disparities in Uruguay's regions: the role of a new system of intergovernmental equalization transfers. *CEPAL Rev.* **2019**(128), 197–220 (2020)
14. Bradbury, D., O'Reilly, P.: Inclusive fiscal reform: ensuring fairness and transparency in the international tax system. *Int. Tax Public Financ.* **25**(6), 1434–1448 (2018)
15. Chung, I.H.: Evaluating fiscal crisis management for local governments as early warning system. *Korean Policy Sciences Rev.* **22**(4), 1–21 (2018)
16. Trstena, A., Mehmeti, I., Krasniqi, S.: Fiscal policy and foreign trade as a determinant of economic development – focus Kosovo. *Int. J. Finance & Banking Studies* (2147–4486), **8**(3), 77–87 (2019)
17. Lee, J.S.: The impact of fiscal conservatism ; a case study of the 2007 national pension system reform in Korea. *Int. Rev. Financial Consumers* **4**(No. 1 Apr 2019), 17–38 (2019)
18. Ocran, D., Broni-Bediako, E., Ofori-Sarpong, G.: Boundary applicability of the ghana's oil block fiscal regimes. *Ghana Mining J.* **19**(2), 70–76 (2019)



# Research on Smart City Development and Construction Driven by Big Data Technology

Yongling Chu<sup>(✉)</sup>, Zhiling Luan, and Shaochun Li

Yantai Vocational College, Yantai 264670, Shandong, China  
394145965@qq.com

**Abstract.** Building a smart city is not only a need of realizing sustainable urban development, but also the strategic choice of improving China's comprehensive competitiveness. The technology of Big Data can supply strong data support for the smart city's construction, and is of great significance to development, construction and management of smart city. This paper points out the existing problems and deficiencies in building smart city's process, and puts forward the framework of smart city management platform on account of the technology of big data, thus maximize the value of big data, push forward the building and growth of smart city.

**Keywords:** Big Data Technology · Smart City · New Generation Information Technology · Sustainable Development

## 1 Preface

The quick growth of message technique makes the construction of smart city possible, and the technology of big data's effective use can effectively push forward the construction and growth of smart city [1]. Countries and governments all over the world have put forward the design concept of smart Earth, smart community and smart city one after another, and use technology to promote their realization. The government of Chinese also pays high attention to the growth of smart cities. From the 100 pilot projects launched during the "Fifteen" period to the present, smart cities have basically covered all major cities of china. The building and growth of smart cities have greatly improved the government's ability to provide government services and communal services, and raised the humanity's index of well-being [2-5].

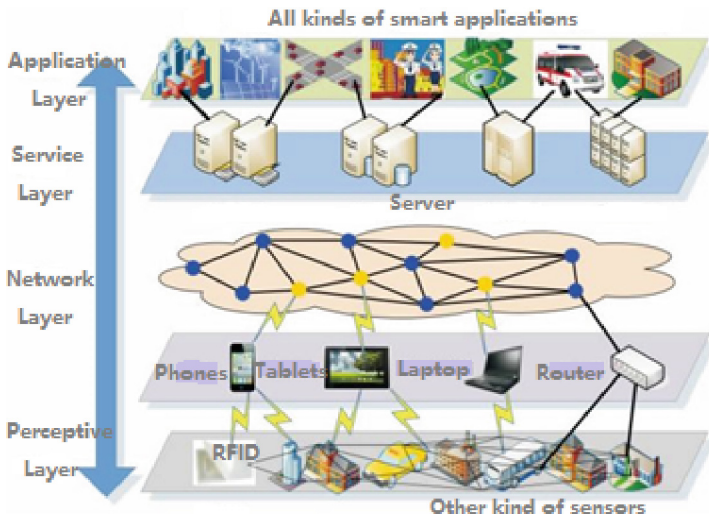
## 2 Related Concepts

### 2.1 Smart City

The conception of a smart city was formally put forward by IBM companies in 2010. Smart City is established on the basic frame of numerical town, which makes full use of

the new technique of IOT, 5G, AI and so on to push forward the construction and growth of city, to realize the depth fusion of message, industrialisation and townization [6].

The building of smart city is mainly restricted by two key factors [7, 8]: one is technology, mainly the growth of new generation message technique among which the most representative is mobile meshwork, IOT, big data and other new techniques; Second, the social factors, including the construction of infrastructure and the social environment of the open degree of innovation. The overall framework of a smart city is seen in Fig. 1:



**Fig. 1.** The general framework of smart city

## 2.2 The Technology of Big Data

Big Data is large-scale and many kinds of massive data, which can not be collected and analyzed by human in a short time, and needs computer technology to collect, clean, store, excavate and process [9, 10]. The Technology of Big Data is the core technology of the recent generation of information technology, the use of big data technique can promote the speed of smart city's building, improve the efficiency of smart city's growth [11]. The whole process of data mining with big data is seen in Fig. 2:

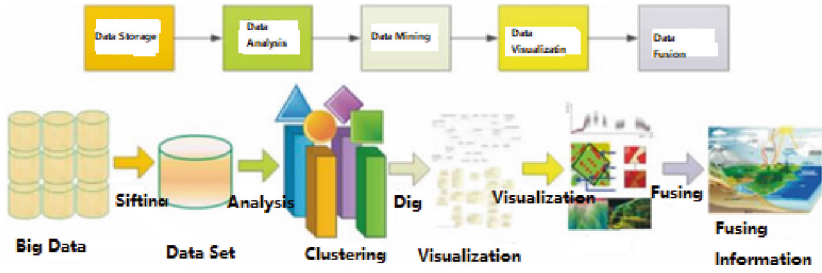


Fig. 2. The whole process of data digging

For the massive data obtained in the building of smart city, it can not only obtain useful message by means of data digging, improve the utilization rate of information, but also provide strong data support for the policy decision of government departments. The apply of big data technology in the building of smart city can accurately identify the target and promote the rapid building and growth of smart city. Figure 3 shows the application of target recognition algorithms in big data technology in different situations [12].



Fig. 3. Multiple applications of target recognition technology

### 3 Problems and Deficiencies in the Building of Smart Cities

The Technology of Big Data carry a big weight in the building of smart city, but because of its complexity and vastness, it brings some difficulties and challenges for the extraction and classification of effective information. At present, the major questions in the building of smart city are as follows [13].

#### 3.1 Information Islands

The construction and operation of smart cities have their own systems in different departments, and each system produces a large amount of data, which can be shared among departments, however, the data between different departments is difficult to achieve effective convergence and integration, resulting in a number of information islands, as shown in Fig. 4. The phenomenon of Information Island makes it impossible to communicate effectively among different departments, which hinders the efficient building and growth of smart city. The deep building and growth of smart city need to eliminate the information island and achieve the high fusion and data's share [14].

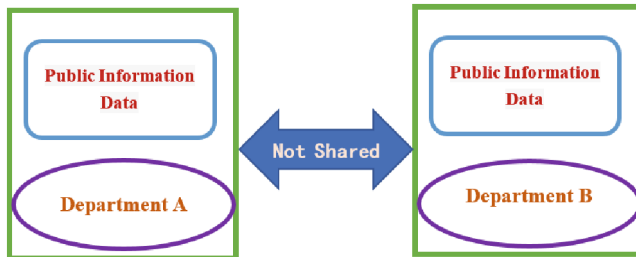


Fig. 4. Information Silo

#### 3.2 Information Security of Data

In pace with the application of big data technique, the construction of smart city has been further developed. While we enjoy the convenience of big data and smart city, the information data's security becomes more important. The relatively safe data in intranet will be attacked by network when it is transmitted between different networks, which will result in many security problems such as information loss, data leakage, network disruption and so on. Big Data Technology is a dual blades knife. It not only promotes the construction of smart city and benefits the society and people, but also has the risk factor of security leak. This is the chief question to be settled in the building and development of smart city propelled by the technology of big data [15].



## **4 Development and Building of Smart City Driven by Big Data Technology**

In pace with the further development of smart city, the technique of big data has been integrated into all levels of smart city's growth. In view of the questions in the building and development of smart city, it is necessary to strengthen the communication and contact among government departments, enterprises and trade associations, break the information island, integrate data resources and realize data sharing. At the same time, create a security system, fortify the training of big data's talents, truly achieve the goal of "Three-way win-win", in the drive of big data's technique to promote the further construction and growth of smart city.

### **4.1 Smart City Management Platform Driven by Big Data Technology**

Building a smart city management platform driven by big data technology to realize information resource sharing. Using the IOT, computing based on cloud, and big data techniques to integrate and integrate the data of people, cars, and objects in cities, break down barriers between government and enterprises, eliminate information islands, and finally display it through platforms, for the government, enterprises and individuals to make decisions. Through the building of Smart City Management Platform, push forward the sustainable growth of smart city.

### **4.2 Smart City Management Platform Architecture Based on Big Data**

The overall architecture of smart city management platform in view of big data chiefly contains infrastructure layer, record layer, sustain layer and application layer, as shown in Fig. 5. Among them, the basic facilities layer is the ensurance of security, through the establishment of a comprehensive infrastructure for the building of smart cities to lay a good foundation. The data layer is to collect, process and integrate all the data in the smart city building's process, eliminate message islands and realize data sharing. The support layer is situated between the data layer and the application layer, and provides support services for the application layer. The utilization layer is the actual application layer of each portal website and each service platform. The government, the industry enterprise, the individual uses the integrated data to carry on the decision-making, has improved the work efficiency, has promoted the service quality [16].

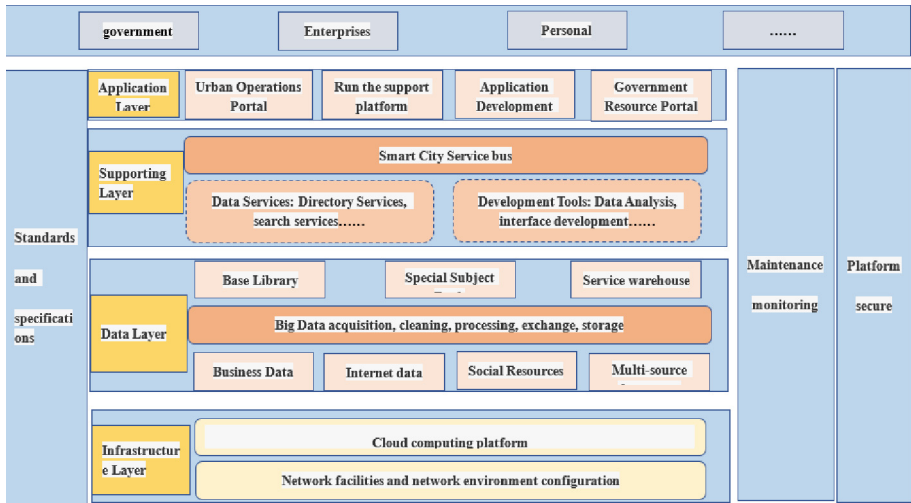


Fig. 5. The overall architecture of smart city management platform in view of big data

## 5 Conclusions

The level of informationization and intellectualization of a city has become one of the important indicators to measure a country’s Comprehensive National Power, degree of modernization, international competitiveness and economic growth potential. The Big Data’s technique is not only the cornerstone of smart city’s building, but also the source of driving the development of urban digital economy. The construction of smart city propelled by big data technology can integrate the information resources of different departments, achieve data’s share and mining, and push forward the decision-making ability and work efficiency of government departments.

**Acknowledgements.** 2021 Shandong Institute of Higher Education Special Topics in Higher Education Research “The Construction of the Monitoring and Evaluation Model of Higher Vocational Teaching Quality Based on Big Data under the Background of “Double Higher Education” (CX-0006)”.

2022 Shandong Province Key topics of culture and art “Research on the development of regional animation industry cluster supported by VR technology in Shandong animation innovation education”.

2022 Annual Project of Yantai Vocational Education “Research on Vocational Education Digital Transformation and Intelligent Upgrading of ‘Industrial Internet plus’ ”.

## References

1. Han, H.: Application of Deep Learning for Level of Engagement in Civic Activities Prediction in Emerging Adulthood for Smart City Development (2022)
2. Wang, H., Li, W., Li, M., et al.: Intelligent selection of delivery parties for fresh agricultural product based on third-party logistics in smart city. Sustainable energy Technol. Assessments (Aug. Pt.B), 52 (2022)

3. Roshan, R., Rishi, O.P.: Design and development of multi-objective hybrid clustering framework for smart city in India using internet of things. *J. Information & Knowledge Manage.* 2250064 (2022)
4. Maksimchuk, O.V., Borisova, N.I., Panov, D.B., et al.: Sustainable development of road transport infrastructure by including elements of reverse traffic in the “smart city” program. In: *IOP Conference Series: Earth and Environmental Science* **988**(2), 022072 (2022)
5. Idrees, S.M.: A fog computing model for VANET to reduce latency and delay using 5G network in smart city transportation. *Applied Sciences* **12**(4), 2083 (2022)
6. Popova, Y., Zagulova, D.: UTAUT Model for Smart City Concept Implementation: Use of Web Applications by Residents for Everyday Operations (2022)
7. Jnr, B.A., Atkinson, R.: Exploring Data Driven Initiatives for Smart City Development: Empirical Evidence from Techno-Stakeholders’ Perspective (2022)
8. Bykov, N.V.: Impact of Counteracting Vehicles on the Characteristics of a Smart City Transport System (2022)
9. Kosti, Z., Angus, A., Yang, Z., et al.: Smart City Intersections: Intelligence Nodes for Future Metropolises. *arXiv e-prints* (2022)
10. Vaz, A., Ferreira, F., Pereira, L.F., et al.: Strategic visualization: the (real) usefulness of cognitive mapping in smart city conceptualization. *Manage. Decision* **60**(4), 916939 (2022)
11. Al-Dabbagh, R.: Dubai, the sustainable, smart city. *Renewable Energy and Environ. Sustainability* **7**, 3 (2022)
12. Putra, K.T., Chen, H.C., Prayitno, et al.: Federated compressed learning edge computing framework with ensuring data privacy for PM2.5 prediction in smart city sensing applications. *Sensors* **21**(13), 4586 (2021)
13. Du, M., Zhang, X., Mora, L., et al.: Strategic Planning for Smart City Development: Assessing Spatial Inequalities in the Basic Service Provision of Metropolitan Cities (2021)
14. Kitchin, R.: Aaron Shapiro 2020: design, control, predict: logistical governance in the smart city. Minnesota University Press, Minneapolis, MN. *Int. J. Urban and Regional Res.* **45**(6), 10811082 (2021)
15. Lenkoe, L., Kotze, B., Veldtsman, P.: Simulation of hexagon spatial image datasets for free motion in a simulator for smart city bidirectional navigation purposes. *VISUAL 2021, The Sixth International Conference on Applications and Systems of Visual Paradigms* (2021)
16. Peng, S.L.: Smart and future applications of internet of multimedia things (IoMT) using big data analytics. *Sensors* **22** (2022)



# Discipline Inspection and Supervision Big Data Information Query Platform Based on Distributed Computing

Changyu Liu<sup>1</sup>(✉), Jinhua Chen<sup>2</sup>, Hongchang Wen<sup>2</sup>, Fanjun Zeng<sup>2</sup>,  
and Sanjiong Wang<sup>2</sup>

<sup>1</sup> State Grid Corporation of China, Beijing 100031, China  
xslw2022@126.com

<sup>2</sup> Beijing Guodiantong Network Technology Co., Ltd., Beijing 100070, China

**Abstract.** The DIAS BD information inquiry platform uses scientific information technology to create a stable cadre archive, reference library and database. Carry out discipline inspection and supervision(DIAS) activities at all levels, adjust the reporting and filing process, and solve the problems of repeated input, repeated calculation, slow process and being unable to keep up with the current DIAS work. It can be divided into different parts according to different business needs and network areas, basically according to the needs of different departments and different management levels. With the strong development of the Internet, it is necessary to combine DIAS with big data(BD) information query platform. This paper designs and analyzes the DIAS BD information query platform based on distributed computing(DC), and constructs an efficient data access model and a perfect security system for the DIAS information platform, which can ensure the data reliability of the system.

**Keywords:** Distributed Computing · Discipline Inspection and Supervision · Big Data · Information Query Platform

## 1 Introduction

With the rapid development of information technology, the relevant government departments of our country have also carried out the construction of government affairs platform in succession, in order to promote the information construction of relevant government departments in order to improve work efficiency, including the DIAS Bureau. DIAS organs at all levels will produce a lot of data on the construction of a clean and honest government and DIAS in their actual work. To this end, this paper builds a DIAS BD information query platform based on DC.

Promoting the informatization construction of DIAS business has become the future development trend. In this paper, the DIAS BD information query platform is built based on DC to support system users to process related businesses when they can connect to the DIAS information platform web server. According to the actual application, the access methods of intranet users and extranet users are different. The extranet has better

requirements for network security. Intrusion data can be filtered through routers and hardware firewalls, Ensure that the external network environment is relatively safe. The data of the intranet belongs to the sensitive data of government affairs. The hot standby scheme of the data is established through the disk array to ensure that when one system is abnormal, the other system continues to work [1].

This paper analyzes the needs of the DIAS information platform: the needs analysis stage is the preliminary work of the design of the system. Through the investigation, the business content is transformed into the needs target, so as to further determine the function, performance and other contents of the system to be designed. The DIAS BD supervision information query platform constructed by this topic is composed of security protection business architecture, data architecture, technical architecture and other parts; Design and implementation of DIAS information platform: scientifically describe the business matters of the main functional modules of the system through the application of sequence diagram, simplify the system development process, and take this as a reference to provide a basis for the system implementation [2, 3].

## 2 DIAS Based on DC

### 2.1 DC

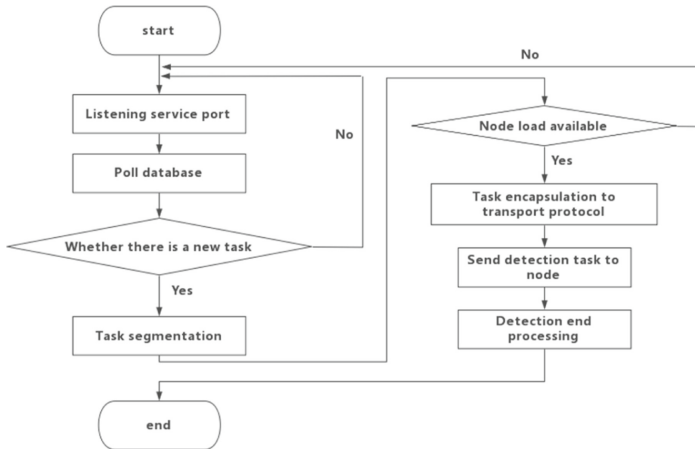
#### (1) implementation of distributed platform system

The distributed platform system is implemented by Python language in Linux environment. The distributed platform communication is based on twisted network development framework, and the database management software adopts mysql. It realizes several functional modules: detection task separation, task scheduling, task distribution, detection result processing, process scheduling, log management, and data storage.

The Config class of the master side is responsible for reading the configuration file information of the master side, including the given service port number, listening port number, database connection user, password, detection task type identification definition and other information. It has certain benefits for infrequent data manipulation and system expansion. Dnstypeparse class and dnspark class respectively realize the IP task list segmentation of DNS distribution detection and the domain name list segmentation of DNS server configuration detection. The schedule class implements the task scheduling function, selects the unfinished tasks in the database, and arranges the task distribution according to the load information on the slave end. Process implements the process scheduling function, including the methods of creating, suspending, waking up, and killing processes. The logger class implements the log management function, and the log information includes system errors and operation conditions [4].

The detection tasks and configuration files of the master end need to be submitted manually. After the configuration is completed, the master end starts the service and starts listening to the service port. The whole task scheduling process can be realized only when there are unfinished detection tasks and online slave end can be scheduled.

Implementation of detection task scheduling: detection task scheduling is mainly task list reading. According to the hardware load information fed back by the slave end, decision-making task distribution achieves the load balance of the slave end. The task scheduling process is shown in Fig. 1.



**Fig. 1.** Master task scheduling flow chart

The first step of task scheduling is task separation, which is implemented in jobparse file. First, it is sorted in descending order according to task priority, in which the priority is divided into four levels, identified by four numbers of 0–3. Read the location information of the task file from the task list to be assigned with the highest priority, find the target task file, take the domain name task separation as an example, read the domain name information line by line from the task file, and write the information into the task table to be assigned [5, 6]. The task distribution part is implemented in the schedule. The task distribution first calculates the load information of the online slave node. The slave node is arranged in ascending order according to the load coefficient, and sends the task to the slave node.

- (2) Instruction issuance: This module calls some methods of sys, argparse, datetime, socket and structmaster classes in Python class library. The instructions are in the form of a command line, where “-l” represents the listing of online slave nodes, “-s” represents the action of sending instructions, “-j” represents the task list sent, and the task file path is specified after the parameter.
- (3) Data processing: the master end is responsible for processing the data information uploaded by the slave end. The effective data mainly includes the hardware load information sent by the slave node regularly and the DNS information detection data result set. DNS information detection results are formatted and stored. Here, the information category is distinguished by the first byte data of the packet. The value of the first byte of the packet “R” represents the return detection result, and the value “H” represents the load data of the slave node.
- (4) Log management: value 1 is marked as system operation error, and value 2 is marked as system operation record. Check the whole system running process by calling the

logging class library integrated by python. The generated log information is written to the error and application log files respectively according to the information type.

- (5) Data storage: pymysql realizes the function of data addition, deletion, modification and query. This module encapsulates the database operation in pymysql class, and provides the contents of query statements to manipulate the database.
- (6) Network topology cloud DC mode

In the network topology of this model,  $t$  is the end cloud collaboration task coordinator, each computer node is connected to the end cloud collaboration coordinator through the network link  $L$ , and the collaboration cloud coordinator  $D$  is responsible for the division of the set. Assuming that  $t_i$  is the time when the current task of each computing node is completed, the calculation formula is as follows:

$$T = \min\{\max\{T_i, 1 \leq i \leq z\}\} \quad (1)$$

Set the end time  $t_i$  for completing the calculation task, which is equal to the time point  $t_{i-1}$  after  $I-1$  tasks before task  $I$  are all completed and task distribution plus task processing time.  $T_i$  is expressed as the following formula:

According to the above formula, we can then calculate the size of each  $RI$ , and finally get the solution formula of task  $I$  of each node as follows:

## 2.2 Demand Analysis of BD Information Query System of DIAS Bureau

- (1) Business requirements

The purpose of the DIAS Bureau to change the existing management mode and promote the informatization construction of DIAS business is to improve efficiency and make the DIAS business standardized, scientific and networked. On the DIAS information platform, there are interactive platforms and reporting platforms, which can make related businesses more efficient [7, 8]. Through the understanding of the discipline inspection work and the research on the realization of the objectives of the system, it is analyzed and determined that the system should have the following business operations:

The system needs to provide an online interactive platform so that people can consult online in order to get a reply in time, which helps to reduce the number of illegal petitions; The system needs to support online reporting and simplify the reporting process. The whistleblower only needs to submit the reporting information in the system, and the staff of the DIAS Bureau can conduct an investigation according to the materials after receiving the report; It has a video broadcasting mechanism to enrich the integrity education methods of the DIAS Bureau, so that civil servants can learn more integrity knowledge in the information platform, and the learning methods are diversified, which can enhance the learning interest of civil servants [9]. It should have all operation functions with clear authority attribution, and ensure that system users perform their responsibilities through authority management, which indirectly ensures the normal and orderly operation of the system; The data of the system is relatively sensitive, especially some reports from the public. The protection of this information needs to be completed by means of data backup and recovery provided by the system; The system should ensure that users

can easily operate various functions, and the operation of various functions should be based on specific business content.

(2) construction of BD information query platform

At present, the architecture method of application system is developing towards loose coupling and distribution. SOA and WOA integrate the characteristics and advantages of these two aspects and absorb the advantages of both aspects, which has become a popular technology of current application system architecture. The specific content consists of five parts:

Foundation support system. The foundation support system is the foundation for the normal operation of a system. Including network facilities for government affairs, hardware platforms composed of various servers and switches, tapes and disks for storing and backing up data, etc. Since the project relies on the government extranet of the Discipline Inspection Commission, these basic software do not need to be re selected [10, 11].

Provincial law enforcement center database. The data storage and management center of the system adopts the database of the provincial law enforcement center. The law enforcement data of all administrative departments in the province are centrally stored here. It involves system database, law enforcement database and discretionary database.

Electronic supervision system for Discipline Inspection. The construction basis of discipline inspection electronic supervision system is tykybos Jee middleware external extension and function design, mainly to complete the DIAS information collection of administrative law enforcement departments, DIAS of administrative law enforcement departments and other functions.

Information portal. Based on the online public portal platform, the public can query the disclosure of various DIAS information through this platform, and functional departments and leaders at all levels can also implement office work, statistics and decision-making through this platform.

Safety assurance system. Based on information security technology guarantee means and guided by the issuance and implementation of security management provisions, ensure the safe and effective operation of the application system in the current interconnected network environment [12].

### **3 General Framework of DIAS BD Information Query System based on DC**

#### **3.1 Project Overview**

With the in-depth promotion of the company's "digital state grid" strategy, enterprise informatization is entering the era of cloud storage and BD. All kinds of businesses are basically handled online. Massive data records the company's reform and development process, which not only poses challenges for DIAS, but also brings new opportunities. On the one hand, DIAS personnel are constantly in contact with the business systems of the State Grid in the process of daily office work and case handling, and have a high level of information application and operation; On the other hand, under the requirements of



confidentiality and rapid response, the requirements for rapid preliminary verification of problem clues are becoming higher and higher. There is an urgent need to achieve efficient preliminary verification of problem clues to avoid the loss of invalid problem clues to DIAS. DIAS work keep pace with the times, make full use of digital information technology, build an internal inspection platform, and query the commonly used data of each business system within a controllable range, which can effectively enrich the methods of supervision and discipline enforcement, and escort the healthy, stable and sustainable development of the company.

### 3.2 Architecture Compliance

#### (1) business structure

Sort out the business structure of the system in this period according to the construction objectives of this period, which is mainly divided into two parts: comprehensive management and DIAS information query business. Among them, integrated management is a service-oriented integrated management to support information query business, mainly including application query management and knowledge management; The inquiry of DIAS information can be divided into 12 business directions according to needs, such as the verification of system and standard information, the verification of important decision-making information, and the verification of internal and external enterprise subject information is shown in Fig. 2.

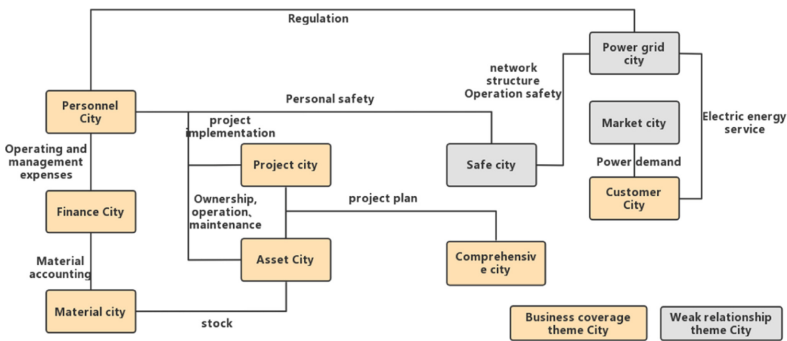


Fig. 2. Schematic diagram of business structure

#### (2) data architecture

The overall data of the system in this period follows the architecture design of the internal inspection platform, and the technical architecture follows the unified technical requirements of the State Grid. Business data mainly involves personnel field, financial field, material field, project field, asset field, comprehensive field and customer field.

The data of the internal inspection platform mainly involves three data topics: organization management, personnel management and training development in the personnel field; In the financial field, there are two data topics: fiscal and tax management and cost management; Purchase contract and purchase demand fulfillment are two data topics in the material field; 1 data subject of the project plan in the project field; One data subject of asset information in the asset field; One data subject of asset information in the asset field; Three data topics in the integrated domain: laws and regulations, external data, and unstructured data; In the customer domain, there are four data topics: customer public, integrated energy services, measurement, and customer service.

(3) database demand analysis

The users of the DIAS system of a large state-owned enterprise are system administrators, ordinary users of the system and users of the investigation team. Different user roles have different functional requirements for the system in the system, which is mainly reflected in the different functions of adding, modifying, deleting, updating and querying data in the underlying database of the system. Therefore, it is necessary to analyze the database requirements of various users, which is the guarantee for the smooth construction of the database in the next step. According to the above demand analysis results, to sum up, the data required by the database of the DIAS system of a large state-owned enterprise mainly include:

User information table: record the user's account number, password, and gender, affiliated unit, and system permission information extracted as personal information settings. Since the system users are the staff responsible for the DIAS of the entire large state-owned enterprises, in order to ensure that they are seriously responsible for the investigation results, the user information also needs to enter the user's real identity information: ID card number. However, because the ID card number is long, it is not easy to remember and apply, so each user can be set with a unique number, which is different from other users, And take this number as the primary key of the user information table.

File information table: the file information table is used to record the data information of each violation of laws and disciplines registered by ordinary users in the system. The main fields of the document information table include document number, registration personnel number, nature of the report (anonymous or real name), entry date, name of the informant, contact information of the informant, Department of the informant, position of the informant, verification code, political appearance, gender, Department, position, disciplinary violations and other categories of the informant. Among them, the document number is the primary key of the document information table.

(4) technical architecture

In the implementation of the technical architecture, it follows the technical design specifications of the State Grid, adopts software component-based and dynamic technology, adopts a consistent data model, and according to the performance layer, service layer, technical support layer, storage layer and infrastructure layer. Components are integrated and reused in the company's internal collaborative work at all levels to meet the company's different business needs, provide

efficient and flexible business support for each work unit, and provide technologically advanced service platforms and simplified business planning capabilities for internal businesses and managers. The technical architecture is shown in Fig. 3:

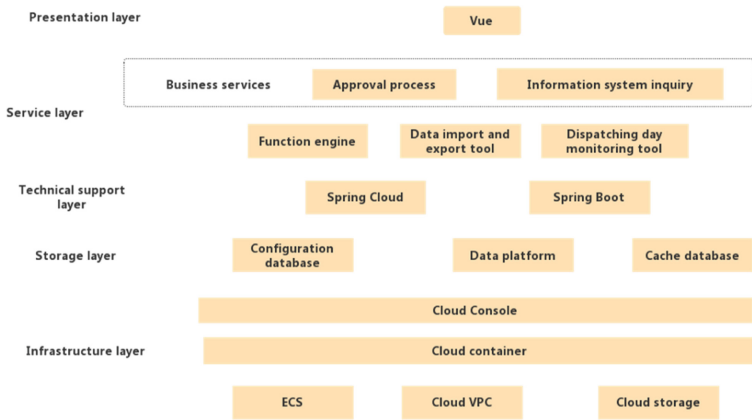


Fig. 3. Schematic diagram of technical architecture

### 3.3 Technical Architecture Design

This system is developed based on J2EE platform. The architecture of this system is divided into four levels, namely, customer level, page level, business level and EIS level. Each level undertakes different tasks in the DIAS information platform.

**Customer layer:** this layer is for direct users. Users of this system will use the browser as an operation tool to analyze functions through the browser. Customers can use the corresponding specific operations. The client represented by the browser conforms to the b/s structure model, and the interface layout through the browser will be applied to specific technologies such as client script.

**Page layer:** users' business requests can be directly accepted through the page layer. The page layer uses struts model, which makes the specific operation page and program code exist independently. This mechanism makes the code structure easier to adjust, and there will be no relevant program code in the client, which indirectly ensures the interface security of the system.

**Business layer:** it provides the level of the main business operations of the system, establishes the main business methods of each functional module, and applies the object-oriented programming idea to write the polymorphic behavior of the method. The Java code in the business layer will execute various business requests issued by the page layer. Each business in this layer has a certain degree of logic. Logic is the premise of good or bad business execution, and the creation of business logic is closely related to the function of the system.

EIS layer: this layer stores the data information involved in the DIAS information platform. The system will adopt SQLServer2005 database, which is mainly characterized by relational data storage.

logical structure design

The system functions are divided into two levels according to the main function modules and sub function modules. According to the modular function division idea, the system mainly includes five function modules and twenty-six sub function modules. The DIAS information platform needs to provide the following basic functions:

Safety protection: ensure the effective operation of various functions of the system, ensure the relative safety of system data, ensure the smooth transmission of system data, and ensure the clear authority of system users.

Discipline inspection news management: the discipline inspection news management module mainly manages the business related to discipline inspection news, so as to show the rich discipline inspection news to civil servants or the people. It consists of news category management, news category management, news management, website search and other functions.

Supervision and reporting: the main function of the supervision and reporting module is to handle the supervision and reporting business. It can provide an online reporting platform for the people, which is composed of real name reporting, anonymous reporting, reporting investigation, reporting acceptance, reporting query, reporting deletion and other functions,

Interactive communication: the main function of the interactive communication module is to build an online communication platform, which can solve the people's questions about DIAS. It consists of message submission, message reply, message deletion, message query and other functions.

## **4 Business Description**

### **4.1 Business Objectives**

- (1) relying on the data center, build the company's internal inspection platform

In accordance with the idea of unified construction, unified management and hierarchical implementation, relying on the company's data center, carry out the construction of internal inspection platform, so as to realize the full coverage of query scope and the gradual extension and refinement of query ability.

- (2) rely on the business information system to realize the effective supplement of platform query

On the basis of the construction of the internal inspection platform, for the query items that are difficult to be achieved due to the fact that some data in the data platform cannot be directly provided to the external system due to confidentiality, we rely on various business systems as a supplementary query means. The DIAS institution shall obtain or entrust the inquiry from the relevant business department in the form of a letter. The relevant business department shall designate the inquiry cooperation personnel, set up a quick approval procedure, cooperate with the inquiry according to the relevant needs, and pay attention to strictly control the scope of knowledge.

- (3) connect the data channels between the headquarters and units at all levels to realize the sharing of query resources and information

Based on the data service ability of the two-level data, according to the principle of sharing and sharing, and on the premise of ensuring compliance and security, the internal inspection platform realizes the shared query of the two-level data through the role data query authority control, so as to maximize the overall utilization of query resources.

## 4.2 Management Mode

The key work of project management is progress management. According to the overall progress requirements of the project, we should reasonably formulate the milestone plan, and implement the management work of “prediction notification coordination confirmation” in the control of milestone nodes. Second, quality management, carry out the whole process quality management of the project, and strictly control the quality of key links such as demand review, outline design review, system testing, online trial operation, project acceptance, etc. Third, establish a closed-loop management mechanism and a supervision and feedback mechanism for project risks and problems, clarify the division of organizational responsibilities for risk management and problem management, formulate risk management strategies, and carry out project level risk and problem management. Fourth, carry out security work at all stages of the project to ensure the design safety, construction safety, implementation safety and operation safety of the information system, and ensure the safety of the whole life cycle of the information system.

### Safety management system

Data is the connotation of information assets, and data security includes data integrity and confidentiality. Network security management is an important part of ensuring system security. Preventing network intrusion from LAN and WAN is a part of daily work that must be done to ensure system security. Prevent the existence of risk factors with unclear responsibilities and rights and chaotic management. Therefore, in order to ensure the safety of the system, in addition to doing enough work in the realization of technology, we must also rely on strict and standardized safety management provisions to do further safeguard measures.

## 5 Conclusions

At present, the DIAS information platform has been built and successfully applied to the DIAS Bureau. After a period of application, the system has played a good role. The DIAS Bureau has received more reports, and the handling efficiency of reported events is higher, which is in line with the basic objectives set by the management of the DIAS Bureau, but some deficiencies of the system have also been found in the application process, There is no corresponding notification mechanism when the people submit the report information and submit it for acceptance. The corresponding users need to take the initiative to check in the system to obtain the results. In the future, the notification mechanism can be introduced, which will help to improve the work efficiency of the

system; For some added functions, the verification of field format is not rigorous enough, especially the field length, which needs to be corrected in the future. The construction of DIAS BD information query platform based on DC needs further research.

## References

1. Wang, H., Mu, et al.: Management and instant query of distributed oil and gas production dynamic data. *Petroleum Exploration Dev.* **46**(05), 169–176 (2019)
2. Gemmill, J.: Dataset popularity prediction for caching of CMS BD. *Comput. Rev.* **60**(3), 133 (2019)
3. Fiore, S., Elia, D., Pires, C.E., et al.: An integrated big and fast data analytics platform for smart urban transportation management. *IEEE Access* **7**, 117652–117677 (2019)
4. Sandhu, A.K.: BD with cloud computing: discussions and challenges. *BD Mining and Analytics* **5**(1), 32–40 (2022)
5. Salloum, S., Huang, J.Z., He, Y.: Random sample partition: a distributed data model for BD analysis. *IEEE Trans. Industr. Inf.* **15**(11), 5846–5854 (2019)
6. Celesti, A., Fazio, M.: A framework for real time end to end monitoring and BD oriented management of smart environments. *J. Parallel & DC* **132**(OCT.), 262–273 (2019)
7. Uplavikar, N., Malin, B.A., Wei, J.: Lucene-P2: a distributed platform for privacy-preserving text-based search. *IEEE Trans. Dependable and Secure Comput.* **18**(6), 2801–2819 (2020)
8. Hou, J., Zhu, Y., Du, S., et al.: Design and implementation of reconfigurable acceleration for in-memory distributed BD computing. *Future Generation Comput. Syst.* **92**, 68–75 (2019)
9. Shakhovska, N., Boyko, N., Zasoba, Y., et al.: BD processing technologies in distributed information systems. *Procedia Comput. Sci.* **160**(2), 561–566 (2019)
10. Misra, H.: Smart sustainable cities of the future: the untapped potential of BD analytics and context-aware computing for advancing sustainability. *Comput. Rev.* **60**(4), 157 (2019)
11. Aksenov, V., Kuznetsov, P.: Review of the third summer school on the practice and theory of DC SPTDC 2020. *ACM SIGACT News* **51**(4), 82–84 (2021)
12. Yadav, S., Mohan, R., Yadav, P.K.: Fuzzy based task allocation technique in DC system. *Int. J. Inf. Technol.* **11**(1), 13–20 (2019)



# Digital Audit Platform Based on Visual Data Analysis

Qi Liu<sup>1</sup>, Jinhua Chen<sup>2</sup>, Hongchang Wen<sup>2</sup>, Guodong Qi<sup>2</sup>(✉), and Yujing Li<sup>2</sup>

<sup>1</sup> State-Owned Assets Supervision and Administration Commission of the State Council, Beijing 100053, China

<sup>2</sup> Beijing Guodiantong Network Technology Co., Ltd., Beijing 100070, China  
ggd0501@163.com

**Abstract.** With the continuous change of technology in the new era, digitization is sweeping China and even the world. Enterprises generate various types of data in business innovation and upgrading, resulting in data gradually becoming a key production factor for enterprise development. As the supervisor of an enterprise, internal audit should take the initiative to grasp and lead the trend of digital technology transformation in the face of increasing audit data, and carry out digital audit transformation. The main purpose of this paper is to conduct research on the construction of a digital audit platform based on visual data analysis. This paper mainly focuses on the system construction goal of “enabling business and improving management efficiency”, in accordance with the work idea of “planning as the guide, demand-oriented”, and facing the needs of different audit users and work management, to build a system with “open, integrated, A digital audit platform with dynamic and intelligent features. The digital audit platform has achieved full coverage of State Grid’s main business audit supervision, and the level of structure, data and intelligence has been continuously improved, providing strong system support for digital auditing.

**Keywords:** Big Data · Visual Analysis · Digital Audit Platform · Audit Informatization

## 1 Introduction

The construction of the digital audit platform must implement the requirements of the Central Audit Committee for strong scientific and technological auditing, deepen the reform of the management system, emancipate the mind, keep pace with the times, improve the management thinking, and timely reflect and display the new situation, new problems and new trends in many economic fields and society. Adhere to scientific and technological management of energy, strengthen the construction of management details; to speed up the construction and implementation of internal control information. Build a “comprehensive enterprise” information system that integrates “three important and one big” decision-making, investment, finance, capital, operation, internal control and other business systems [1].

In related research, Palviainen mentioned that machine learning (ML) solutions are rapidly evolving and increasingly capable of performing automatic visual data processing (AVDP) tasks. The purpose of the study is to examine the skills and insights of companies using different types of AVDP solutions in their business. Kruesi et al. to evaluate the opportunity to open a biomedical knowledge base (OBR) using a distributed network of the knowledge management system (KMS) conceptual framework. An innovative KMS conceptual framework is proposed to guide the transition from traditional, siloed approaches to sustainable OBR. It turns out that the proposed KMS framework aligns the OBR requirements with the people, process, technology, and content elements of the KM standard [2].

This paper studies the construction of digital audit platform based on visual data analysis. This paper mainly introduces the digital audit platform and its advantageous applications, and analyzes the reasons for its construction and business goals. Explore the construction of an audit tool platform and carry out online auditing of key businesses; State Grid's implementation work requires proactively adapting to new challenges. Central SOEs took the lead in proposing digital auditing as the main carrier for the construction and application of audit informatization, making the construction of audit informatization enter the fast lane.

## 2 Design Research

### 2.1 Digital Audit Platform

The digital audit platform implements the construction concept of "openness, integration, dynamic and intelligence", and fully supports the online audit work.

#### (1) Platform Introduction

The digital management platform consists of management control area, control operation area and main database. Among them, the management area is based on audit project management and daily audit management, and realizes the electronic management of the whole process of the project; the operating environment is based on the independent analysis library and intelligent analysis model library, and provides a variety of professionals. The audit analysis service fully supports the audit business needs; the data area is based on the actual production data of the secondary data center of the State Grid, and provides all the data sources required for the evaluation. The digital audit platform has achieved full coverage of State Grid's main business audit supervision, and the level of structure, data and intelligence has been continuously improved, providing strong system support for digital auditing [3].

#### (2) Platform Advantages

##### 1) Intelligent project process

Through the built-in audit template library, standardized entry audit knowledge base, mining and utilization of historical audit results, integrating internal and external audit data, analyzing the content elements, dependent data and dependent logic of the core links of the audit project operation process, the intelligence of the core links of the audit project is realized [4].



## 2) Online audit work

The digital audit platform provides auditors with a data analysis platform and a variety of technical means. Using the audit model library and independent analysis library, auditors can rely on various business data in the data center to complete remote online audit work.

## 3) Diversification of audit data

The audit data domain is based on the data center of the State Grid Corporation, supplemented by the audit business database, aggregates audit business data, constructs an audit intermediate table, and provides comprehensive data support for digital auditing [5].

## (3) Application Results

At present, the digital audit platform has been fully deployed and applied in the headquarters of State Grid Corporation of China, 6 branches, 27 provincial power companies, and 39 directly affiliated units, and has been extended to large state-owned enterprises such as National Pipe Network, China Construction Technology, and China Green Development. It has strongly supported the construction of audit digital transformation and has been widely praised by customers.

Improve business structure, strengthen key core technology business research, and improve development quality and efficiency. The digital audit platform will combine the advantages of “business + data” to achieve a large-scale development layout, build a data application and service product system, and drive the digital decision-making of the power grid.

## 2.2 Reasons for Building

In terms of information system audit, the company's information system audit work is still in its infancy. With the steady progress of the construction of the Internet of Things in electric power, the company has built a number of information systems covering the company's entire business after more than ten years of informatization construction. Carrying out a special audit of informatization construction, combined with the headquarters economic responsibility audit project, taking informatization construction and information system security as an important audit content, using digital audit tools to broaden the way of audit evidence collection, allowing auditors to remotely view information system logs, log in, operation process and other data, improve the efficiency and quality of information system audit [6].

In terms of the improvement of platform management functions, the current audit project operation has low correlation and weak intelligence in the whole process; the degree of data sharing is low, and multi-point linkage data analysis cannot be realized, and there is no strong support for in-depth analysis and accurate discovery of problems; basic knowledge There is no intelligent correlation reference for the result data such as database and audit operation domain model operation doubts, and it is not automatically pushed to the audit record and draft preparation stage, forming the basic basis for audit record and draft preparation, and there is an information island between the audit management process and the operation model data; In terms of audit talents, there are more than 3,000 full-time and part-time audit talents in the audit talent pool of State Grid. The auditors build audit teams mainly based on personal experience, lack of scientific analysis data

support for personnel resumes, past experience and professional expertise, and it is difficult to accurately match the needs. Auditors cannot maximize their own advantages of existing auditors. In terms of cross-domain auditing, due to changes in related audit services, it is often difficult to find problems in the audit of a single business domain. It is necessary to further expand the construction of a cross-domain comprehensive analysis model, strengthen the horizontal integration of the audit business field, and open up the vertical direction from the headquarters to the provincial (city) company. Model management to meet the new requirements for cross-domain audit improvement. Comply with the requirements of “Phase III construction and deployment of digital audit platform and cross-domain business scenarios” [7].

At present, the discovery of audit doubts still relies on a large number of manual inspections, resulting in the current situation of heavy workload, heavy tasks, and time-consuming and laborious work. Combined with the application of unstructured data processing of natural language processing and related theories, and artificial intelligence technology to the audit process, various Automatic extraction of key information from unstructured data such as unstructured audit documents, effectively solving the blind spots and pain points caused by automated audit screening, real-time judgment of doubts, and sampling audits, and build unstructured data conversion tools to achieve unstructured Transform into structure and realize the goal of in-depth application of unstructured data.

### 2.3 Business Objectives

In terms of main business audit of directly affiliated units, firstly, the audit field of directly affiliated units expanded, including international business audit, general contracting business audit, equipment manufacturing business, pumped storage business, ZTE business, ICT business, and general aviation business audit, realizing digital auditing. Full coverage of the company’s audit unit and audit content. The second is to improve and improve the financial business audit, in accordance with the financial industry regulatory requirements, combined with the business audit needs of various financial units, based on the management process of financial business, to achieve the whole process of financial business management and supervision.

Information system audit: including information system construction, implementation of information security control strategies and measures, information technology risk assessment, information work risk assessment, etc., to promote the quality and efficiency of the company’s information construction [8].

Monitoring and early warning: Based on the digital audit platform, it provides audit monitoring and early warning functions, and issues early warning signals. Through the establishment of a complete set of business monitoring indicators, it conducts real-time monitoring and analysis, and plays a role in dynamic monitoring and early warning.

In terms of platform management function improvement: based on an “open, dynamic, integrated, and intelligent” digital audit platform, it is oriented to prevent risks, promote governance, feed back the construction of the energy Internet, and escort the realization of the company’s strategic goals. It has carried out in-depth digital audit work in terms of efficiency improvement, audit platform management innovation and audit business data governance. Through the optimization and improvement of the functions of the management domain of the digital audit platform, the whole-process correlation and

intelligence of the audit project operations have been strengthened, and the construction level of the basic knowledge base has been further improved; advanced application optimization includes cross-domain scene audit model, intelligent audit There are two parts of the work, starting from the requirements of function, friendliness, and performance, with a total of 6 requirements points, mode default value modification and optimization, adaptation of Bi analysis tools, table Chinese label function, intelligent model library model curing customization function, provincial (city)) Model Execution Status Statistics and Smart Model Library Released Model Maintenance. Through this iterative upgrade, the usability and capabilities of the system have been improved, and the overall optimization and transformation of the audit platform has been completed. Cross-domain audits include marketing cross-domain audit scenarios, material cross-domain scenarios, engineering cross-domain scenarios, and human resources cross-domain scenarios, enabling digital auditing to fully cover the company's audit units and audit content; continuous audit supervision and early warning, including the State Grid Audit Department Supervision and early warning topics and regional center supervision and early warning topics, etc., promote the quality and efficiency of the company's informatization construction, further implement the concept of data assets, build online unstructured data conversion tools, give full play to unstructured file processing capabilities, and improve data management and processing capabilities, to realize all-round data mining and utilization [9].

## 2.4 Implementation of BP Neural Network

- (1) Network initialization.
- (2) Hidden layer output calculation.

The input vector  $X$ , the weight  $\omega_{ij}$  and the threshold  $a_j$  ( $j = 1, 2, \dots, l$ ) are calculated with reference to Formula 1.

$$H_j = f\left(\sum_{i=1}^n \omega_{ij}x_i - a_j\right), j = 1, 2, \dots, l \quad (1)$$

In (1),  $n$  is used to represent the hidden layer node, and the excitation function is  $f(*)$ .

- (3) Output layer output calculation.

Substitute  $H_j$ ,  $\omega_{jk}$ ,  $b_k$  ( $k = 1, 2, \dots, m$ ) into formula 2 for calculation, and obtain the output layer result  $O_k$ .

$$O_k = \sum_{j=1}^l H_j\omega_{jk} - b_k, k = 1, 2, \dots, m \quad (2)$$

- (4) Error calculation.

As shown in Eq. 3, the difference between the output layer result  $O_k$  and the expected value  $Y_k$  is made to obtain the prediction error  $e_k$ .

$$e_k = Y_k - O_k, k = 1, 2, \dots, m \quad (3)$$

- (5) Update the weights.

Based on the prediction error  $e_k$ , according to Eq. 4 and Eq. 5, the adjusted connection weights  $\omega_{ij}$  and  $\omega_{jk}$  are obtained.

$$\omega_{ij} = \omega_{ij} + \eta H_j (1 - H_j) x_i \sum_{k=1}^m \omega_{jk} e_k, i = 1, 2, \dots, n; j = 1, 2, \dots, l \quad (4)$$

$$\omega_{jk} = \omega_{jk} + \eta H_j e_k, j = 1, 2, \dots, l; k = 1, 2, \dots, m \quad (5)$$

In the above formulas (4) and (5), the  $\eta$  variable is the learning rate.

### 3 Experimental Study

#### 3.1 Highlights of Digital Audit Platform Construction

- (1) Focus on the data foundation and build a full-service element data environment

The audit data domain is to support the audit operation model and management decision analysis in an all-round way. We always take the construction of audit data mart as the top priority of platform construction. Up to now, the company has deployed 246 audit intermediate tables in the data domain of the two-level digital audit platform, covering the five majors of company engineering, materials, human resources, finance and marketing., involving more than 8,000 core business fields of 32 sets of business management systems are fully accessed. At the same time, the audit data resource directory is constructed according to specialties and scenarios, and the data access interface is improved, so that the data can be understood, found accurately, and used quickly. Auditors carry out multi-dimensional data analysis to support project operations and provide strong data support. In the process of promoting the construction, the data domain also feeds back the digital transformation of the company [10].

- (2) Management domain project process intelligence

The audit management domain is the basic platform for audit project implementation and daily audit management. Through the built-in audit template library, standardized entry audit knowledge base, mining and utilizing historical audit results, integrating internal and external audit data, analyzing the content elements, dependent data and dependent logic of the core links of the audit project operation process, the automation of the core links of the audit project is realized.

- (3) Integration of work domain and audit business

The audit operation domain is based on the platform data domain foundation, and provides various professional audit analysis models and data analysis tools for auditors at different levels to fully support audit project operations. It mainly includes audit model library and autonomous analysis library.

Relying on the “two databases” to concretize the audit model, connect the audit business and enterprise data in series, realize the “digital” presentation of audit logic and audit methods, and become the core vitality of the digital audit platform.

### 3.2 Overall Architecture of the Digital Audit Platform

(1) Business Architecture as shown in Fig. 1:

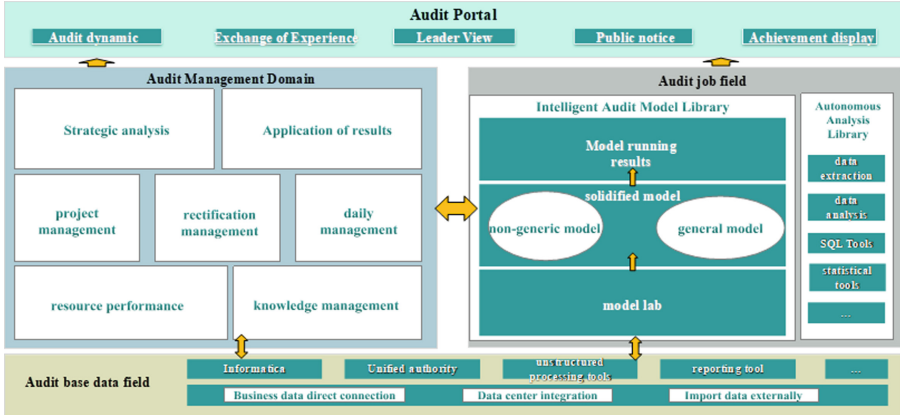


Fig. 1. Business architecture of the digital audit platform

(2) Data Architecture

The overall data architecture of the system in this issue follows the overall data architecture design of the digital audit platform, and the system data model design follows the SG-CIM model design. The business data mainly involves the financial domain, material domain, market domain, personnel domain, power grid domain, and project domain. The data architecture diagram is as follows Fig. 2:

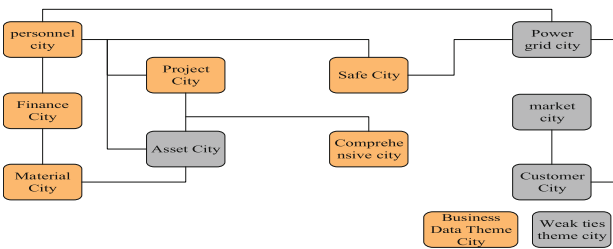


Fig. 2. Data architecture diagram of the first-level subject domain of the digital audit platform system

### (3) Technical Architecture

**Data layer:** Integrate the data center (analysis domain), build a digital audit platform audit data mart based on the enterprise data warehouse (SG-CIM4.0) and the data center.

Build an independent audit business library to store the data generated by the audit management of the digital audit platform and the configuration data of audit jobs (audit models and advanced applications in various fields).

**Public service layer:** The platform function module of the digital audit platform, that is, the core functional components, including full-text search, instant messaging and user behavior logs.

**Application layer:** It includes functional modules such as digital audit platform audit management, audit models in various fields, and advanced applications. Use SG-UAP3.0 development platform for development.

**Presentation layer:** Adopt the popular Bootstrap, Echarts, HTML5 and other technologies in the industry. Bootstrap is a front-end framework used to control page style and layout; Echarts is used for chart display of digital audit platform decision analysis and other functions; HTML5 is the latest revision of HTML, used to mark page elements.

**System integration:** including integration with portal, ISC, I6000, data center and big data platform. The portal adopts the interface integration method, the ISC, I6000 and big data platform adopts the application integration method, and the data center adopts the data integration method.

## 4 Experiment Analysis

### 4.1 Capacity Planning

The original data of the cross-domain audit and information system audit of the audit objects of the digital audit platform is stored in the data center shared domain, the audit result data is stored in the data center analysis domain, and the audit comprehensive management data is stored in the audit's own database. Based on the current stock data and service life of the audit comprehensive management system of the digital audit platform, the capacity of the audit business database is estimated as shown in Tables 1, 2 and 3:

**Table 1.** Estimated number of users

project		Digital Audit Platform
Phase II	total number of users	5000
	Maximum number of online users	200
	System concurrency	50
Phase III	total number of users	5000

(continued)

**Table 1.** (continued)

project		Digital Audit Platform
	Maximum number of online users	200
	System concurrency	50

**Table 2.** Headquarters data capacity estimation table

content		Audit business library		total
Data Classification		Comprehensive improvement of platform management	Audit job configuration data	
Phase II	Data volume (GB/year)	175	15	190
	Estimated storage life (years)	5	5	5
	Estimated total amount of stored data (GB)	875	75	950
Phase III	Estimated data volume (GB/year)	500	10	210
	Estimated storage life (years)	5	5	5
	Estimated total amount of stored data (GB)	2500	50	1050

**Table 3.** Provincial company analysis domain data capacity estimation table

content		Audit business library	Audit synthesis
Data Classification		Audit job result data	Comprehensive improvement of platform management
Phase II	Data volume (GB/year)	23	175
	Estimated storage life (years)	5	5
	Estimated total amount of stored data (GB)	115	875
Phase III	Estimated data volume (GB/year)	20	500

(continued)

**Table 3.** (continued)

content		Audit business library	Audit synthesis
Data Classification		Audit job result data	Comprehensive improvement of platform management
	Estimated storage life (years)	5	5
	Estimated total amount of stored data (GB)	100	2500

#### 4.2 The Future Construction Direction of the Digital Audit Platform

Prospects for the next stage of work: Focus on the implementation of the company's strategy and the implementation of major decisions and arrangements of the party group, adhere to the leadership of party building, ensure the performance of audit responsibilities with technology-enforced audits, accelerate the transformation and upgrading of audit work, and strengthen the company's management risks, key areas, and key links. Implement effective supervision and promote the modernization of the corporate governance system and governance capacity.

(1) Improve the digital audit data mart

Based on the company's digital infrastructure construction (including data center and SG-CIM), iteratively upgrades the audit intermediate table, and expands the scope of the platform's data to the core business data of financial, international, emerging and other units through the third-phase project of the platform; expands non-structural Data conversion and technology application will be enriched to enrich audit data sources; the standardization of audit data marts will be strengthened, and the audit intermediate tables will be iteratively upgraded. At the same time, this will be the starting point to continuously improve the company's data quality and promote the company's digital transformation and development [11].

(2) Functions of iterative digital audit platform

R&D and deployment of directly subordinate unit work domains, provide general models and advanced applications of "two databases"; digitize important audit business processes such as audit project management, human resources, and problem rectification, build an audit work management standard system, and realize the interconnection of audit operations and management. Interoperability and data-driven; focus on key areas of finance, marketing and other core business areas such as quality improvement, asset effectiveness, electricity bill recovery, etc., and build cross-domain comprehensive analysis and early warning monitoring models; strengthen internal control audits and add information system audit modules[12].

(3) Upgrade the digital operation mode

Continue to strengthen the in-depth application of the digital audit platform in audit projects, cultivate the data thinking of auditors, give full play to the leading role of the headquarters in digitalization, and consolidate off-site data analysis based on



audit projects; give full play to the digital audit platform. The functional advantages of general models and advanced applications promote the in-depth development of digital audit supervision to cross-domain audit and early warning audit.

## 5 Conclusions

Digital auditing can play a greater role in supervising the production and operation of enterprises and preventing enterprise risks. Therefore, enterprises should strengthen the construction of digital auditing, play the auditing function through the digitalization of auditing, and realize the “full coverage” requirement of auditing proposed by the National Audit Office. At the same time, the development of digital auditing has brought new challenges to the organizational structure and management mode of the company’s internal control audit. Only by continuously innovating auditing methods and applying digital means in the auditing process can the effectiveness of auditing be improved, so as to meet comprehensive audit coverage overall requirements.

## References

1. El-Sherif, D.M., Abouzid, M.: Analysis of mHealth research: mapping the relationship between mobile apps technology and healthcare during COVID-19 outbreak. *Glob. Health* **18**(1), 1–11 (2022)
2. Harnett, J.T., Vithlani, S., Sobhdam, S., et al.: National audit of antidote stocking in UK emergency departments. *Eur. J. Hosp. Pharm.* **28**(4), 217–222 (2021)
3. Palviainen, M., Harviainen, T., Lopez, M.B., et al.: Boosting business with machine learning based automated visual data processing: results of finnish company interviews. *IEEE Access* **8**, 99171–99179 (2020)
4. Kruesi, L., Burstein, F., Tanner, K.: A knowledge management system framework for an open biomedical repository: communities, collaboration and corroboration. *J. Knowl. Manag.* **24**(10), 2553–2572 (2020)
5. Youngs, P.J.: Decision to delivery time. *Neurol. Neurochir. Pol.* **16**(4), 1330–1333 (2021)
6. Bellgard, M.I.: ERDMAS: an exemplar-driven institutional research data management and analysis strategy. *Int. J. Information Manage.* **50**, 337–340 (2020)
7. Dka, B., Has, B., He, C., et al.: Visualization and isolation of zone-specific murine hepatocytes that maintain distinct cytochrome P450 oxidase expression in primary culture. *Biochemical and Biophysical Research Commun.* **528**(3), 420–425 (2020)
8. Pereira, M., Burns, D., Orfeo, D., et al.: 3-D multistatic ground penetrating radar imaging for augmented reality visualization. *IEEE Trans. Geoscience and Remote Sensing* **58**(8), 5666–5675 (2020)
9. Kelly, J.W., Ostrander, A.G., Lim, A.F., et al.: Teleporting through virtual environments: Effects of path scale and environment scale on spatial updating. *IEEE Trans. Visualization Comp. Graphics* **26**(5), 1841–1850 (2020)
10. Behrisch, M., Schreck, T., Pfister, H.: GUIRO: user-guided matrix reordering. *IEEE Trans. Visual Comput. Graphics* **26**(1), 184–194 (2020)
11. Selvaraju, R.R., Cogswell, M., Das, A., et al.: Grad-CAM: visual explanations from deep networks via gradient-based localization. *Int. J. Comput. Vision* **128**(2), 336–359 (2020)
12. Ramlawi, N., Bharadwaj, N.A., Ewoldt, R.H.: The weakly nonlinear response and nonaffine interpretation of the Johnson–Segalman/Gordon–Schowalter model. *J. Rheol.* **64**(6), 1409–1424 (2020)



# The Division of Labor and Matching System of Teaching Reform Model Based on Multi-objective Classification Algorithm

Min Liu<sup>(✉)</sup>

Shandong Vocational and Technical College of Commerce, Jinan 250001, Shandong, China  
lm13953178822@163.com

**Abstract.** With the progress of technology and the development of society in our country, the country's demand for high-quality talents is also increasing. Modern education presents many forms of teaching concepts and teaching methods, but my country's current teaching mode (TM) cannot meet the needs of the needs of high-quality talents, so it is necessary to reform the TM. This paper is based on the multi-objective algorithm (MOA) to design the teaching reform mode system. This paper firstly explains the difficulty of solving the MOA, the related concepts and existing problems of the TM, and then designs the teaching reform mode system, and finally analyzes the results of the system implementation. In the analysis of the effect of the reform of the TM, it was found that the proportion of students in Class A with scores between 70–80 and 80–90 was relatively high, accounting for 27% and 38%, respectively. The effect of the new TM reform is remarkable.

**Keywords:** Multi-Objective Algorithm · Teaching Reform · Teaching Mode · Teaching System Design

## 1 Introduction

Looking back at history, for every major reform of TM, the Ministry of Education spends a lot of energy and financial resources to invest in curriculum planning and design, in order to achieve a more perfect TM. However, the design and research of the TM system based on the MOA is relatively small, and the matching evaluation system and the system reform are relatively lagging behind, which directly leads to the appearance of the use of each module in the reformed TM. What's more, because the school level has made great efforts to implement the teaching reform model, it basically adopts a "one-size-fits-all" model and enforces the reform, which leads to the situation that each teaching module is forcibly copied and the classroom efficiency is low in all disciplines and years. It is impossible to do a specific analysis of the specific situation at all [1].

At present, many scholars have carried out in-depth research on the system design of TM based on MOA, and have achieved good results. For example, Pawusik M et al. Proposed NSGA-II. The strategy in this algorithm is widely used to solve multi-objective problems. The main idea is to first stratify individuals in the target space through fast non-dominated sorting, and then use elite retention. The strategy saves the optimal solution

to improve the exploration performance of the algorithm, and finally uses the crowding distance of individuals in the target space as the individual selection mechanism to improve the distribution of the solution in the target space. This idea has been used for reference by a large number of algorithms [2]. Elsedimy E I proposed a multi-objective optimization algorithm based on enhanced levy to efficiently solve the VM placement problem as well as for improving non-dominated virtual machines in archives, using a mechanism to maintain archives [3]. Although MOAs are popularized around the world, there are few studies that have been successfully applied to teaching reform models. However, with the reform of TMs, the integration of MOAs can promote the reform and development of TMs.

This paper first analyzes the difficulties of multi-objective solution, puts forward the relevant concepts of TM and the problems existing in the current TM, and then designs the TM reform system based on the MOA. Finally, after the system design is completed, the teaching reform the implementation results of the system are analyzed. Through the analysis of relevant data, it is found that the implementation effect of the teaching system is remarkable. Finally, suggestions for the teaching reform model are put forward.

## 2 Related Concepts

### 2.1 Difficulties in Solving Multi-objective Optimization Problems (OM)

Solving multi-objective OM is to find an approximate solution of a set of PF after considering multiple objectives at the same time. For this-group should try to satisfy:

- (1) The obtained solution set is as close to PF as possible.
- (2) Keep the solution diversity as much as possible in the target space.

For the first point, the requirement to solve the set as close to PF as possible is to ensure that the algorithm has a good convergence, and convergence is an important criterion for evaluating the performance of the algorithm in all OMs. For the second point: keep the solution diversity. Since the multi-objective OM obtains a set of optimal solutions in the solution space in the target space, we hope that the obtained solution can cover as complete and uniform as possible while ensuring the convergence performance on PF [4, 5]. In the multi-objective OM, the diversity of the solution set refers to the obtained--the Euclidean distance between the solutions in the group solutions is relatively far, so that even if fewer solutions are used, the PF can be expressed to a large extent. The information contained in it is more conducive to the diversified choices of decision makers. The specific measure of the diversity of solutions depends on the uniformity of the Euclidean distance between the solutions and the coverage area of the optimal solution obtained on the PF [6].

### 2.2 TMI

TM refers to the specific teaching activities organized under the guidance of a certain thought [7]. As a way of teaching activities, TM can deal with the relationship between

the elements and the whole of teaching activities as a whole, so as to highlight the regularity and controllability of teaching activities. Different regions, different periods, and different levels of education often have different TMs, and different TMs are also required for different disciplines or people with different educational backgrounds. They are not the same, the TM must be consistent with the training objectives, and the TM should be selected reasonably according to the actual situation [8].

The TM is not a rigid teaching procedure, but an open subsystem. As educational ideas become richer day by day, the educational system is constantly updated and developed, and the TM is also constantly developing and changing [9]. Teaching practice is the logical starting point for the generation of TMs, which provides feasibility for the generation of new TMs, and is also the process of verifying and improving TMs, providing experience for the correct selection or application of a TM [10].

### **2.3 Problems Existing in the Current TM**

#### (1) The teaching method is relatively simple

At present, the teaching method is still mainly based on traditional classroom teaching. The advantage of teaching method teaching is that it can explain the relevant theoretical knowledge of mechanical manufacturing technology to students at the fastest speed. However, this teaching method is only for the purpose of instilling knowledge. Develop students' practical ability.

#### (2) The teaching assessment method is not scientific

At present, the assessment method of the course is still the evaluation method of the paper assessment. This kind of assessment method of written answer sheets causes students to completely ignore the key teaching content of the course in the learning process, and only pay attention to the key test content planned by the teacher at the end of the semester. Usually, I don't care at all in my studies. During the exam, I only want the scores on the paper to pass. I hurriedly answer the paper without pursuing the quality of the paper, and forget it after the test.

#### (3) Ability training is neglected

At present, the teaching method ignores practical teaching, and the teaching of relevant knowledge is mainly through the teacher's explanation of the knowledge in the classroom, did not translate the knowledge learned into practice, which directly led to the slow progress of innovation ability development.

## **3 System Design**

### **3.1 System Design Goals**

The users of the teaching reform system include students, teachers and administrators, and they all have different requirements for system functions. The main users of the

system are students. They want to learn anytime and anywhere, and can communicate and communicate online. The system can store the progress of learning and test the results of learning. Teachers can work at any time, communicate with students academically, mark examination papers, and manage students' grades. The system administrator has the greatest authority, and must carry out database and management on the database at any time. Because the TM system mainly serves students, the design of the TM system is mainly based on student users, and the students should do the following: first, students can view the courseware in the system, and the system intuitively displays the course chapters. The system has a clear hierarchical structure and does not There will be cases where the function points cannot be found. Secondly, after learning a chapter, you can deepen your memory and consolidate your learning results through exercises. Finally, when you encounter difficult problems, you can solve the learning problems encountered through online real-time communication. The sharing and reuse of resources are realized through the system. Therefore, we designed the overall structure diagram of the system as shown in Fig. 1.

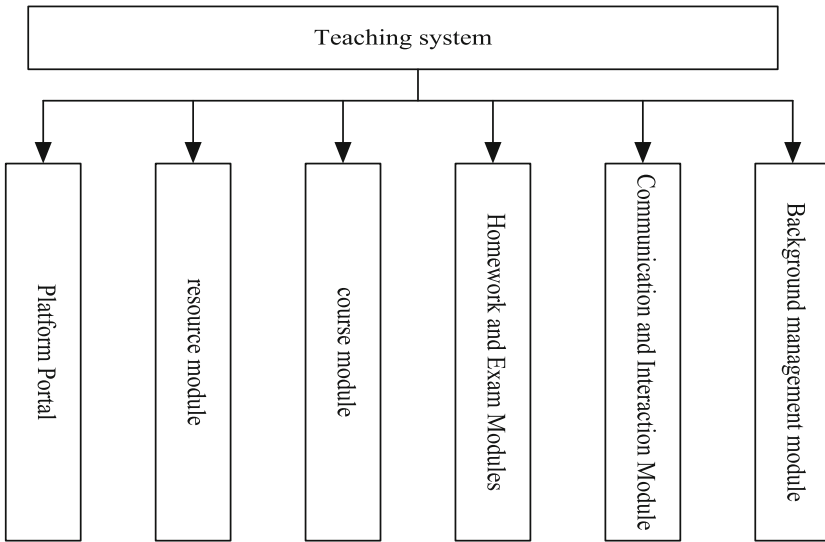


Fig. 1. Overall structure of the system

### 3.2 Evaluation Indicators for Multi-objective OMs

Because the objectives to be optimized in multi-objective OMs often conflict with each other, for example, in the maximization problem, when one of the optimized objectives is improved, it will inevitably lead to the decline of other objective values. Therefore, it is impossible to directly the performance of the multi-objective optimization algorithm is measured by comparing the objective value of the obtained solution. According to the different evaluation indicators, the performance of the algorithm is different. In order

to consider the convergence and diversity of the solution at the same time: IGD and HV, these evaluation indicators can take into account the functions of the above two evaluation indicators, and measure the convergence performance and diversity of the algorithm at the same time.

(1) Inverse Iteration Distance (IGD)

$$IGD(O, P^*) = \frac{\sum_{q \in P^*} d(q, O)}{|P^*|} \quad (1)$$

where  $O$  represents the PF obtained by the MOA,  $P^*$  represents the real PF, and  $d(q, O)$  represents the Euclidean distance of the solution between  $q$  and  $O$ . It can be seen from the formula that the smaller the IGD value, the closer the obtained PF is to the reference point uniformly distributed in the target space, the better the convergence and diversity of the algorithm.

Hyper-Volume Difference (HV)

$$HV(P) = vol(U_{q \in O} [q_1, z_1^*] \times \dots [q_m, z_m^*]) \quad (2)$$

where  $vol$  represents the Lebesgue measure, and  $O$  represents the approximate solution obtained by the algorithm in the target space. In short, the HV index is to evaluate the performance of the algorithm by calculating the volume of the region dominated by the approximate solution set  $O$  in the target space with  $z^*$  as the boundary. The larger the HV index value, the larger the hypervolume dominated by the approximate solution set, which proves that the closer the solution set is to the true PF, the better the algorithm performance.

## 4 System Implementation

### 4.1 Analysis of the Implementation Effect of the Teaching Reform Model

Aiming at the implementation of the TM reform based on the MOA, this study selected two classes A and B with the same enrollment time, similar class size, and no significant difference in academic performance in a school to compare the system design effect. Among them, there are 48 students in class A and 52 students in class B. Class A is taught by this TM system, and class B is taught by traditional TM. The design of the system is evaluated through the academic performance levels of the two classes at the end of the term. The effect of the implementation of TMI reform. Analysis of the mastery of basic knowledge Because class A adopts this TM to guide teaching, although the assessment and evaluation method is different from that of class B, which adopts traditional TM, both of them use the test papers issued by the school at the end of the term, so here Only the results of the final exam papers of the two classes were compared, and the teaching effect of the basic knowledge part was analyzed. Table 1 and 2 are shown in Table 1 and Table 2.

By analyzing the score statistics of the two classes A and B, we can draw a comparison between the two classes using the same test paper for the test, the proportion of the

**Table 1.** Statistical table of grades in class A

Score range	Below 60	60–70	70–80	80–90	90–100
Number of people	3	5	13	18	9
The proportion	6%	10%	27%	38%	19%

**Table 2.** Statistical table of grades in class B

Score range	Below 60	60–70	70–80	80–90	90–100
Number of people	8	17	18	7	2
The proportion	15%	33%	35%	13%	4%

students in class A between 70–80 and 80–90. High, 27% and 38%, respectively; Class B has a higher proportion between 60–70 points and 70–80 points, 33% and 35%, respectively. It can be seen from the data statistics that compared with the B class that adopts the traditional TM, the A class that adopts this TM has better grades. This shows that the application of this TM to the A-class related knowledge is better, so that the students' final exam scores have been improved, and in terms of the basic knowledge part of the TM reform, good results have been achieved.

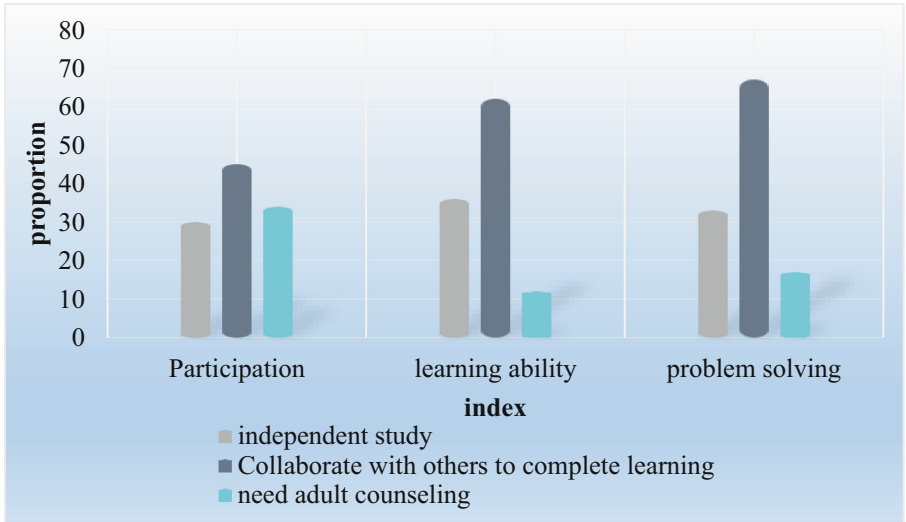
#### 4.2 Students' Participation in Teaching Reform Model

On the student side, when participating in each module after the reform of the TMI, children with better academic performance have significantly higher participation rates, while underachievers are basically unable to participate and dare not participate at all. This is positively correlated with the level of students' own learning ability and study habits. Figure 2 shows the relationship between students' participation, learning ability, and learning style after the reform of the TM.

The main concern of parents is whether their children can learn and grow in a relatively relaxed and pleasant environment. Students have a higher voice to reduce their burden, and they are opposed to the sea-of-question tactics and "cramming" teaching. They are also concerned about the school's philosophy and teachers' teaching and research. The ability is generally more trusted, and I believe that as long as it is good for the child and the way the child likes, they will support it. Of course, the expectations for grades are also significantly higher. It is still expected that in the new classroom reform TMI, students' academic performance can be significantly improved, and they also have a more objective understanding of the current situation of children.

#### 4.3 Countermeasures and Suggestions

The reform of the TM based on the MOA brings not only a simple change of the central role of the classroom, but also a comprehensive innovation of the educational concept.



**Fig. 2.** Distribution map of the relationship between academic participation and learning ability under the new TMI

This kind of innovation will penetrate into all aspects of the teaching process, forming a chain reaction and producing various new changes. Therefore, this paper puts forward the following three suggestions.

- (1) Cultivate students' awareness of being the subject of the classroom. Teachers should carry out the awareness that students are the subject of the classroom throughout the entire classroom teaching process, strengthen the guidance of students, allow students to join in the classroom, and promote students' motivation. Learning by experience; teachers give students tasks and give students a certain amount of time, so that students can gradually break down tasks. In the process of analyzing problems, students are encouraged to extensively collect and achieve this task by consulting literature and using information technology such as the Internet. The data related to the task goal, urge students to independently discover and solve problems, enable students to gradually establish the awareness of independent inquiry and investigation, cultivate their ability to collect literature, and help students to complete the selection of papers and projects in the future. Expand the accumulated material.
- (2) Teachers should improve the level of theoretical knowledge and practical ability. In teaching, they can adjust the teaching plan according to the needs of students. In the system design, they can fully learn from the educational concepts and methods of Western countries to promote the mature teacher skills training model.
- (3) In the process of TM reform, teachers should hold regular communication activities to improve their own shortcomings through communication, so as to achieve the goal of condensing the results of TM reform. The reform of TM is a gradual process, and any reform cannot be achieved in one step. Therefore, it will take a certain amount of time to fully realize the multi-objective TM.



## 5 Conclusions

In this paper, the TM system is designed based on the theory of MOA. The problems existing in the current TM are that the teaching method is relatively simple, the assessment method is unscientific, and the ability training is neglected. Therefore, three suggestions are finally put forward: cultivating students' awareness of being the main body of the classroom, strengthening teacher training and strengthening the management of the TM reform process. Although this paper proposes the design framework of the TM system based on the MOA, due to its limited ability, there are many deficiencies in the research on the teaching reform mode, but how to reform the TM is still a problem worth exploring.

## References

1. Boulanouar', K., Hadjali, A., Lagha, M.: Trends summarization of times series: a multi-objective genetic algorithm-based model. *J. Smart Environ. Green Comput.* **2**(1), 19–33 (2022)
2. Pawlusik, M., Szałpczyński, R., Karczewski, A.: Optimising rig design for sailing yachts with evolutionary multi-objective algorithm. *Polish Maritime Res.* **27**(4), 36–49 (2020)
3. Elsedimy, E.I., Algarni, F.: An enhanced MOA for virtual machine placement in the cloud computing environment. *J. Comput. Theor. Nanosci.* **16**(5), 1821–1827 (2019)
4. Ahmadi, H., Rajaei, A., Nayeripour, M., et al.: A hybrid control method to improve LVRT and FRT in DFIG by using the MOA of krill and the fuzzy logic. *Iranian J. Electr. Electron. Eng.* **14**(4), 330–341 (2018)
5. Žalik, K.R., Žalik, B.: Multi-objective evolutionary algorithm using problem-specific genetic operators for community detection in networks. *Neural Comput. Appl.* **30**(2), 1–14 (2017)
6. Olteanu, A., Pietraru, R.N., Olarescu, S.M., et al.: Innovations in the educational process in technical universities based on an ontology for interactive teaching system. *Revue Roumaine des Sciences Techniques - Serie Électrotechnique et Énergétique* **66**(1), 53–58 (2021)
7. Rublev, V.S., Kondakov, M.D.: Automated teaching system “Sets” (research for organizing the 1st part of the project). *Model. Anal. Inform. Syst.* **28**(1), 90–103 (2021)
8. Wei, M., Chung, H.F., Li, D., et al.: The current situation and a review of Chinese library and information science from the perspective of the teaching system. *J. Educ. Libr. Inf. Sci.* **61**(1), 25–47 (2020)
9. Jang, M.S., Ruy, W.S., Park, C.K., et al.: A study on the prediction system of block matching rework time. *J. Soc. Naval Architect. Korea* **55**(1), 66–74 (2018)
10. DeVaro, J., Gürtler, O.: Advertising and labor market matching: a tour through the times. *J. Labor Econ.* **36**(1), 253–307 (2018). <https://doi.org/10.1086/693872>



# Domestic AI Chip Based on Internet of Things Technology and Artificial Intelligence Algorithm in Power Grid Infrastructure Construction

Weihua Zhong<sup>1</sup> (✉), Jun Liu<sup>2</sup>, Jianxiang Xie<sup>3</sup>, and Ji Zhang<sup>4</sup>

<sup>1</sup> China Southern Power Grid Corporation, Guangzhou 510070, China  
zhongww7648@sina.com

<sup>2</sup> Digital Grid Research Institute of China Southern Power Grid, Guangzhou 510070, China

<sup>3</sup> Guangzhou Power Supply Bureau, Guangdong Power Grid Co., LTD., Guangzhou 510630, China

<sup>4</sup> China Southern Power Grid Big Data Service Co., LTD., Guangzhou 510630, China

**Abstract.** At present, the on-site management of the power grid infrastructure projects of the power supply company still mainly relies on manual management, and only relies on the internal online power grid project management system for business management, which has the problems of low business management efficiency and inconvenient management. The purpose of this paper is to study the application of domestic AI chips in power grid infrastructure construction based on IoT technology and artificial intelligence algorithms. Aiming at the problems existing in the information construction and development of the power grid construction industry, by introducing RFID technology and using self-controllable domestic chips, this paper focuses on the domestic Loongson 1B chip, including on-chip resources and interfaces and the internal structure of the chip. The SM4 encryption algorithm and how to implement data encryption and decryption are introduced. Reduce personnel and physical management risks and costs in the process of engineering construction, standardize the management of enterprise construction projects, and improve the level of information management of power grid construction enterprises.

**Keywords:** Internet of Things Technology · Encryption Algorithm · Domestic Chips · Power Grid Infrastructure Construction

## 1 Introduction

With the continuous development of the domestic power system, power companies all over the country have strengthened the construction of power grid infrastructure. Therefore, the management quality and level of power grid infrastructure construction has become a key factor for power companies to build smart grids [1, 2]. Because the construction sites of power grid infrastructure projects are widely distributed and involve a large number of personnel, there are relatively big problems in management [3]. At present, the management of power grid infrastructure construction in most domestic

power systems is mainly automated through management software [4]. However, because such management software is usually located in the company's internal network, it cannot cover the project management activities at the construction site. As a result, the site management business still mainly relies on manual management and implementation, and there are problems such as low efficiency and low level of automation [5, 6].

On-site operation management has always been a weak link in the management of power grid construction enterprises. Shadare SA explores emerging smart grid technologies such as electronic power interfaces, power line communications, alternating current switches (FACTS) and communications equipment, with a focus on issues related to electrical interference. They also provide some available EMC/EMI reduction procedures [7]. Nguyen T discusses these important guidelines and recent improvements from the perspective of different detection mechanisms, machine security concepts, and mitigation strategies to improve the capabilities of infrastructure and service innovations against cyberattacks. This capability is important because even modest improvements in cyber resilience to cyber threats can lead to substantial savings and rich general benefits [8]. In the construction of the power grid infrastructure, the application of the Internet of Things technology can realize the efficient collection of information points such as personnel and materials in the process of infrastructure construction, and then realize the intelligent management of the power grid [9].

This paper is a research on the application of domestic AI chips of Internet of Things technology and artificial intelligence algorithms in the construction management of power grid infrastructure. Although the management of IoT tag chips is only a technology in the construction management of power grid infrastructure, it has a great impact on the smooth development of the entire project. Pivotal role. On the basis of analyzing the current situation of power grid infrastructure construction management and sorting out the management process, combined with the Internet of Things technology, this paper proposes an informatization construction plan for power grid infrastructure construction management, and combines specific applications to design a power grid infrastructure construction management system based on the Internet of Things. The informatization of infrastructure construction management provides a set of feasible solutions.

## **2 Research on the Application of Domestic AI Chips Based on Internet of Things Technology and Artificial Intelligence Algorithms in Power Grid Infrastructure Construction**

### **2.1 Key Technologies of IoT Tag Chips**

#### **(1) Offset voltage elimination**

At present, methods for eliminating the DC offset voltage of a comparator mainly include the following two voltage storage technologies, an output offset storage technology and an input offset storage technology. Output offset cancellation is to store the output of the differential comparator when the differential input is zero on a capacitor in series with the comparator output; input offset cancellation is to store the comparator's DC offset on a capacitor in series with the comparator input [10, 11].

In the design, the offset voltage is eliminated by using the characteristic that the voltage across the capacitor cannot be abruptly changed. To design a reasonable switched capacitor circuit, first connect the capacitor to the circuit. At this time, the offset voltage will be loaded on the capacitor. In the next working stage, the offset voltage on the capacitor can be offset by the characteristic that the voltage across the capacitor cannot be abruptly changed, so as to achieve the purpose of eliminating the offset. [12, 13].

(2) Low power consumption design

For digital circuit design, methods such as reducing the output load capacitance, reducing the power supply voltage, reducing the operating frequency, and reducing the signal inversion frequency can effectively reduce the power consumption [14]. The design process and performance indicators determine the operating frequency and power supply voltage of the circuit, so the focus of low-power design should be considered from reducing the signal flip frequency and load capacitance [15, 16]. When designing the circuit layout, the load capacitance can be reduced by reasonable layout, optimizing the relative position of the devices, and shortening the circuit traces. The method of gated clock can effectively reduce the signal flip frequency [17, 18].

## 2.2 Localized Chips

Semiconductor technology is applied to various fields, and the global market is huge. According to statistics, there is a huge demand for chips. Nearly half of the chips are consumed by my country's market. Among them, chips represented by ARM series microprocessors account for about 90% of my country's market share. There is a great security risk in this situation. In order to obtain my country's commercial, military or other confidential information, chip manufacturers will write malicious code and install Trojan viruses on the chip to realize the control of computers, control systems and other important equipment. It is very easy to cause security problems such as data leakage. Therefore, long-term dependence on foreign chips is not a long-term strategy, which will have a major impact on national security. This paper uses the Loongson 1B chip developed by the Institute of Computing Technology, Chinese Academy of Sciences, which is independently developed and designed by my country, and its performance is between the ARM7 series and ARM9 series chips. The security measurement and control terminal completes the data acquisition and power calculation, and there are no harsh conditions for the performance of the core processor, so the Loongson 1B chip can meet the requirements of the design of the paper. At the same time, the use of domestic chips can promote my country's continuous improvement and innovation in chip technology research.

### **2.3 The Overall Technical Framework Design of the Application of RFID Technology in Power Grid Construction Management**

- (1) Personnel attendance registration management  
The application of RFID technology can realize the identification and positioning of personnel entering the construction site. From the time when the personnel enter the site gate to the time they leave the site, the RFID tag is the only identification.
- (2) Electronic history management of machinery and equipment  
In the management of power grid construction, RFID technology is introduced to assign a set of electronic tags to each machine and equipment to record its basic information and real-time operation status information, so as to monitor it in real time, which greatly simplifies the management process and eliminates existing problems in the management process. In some blind areas, from passive management to active supervision, it provides timely and accurate data for the operation and maintenance of equipment, thereby maximizing the utilization rate of equipment.
- (3) Real-time location reporting  
Using the RTLS communication module of RFID technology, combined with the current mainstream android mobile intelligent terminal advanced technology, it can complete the function of real-time viewing of the positioning target position, activity trajectory and distribution.

## **3 Investigation and Research on the Application of Domestic AI Chips in Power Grid Infrastructure Construction Based on Internet of Things Technology and Artificial Intelligence Algorithms**

### **3.1 System Development Platform**

In order to better integrate the power grid infrastructure management and control system with the SAP system, the SAP Net Weaver platform based on the J2EE architecture is preferred for development in the technology selection, and the system is also required to be developed with the same SDK version of Java technology to ensure Application Consistency. (Because this project is a complex enterprise-level application, it has very high requirements on the reliability, performance and availability of the system).

### **3.2 Security Unit**

The SM4 encryption algorithm is a commercial block cipher published by the State Cryptography Administration of China. The block length and key length are 128Bit. Both data encryption and key expansion use a 32-round non-linear iterative structure. The structure of encryption and decryption is the same and the same key is used.. The SM4 algorithm is improved and designed on FPGA and applied to Ethernet communication.

For the input 128Bit message, first divide it into four 32Bit data, set it as (X0, X1, X2, X3), and the output ciphertext is (Y0, Y1, Y2, Y3). For the i-th round of encryption

transformation, the operation process is as follows:

$$\begin{aligned} X_{i+4} &= F(X_i, X_{i+1}, X_{i+2}, X_{i+3}, rk_i) \\ &= X_i \oplus T(X_{i+1} \oplus X_{i+2} \oplus X_{i+3} \oplus rk_i), i = 0, 1, \dots, 31 \end{aligned} \quad (1)$$

After 32 iterations, the output is:

$$(Y_0, Y_1, Y_2, Y_3) = (X_{35}, X_{34}, X_{33}, X_{32}) \quad (2)$$

where  $\oplus$  represents the XOR operation;  $rki$  is the round key, which is generated by the initial key through the key expansion algorithm operation;  $T(\cdot)$  is the synthetic permutation function, which is composed of the nonlinear transformation  $\tau$  and the linear transformation  $L$ .

In practical applications, the SM4 encryption algorithm is mainly used for the encryption and decryption of a large amount of data, and requires high real-time performance. Therefore, the SM4 hardware acceleration engine integrated in the security chip adopts a parallel design, and the hardware circuit only needs 35 clock cycles to complete an encryption and decryption operation. The input signal of SM4 acceleration engine is 256Bit input data and 32Bit control signal; the output signal is 256Bit output result and 32Bit status signal.

## 4 Analysis and Research on the Application of Domestic AI Chips Based on Internet of Things Technology and Artificial Intelligence Algorithms in Power Grid Infrastructure Construction

### 4.1 FPGA Implementation of Security Chip

Create an FPGA project through Vivado, and import all the code files of the designed security chip SoC into the project. When the SoC is transplanted to the FPGA platform, the memory size is modified to 32KB. In addition, because the PUF circuit requires complete symmetry of layout and routing, it is difficult to achieve complete symmetry of layout and routing in FPGA due to the limitation of slice position and trace length. However, through manual constraints, the symmetry of MUX layout can be achieved. There are two common methods for manually constraining the MUX. The first is to specify the specific location of the MUX through the XDC constraint file, and the second is to write a tcl script to implement the constraints of the MUX. This paper adopts the first method to constrain the delay arbitration module of the PUF circuit.

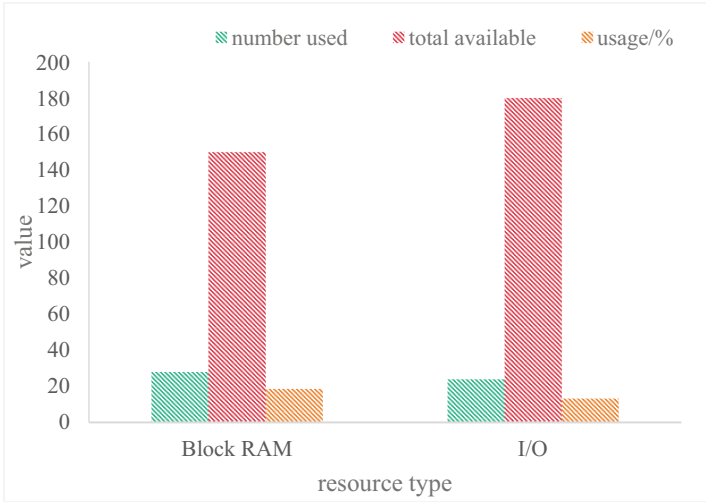
The rest of the modules except PUF are automatically completed by Xilinx's synthesis and place-and-route tools. After the design is implemented, the resources of the FPGA occupied by the entire SoC are shown in Table 1, and the occupation of Block RAM and I/O is shown in Fig. 1.

### 4.2 Application of Power Grid Infrastructure Construction Management System

The system consists of a data center, a construction site positioning base station and a radio frequency identification card. The construction site base station is equipped with

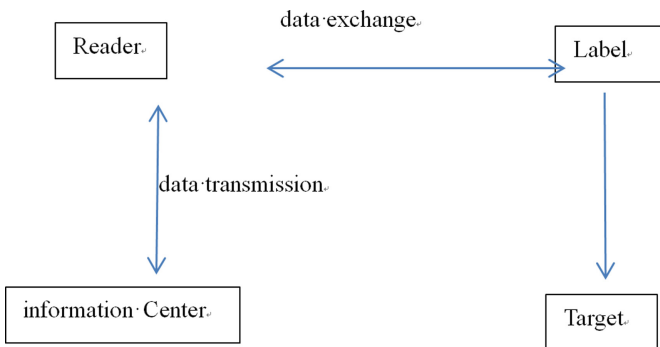
**Table 1.** FPGA resource usage table

Resource Type	number used	total available	usage/%
LUT	28641	38500	74.3
Flip-Flop	16874	100000	16.8
Block RAM	28	150	18.6
I/O	24	180	13.3



**Fig. 1.** Block RAM and I/O usage

an RFID reader. The radio frequency identification card is an RFID electronic tag, which is worn on the construction site staff or attached to the mechanical equipment on site. Through the SPOTON signal strength algorithm, the specific location of the electronic tag can be calculated to achieve positioning. The positioning model is shown in Fig. 2.



**Fig. 2.** RFID positioning system model diagram

Combining RFID technology with the management system, a regional engineering information traceability card is set up on the construction site. By scanning the traceability card, the daily quality and safety inspection information on the construction site can be extracted from the system and displayed to the terminal. Inspection information, as well as the construction risk level of the area, can be traced back to the inspector information.

## 5 Conclusions

Power grid infrastructure construction management is the highlight of the daily construction business of power companies, which not only reflects the management level of a company to a large extent, but also relates to whether the strategic goals of infrastructure construction can be successfully achieved. The system in this paper has been deployed and applied in the company's power grid construction infrastructure project management site. The APP has been installed and deployed in the Android mobile devices of all construction site personnel, and the server function has been implemented in the company's information room internal network. The operation of the system has reached the expected development requirements and goals, solved the problems and deficiencies in the current business management, realized the digitalization and automatic management of the company's power grid construction infrastructure construction management, and improved the company's construction management efficiency and automation level.

## References

1. Bhagwatikar, G.: IoT-based implementation of field area network using smart grid communication infrastructure. *Smart Cities* **1**(1), 176 (2018)
2. Paladino, J.: Electricity infrastructure: the application of grid architecture [in my view]. *IEEE Power Energy Mag.* **17**(5), 100–198 (2019)
3. Raveendran, V., Shanthisree, S.W., Swathy, K., et al.: Vehicle-to-grid ancillary services using intelligent green electric vehicle charging infrastructure in smartgrid. *Int. J. Power Energy Syst* **40**(1), 18–28 (2020)
4. Tomaszewski, M., Ruszczak, B., Michalski, P., et al.: The study of weather conditions favourable to the accretion of icing that pose a threat to transmission power lines. *Int. J. Crit. Infrastruct. Prot.* **25**(JUN), 139–151 (2019)
5. Marten, A.K., Akmatov, V., Sørensen, T.B., et al.: Erratum: Kriegers flak-combined grid solution: coordinated cross-border control of a meshed HVAC/HVDC offshore wind power grid. *IET Renew. Power Gener.* **12**(13), 1493–1499 (2018)
6. Sun, C.C., Hahn, A., Liu, C.C.: Cyber-physical system security of a power grid: state-of-the-art. *Int. J. Electr. Power Energy Syst.* **99**(JUL), 45–56 (2018)
7. Shadare, A.E., Sadiku, M., Musa, S.M.: Electromagnetic compatibility issues in critical smart grid infrastructure. *IEEE Electromagn. Compat. Mag.* **6**(4), 63–70 (2018)
8. Nguyen, T., Wang, S., Alhazmi, M., et al.: Electric power grid resilience to cyber adversaries: state of the art. *IEEE Access*, **PP**(99), 1 (2020)
9. Cohn, J.: When the grid was the grid: the history of North America's brief coast-to-coast interconnected machine. *Proc. IEEE* **107**(1), 232–243 (2018)



10. Marimuthu, K.P., Durairaj, D., Srinivasan, S.K.: Development and implementation of advanced metering infrastructure for efficient energy utilization in smart grid environment. *Int. Trans. Electr. Energy Syst.* **28**(3), e2504.1-e2504.15 (2018)
11. Mueller, C.E., Keil, S.I.: Measuring perceived procedural fairness in the context of power grid expansion: a Rasch modeling approach. *Int. J. Energy Sect. Manage.* **14**(1), 85–107 (2020)
12. Parizy, E.S., Choi, S., Bahrami, H.R.: Grid-specific co-optimization of incentive for generation planning in power systems with renewable energy sources. *IEEE Trans. Sustain. Energy* **11**(2), 947–957 (2020)
13. Soltan, S., Zussman, G.: EXPOSE the line failures following a cyber-physical attack on the power grid. *IEEE Trans. Control Netw. Syst.* **6**(1), 451–461 (2019)
14. James, A.P.: A hybrid memristor–CMOS chip for AI. *Nat. Electron.* **2**(7), 268–269 (2019)
15. Parker, K.: AI on a chip; AI in the cloud. *Consult. Specifying Eng.* **55**(3), a4 (2018)
16. Kani H T, Ergenc I, Polat G, et al. Evaluation of endoscopic mayo score with an artificial intelligence algorithm. *Journal of Crohn s and Colitis*, 2021, 15(Supplement\_1): S195-S196
17. Graewingholt, A., Rossi, P.G.: Retrospective analysis of the effect on interval cancer rate of adding an artificial intelligence algorithm to the reading process for two-dimensional full-field digital mammography. *J. Med. Screen.* **28**(3), 369–371 (2021)
18. Morozov, S.P., Chernyaeva, G.N., Bazhin, A.V., et al.: Validation of diagnostic accuracy of anartificial intelligence algorithm for detecting multiple sclerosis in a city polyclinic setting. *Diagn. Radiol. Radiother.* **11**(2), 58–65 (2020)



# Data Compression Algorithm of Power System

Xinyu Zhang<sup>(✉)</sup>, Yanduo Xia, Shuangwu Quan, and Guozhi Zhang

Dagang Oilfield Electric Power Company, Tianjin 3002800, China

xinwei6444@163.com

**Abstract.** Data compression refers to the use of software performance to improve system utilization without increasing hardware costs. Therefore, in the process of transmitting a large amount of power data, only an appropriate data compression algorithm can reduce the cost of data storage and communication and meet the user's requirements for high-speed data transmission. Aiming at the reconstruction and coding of power data signals, this paper proposes a Huffman compression algorithm, which can compress power energy signals and power quality disturbance signals. As a result, experiments show that the mean square error of the Huffman compression algorithm is small, the signal-to-noise ratio is large, and the compression effect is good. In view of the characteristics of redundant information of power data, this paper proposes a dictionary-based compression algorithm to compress it by searching for redundant feature formats. The advantage of this data compression method is that it can be well compressed according to the characteristics of redundant information in power system data format and units, and the compressed text can keep good format and original useful data. The measured compression ratio of this algorithm is not less than 30%. On the premise that the compression effect is good, the amount of program consumed is relatively small compared to the entire test data, and it is easy to implement on the hardware of the store's data processing system. Data compression reduces the occupation of data storage space, thereby reducing the time required to transmit data.

**Keywords:** Power System · Data Compression · Huffman Compression Algorithm · Dictionary-Based Compression Algorithm

## 1 Introduction

A large amount of power data is stored in the power system. How to achieve efficient power data transmission and increase the storage space of the power system has become an urgent problem to be solved. The data compression algorithm can compress huge data files into smaller files. The compressed data is consistent with the original data, but the size of the data file is changed. Therefore, this paper analyzes the power system data compression algorithm to improve the data transmission speed, so that the power system can store more data.

At present, many scholars have conducted in-depth research on the data compression algorithm of power system, and achieved good research results. For example, some researchers use the Huffman coding algorithm to reconstruct the power data signal, and

the results show that if the coding level of the algorithm is increased, the power data signal can be compressed to allow the power system to store more data for evaluation. The compression ratio index of the Huffman coding algorithm will also increase. The algorithm can extract signal features from the reconstructed power data signal, and the amount of calculation increases. At the same time, with the increase of the number of coding and decoding stages, the signal-to-noise ratio increases., the mean square error after signal reconstruction is reduced. In practical applications, it is necessary to consider the three trade-offs. In the application, the number of encoding and decoding stages can be flexibly controlled according to the needs, and irrelevant information in the data is eliminated [1, 2]. Some experimental studies have shown that the GPRS data transmission speed is related to the size of the transmitted data. When the transmission data is less than 50KB, the data transmission speed is fast, which is ideal. Therefore, in practical applications, when the transmission data is slightly larger, in order to ensure the real-time nature of data transmission., it needs to be effectively compressed in order to improve the efficiency of data transmission [3, 4]. Although the research on power system data algorithms has achieved good results, there are many data compression algorithms, but each algorithm has its own advantages and disadvantages. In order to achieve efficient transmission of power data in the transmission process, it is necessary to further improve the data The performance of the compression algorithm.

This paper first focuses on two common data compression algorithms, namely Huffman compression algorithm, dictionary-based compression algorithm, and other compression algorithms, which are not described in detail in this paper. Then, the compression performance evaluation index is proposed, and then the power data processing system is built. Finally, the simulation experiments are carried out on the two compression algorithms proposed in this paper, and the experiments have verified that the compression effects of the two algorithms are good.

## 2 Data Compression Algorithm and Algorithm-Related Performance Evaluation Indicators

### 2.1 Common Data Compression Algorithms

#### (1) Huffman compression algorithm

The core of Huffman coding is to construct a Huffman tree according to the probability of occurrence of each character, and the codeword of each character can be obtained according to the depth of the tree where each character is located [5]. The performance of the Huffman encoding algorithm is related to the character length and the number of times it appears in the data string. Generally speaking, the longer the character is and the higher its frequency in the data string, the higher the encoding efficiency, and the better the performance of the Huffman compression algorithm. In this way, the overall symbol length is the smallest, and the purpose of compression is achieved. Huffman encoding each character symbol length is different, so Huffman is a prefix encoding [6].

#### (2) Dictionary-based compression algorithm

Huffman coding needs to obtain the statistical information of the source. The dictionary-based compression method takes a different approach. By establishing

a dictionary for the most recently encoded characters, the characters to be encoded are searched in the dictionary to compress the data [7]. The decompression principle of the dictionary lookup redundant feature format method: traverse the compressed file, when the data separator comma is encountered, the corresponding information is recovered according to the dictionary elements in the compression dictionary, and the subsequent characters are moved back, so that the recovered The text maintains a good format and original power data information [8].

The idea of run-length coding is the simplest example based on a dictionary, such as compressing the character AAAABBBBAAAACCC, and run-length coding compresses it into 4A3B4A3C, using the local repetition of characters for compression. Dictionary-based compression algorithms are widely used in text compression, application compression, kernels of general compression algorithms such as WinZip, Rar, and Gzip. Dictionary-based compression algorithms include LZ77 compression algorithm, LZW compression algorithm [9].

The LZ77 compression algorithm is a classic dictionary-based compression algorithm. The algorithm uses a transformable data window to map the data. It can compress the data file without knowing the original file of the data signal. At the same time, it does not need to know the source in advance. Can be encoded. In some respects, dictionary-based compression algorithms can not only compress data directly, but also compress text, images, etc., with a larger range than other compression algorithms [10].

LZW compression algorithm is a widely used compression algorithm based on dictionary compression algorithm. The advantage is that this method only needs to store a relatively small table to restore the corresponding value when storing the data, so the required cost is relatively low, and the encoding and decoding speed of the LZW algorithm is very fast. When the data repetition is large, the compression ratio of the LZW algorithm is better than the compression algorithm based on statistics [11].

(3) Other compression algorithms

Other compression algorithms include wavelet transform, arithmetic coding, Fourier transform, etc., which are not introduced in this article.

## 2.2 Compression Performance Evaluation Index

The evaluation indicators for compressing electrical energy signals include signal compression ratio, mean square error, and signal-to-noise ratio.

(1) The compression ratio of the signal  $R_{CR}$

$$R_{CR} = \frac{R_1}{R_2} \times 100\% \quad (1)$$

$R_1$  represents the data volume of the signal to be stored after compression, and  $R_2$  represents the data volume of the original signal.

(2) Mean square error  $e_{MSE}$

$$e_{MSE} = \frac{\sqrt{\sum_{i=0}^{N-1} [g(i) - h(i)]^2}}{\sqrt{\sum_{i=0}^{N-1} [g(i)]^2}} \times 100\% \tag{2}$$

The larger the  $e_{MSE}$ , the less distortion of the AA signal.

(3) Signal-to-noise ratio  $r_{SNR}$

$$r_{SNR} = 10\log_{10} \left\{ \frac{\sum_{i=0}^{N-1} [g(i)]^2}{\sum_{i=0}^{N-1} [g(i) - h(i)]^2} \right\} \tag{3}$$

Among them,  $g(i)$  is the original signal sampling point,  $h(i)$  is the signal reconstructed from the compressed data, and  $i = 0, 1, \dots, N-1$ . The larger the  $r_{SNR}$ , the smaller the signal distortion [12].

### 3 Power Data Processing System

This paper proposes a power data processing system with TMS320F2812 as the core. As shown in Fig. 1, the collected power data is input from the RJ-45 interface to the TMS320F2812. TMS320F2812 fixed-point digital signal processor is the main control chip, which mainly completes the realization and control of power data compression algorithm, so as to achieve power data compression processing. In the hardware structure of the system, CPLD is connected with TMS320F2812 and RTL8019AS at the same time, while TMS320F2812 is connected with Flash and SRAM. The two-way connection makes data transmission more convenient, and the final data is output through UART. The system has wide application prospects in power system data compression.

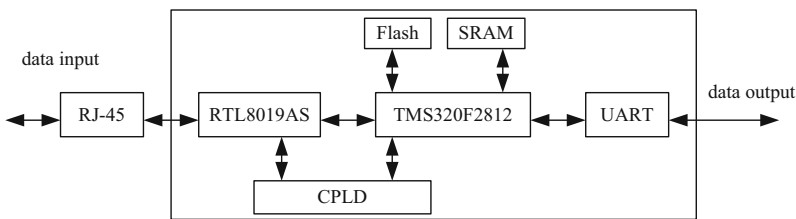


Fig. 1. Hardware structure of power data processing system

RTL8019AS network card hardware driver is mainly composed of initialization program, data sending and receiving driver. The internal current page pointer, boundary pointer, DCR, TCR and other parameters of the RTL8019AS are configured during the initialization process. Among them, including selecting what kind of data packet to receive, setting the buffer threshold range of FIFO, determining the size of the sending and receiving buffer, which interrupt service mode to use, and the determination of the physical address.

The data sending process is mainly by starting the remote DMA and the local DMA, among which, the RTL8019AS automatically completes the data information transmitted from the wire. There are two ways to transfer the data in the RTL8019AS receiving buffer to the memory of the DSPF2812. The first way is that for the data packet corresponding to the boundary pointer BURY, the remote DMA automatically sends it away. After the transmission is successful, the NIC removes the boundary pointer BURY and releases part of the buffer area for continued reception. Because this method is inconvenient to control read and write operations, it is generally not used. The second way is to move the boundary pointer BURY through human operation.

## 4 Simulation Analysis of Power System Data Compression Algorithm

### 4.1 Experimental Simulation of Huffman Compression Algorithm

The 5 samples of the electrical energy signal are compressed by the Huffman compression algorithm, and the performance of the compression algorithm is measured by three indicators: compression ratio, normed mean square error after signal reconstruction and signal-to-noise ratio. Table 1 shows the comparison results of the compression ratio, normed mean square error and signal-to-noise ratio obtained by using the Huffman compression algorithm.

**Table 1.** Energy Signal Compression Results

	compression ratio	mean squared error	Signal to noise ratio
1	1.5638%	1.3192%	37.5217
2	3.5754%	0.8967%	42.8494
3	12.6823%	0.3281%	49.7265
4	17.3076%	0.1895%	53.9244
5	26.6172%	0.07483%	64.1556

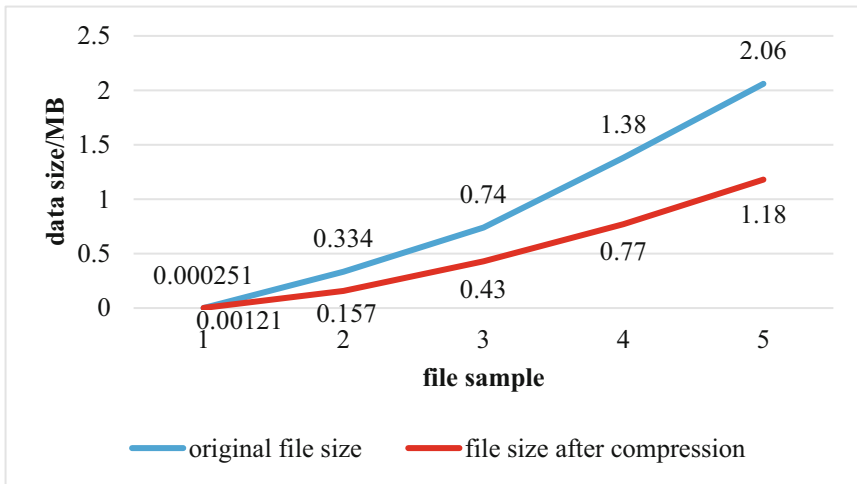
As can be seen from the data in Table 1, using the Huffman compression algorithm to compress the electrical energy data can achieve a good compression effect: first, a high compression ratio can be achieved, and even in the case of a high compression ratio, it can still maintain High signal quality; secondly, as the compression ratio increases, the superiority of the algorithm's compression effect becomes more and more prominent. The Huffman compression algorithm not only compresses relatively small but also has a large signal-to-noise ratio. With the increase of the compression ratio, the normed mean square The error is relatively small, and the signal-to-noise ratio will increase; another outstanding advantage is that the compression ratio can be freely changed.

### 4.2 Experiment Simulation of Dictionary-Based Compression Algorithm

In order to more comprehensively check the compression effect of the dictionary lookup redundant feature format compression algorithm, the front-end data acquisition module is used to receive a large amount of power system data. A total of 5 test files are selected, and the data volume is from small to large, ranging from a few B to a few kB. To a few MB, through the change of the number, the compression effect can be expressed comprehensively and concretely. The five test files are from different time periods respectively, and the above data is compressed and tested on the computer. The file size before and after compression is shown in Table 2 and Fig. 2. In Fig. 2, the file unit is converted (1 kB = 1021 B, 1 MB = 1024 kB).

**Table 2.** Comparison of file sizes before and after compression

original file size	file size after compression	compression ratio
263 B	127 B	38.25%
342 kB	161 kB	36.49%
0.74 MB	0.43 MB	35.61%
1.38 MB	0.77 MB	33.07%
2.06 MB	1.18 MB	31.82%



**Fig. 2.** Comparison of original file and compressed file size

As can be seen from Table 2, after using the dictionary lookup redundant feature format algorithm to compress, when the compression ratio is not less than 30%, the original file size of 263 B is compressed to 127 B, and the original file size is 342 kB. The compressed file size is 161 kB, the original file size of 0.74 MB is compressed to

0.43 MB, the original file size of 1.38 MB is compressed to 0.77 MB, and the original file size of 2.06 MB is compressed to 1.18 MB, which is the same as the original file size. Compared with the amount of data, the compression effect is obvious. And the data compression process can be realized on the power data processing system with TMS320F2812 as the core constructed in this paper, and it is also suitable for power system data compression. After testing, the compression results of this algorithm have reached the expected design compression results, and achieved the purpose of saving storage space and facilitating power data transmission.

## 5 Conclusions

This paper proposes two data compression algorithms, namely Huffman compression algorithm and dictionary-based compression algorithm, and conducts power system data compression simulation experiments on them respectively. In the Huffman compression algorithm simulation experiment, three compression algorithm performance evaluation indicators are used to evaluate the compression effect of the Huffman compression algorithm on the electrical energy signal. The results show that the Huffman compression algorithm has a small compression ratio, a small mean square error, and a large signal-to-noise ratio. The compression algorithm has a good compression effect; in the dictionary-based compression algorithm simulation experiment, comparing the size of the original file and the compressed file, the results show that the dictionary-based compression algorithm also has a good compression effect, and the algorithm can perform data compression on the power data processing system.

## References

1. Sarkar, S.J., Kundu, P.K., Sarkar, G.: Development of lossless compression algorithms for power system operational data. *IET Gener. Transm. Distrib.* **12**(17), 4045–4052 (2018)
2. Sivanandam, L., Periyasamy, S., Oorkavalan, U.M.: Power transition X filling based selective Huffman encoding technique for test-data compression and Scan Power Reduction for SOCs. *Microprocess. Microsyst.* **72**(Feb), 102937.1-102937.10 (2020)
3. Ashraf, M., Mostafa, H., Eladawy, A., et al.: Power adaptive high-resolution neural data compression algorithm (PANDCA). *Microelectron. J.* **88**(Jun), 154–163 (2018)
4. Crispin-Bailey, C., Dai, C., Austin, J.: A 65-nm cmos lossless bio-signal compression circuit with 250 FemtoJoule performance per bit. *IEEE Trans. Biomed. Circuits Syst.* **13**(5), 1087–1100 (2019)
5. Acharya, S., Demarco, C.L.: Low-loss image-based compression for synchrophasor measurements. *IET Gener. Transm. Distrib.* **14**(20), 4571–4579 (2020)
6. Sarkar, S.J., Kundu, P.K., Sarkar, G.: Development of cloud-based power system operational data management system. *Gener. Transmiss. Distrib. IET* **13**(5), 644–651 (2019)
7. Tsai, T.H., Tung, N.C., Lin, D.B.: VLSI implementation of multi-channel ECG lossless compression system. *IEEE Trans. Circuits Syst. II: Expr. Briefs* **PP**(99), 1 (2021)
8. Sarkar, S.J., Kundu, P.K., Sarkar, G.: Development of DBEA compressed data transfer system over power line. *Modell. Measur. Control A* **90**(3), 262–278 (2018)
9. Alvarez, G., Favaro, F., Lecumberry, F., et al.: Wireless EEG system achieving high throughput and reduced energy consumption through lossless and near-lossless compression. *IEEE Trans. Biomed. Circuits Syst.* **PP**(99), 231–241 (2018)



10. Tsigkanos, A., Kranitis, N., Theodoropoulos, D., et al.: High-performance COTS FPGA SoC for parallel hyperspectral image compression with CCSDS-1230-B-1. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **28**(11), 2397–2409 (2020)
11. Assche, J.V., Gielen, G.: Power efficiency comparison of event-driven and fixed-rate signal conversion and compression for biomedical applications. *IEEE Trans. Biomed. Circuits Syst.* **PP**(99), 1 (2020)
12. Tsai, T.H., Hussain, M.A.: VLSI implementation of lossless ECG compression algorithm for low power devices. *IEEE Trans. Circuits Syst. II: Expr. Briefs* **PP**(99), 1 (2020)



# Bank Credit Risk Prediction Based on MCDM and CNN

Yuyue Cong<sup>1,2</sup>(✉)

<sup>1</sup> Xi'an Eurasia University, Xi'an 710065, Shaanxi, China  
cider999@163.com

<sup>2</sup> Xidian University, Xi'an 710126, Shaanxi, China

**Abstract.** Bank bankruptcy will trigger a credit crisis and produce a series of chain reactions, which have great destructive power to a country's economy. Therefore, governments of various countries have implemented strict financial controls on commercial banks, and at the same time have strengthened the research on credit risk identification and control. The purpose of this paper is to study bank credit risk prediction based on MCDM and CNN. Apply convolutional neural network to bank credit risk prediction to serve commercial banks. Through continuous optimization of the algorithm, a more accurate and efficient credit risk identification model is obtained, and the combined algorithm model of CNN and XGBoost is combined with specific practices to give specific practical solutions. At the same time, the MCDM method is used to build a multi-space multi-criteria model comprehensive evaluation framework, and then the evaluation object is predicted and evaluated. The experimental results show that the accuracy rate of the CNN and XGBoost combined algorithm model in predicting non-overdue repayments reaches 90%, and the model application effect Better.

**Keywords:** Bank Credit · Risk Prediction · Convolutional Neural Network · MCDM Method

## 1 Introduction

Credit risk is one of financial risk in financial markets. This is a major problem faced by modern socio-economic entities (especially financial institutions), investors and consumers [1, 2]. It directly affects various activities in modern economic life, and even affects the stable growth of the world economy. Therefore, understanding the credit risk, analyzing the reasons for the formation of credit risk, and preventing credit risk have become an important topic in the field of modern economy and finance [3, 4].

According to their own national conditions, countries have successively improved and innovated the original measurement methods according to their own economic characteristics, in order to find a credit risk measurement method suitable for their own national economic environment, and then improve the management of credit risk [5, 6]. Riyazahmed K empirically examines the impact of managerial efficiency on credit risk in Indian public and private sector banks. Think of return on assets as a proxy for management effectiveness and gross non-performing assets (GNPA) as a proxy for credit

risk. The study used fixed effects and dynamic panel data models to examine effects. Econometric model estimates suggest that asset returns have a negative impact on credit risk. In addition, the impact of asset returns is analyzed through information on microeconomic and macroeconomic variables in the dynamic Generalized Method of Moments (GMM) method. The results remain unchanged after using a dynamic GMM modeled with lagged credit risk and lagged asset returns [7]. Okoli T T uses panel ARDL (1, 1) ensemble average group, average group and dynamic fixed effect estimator to study the bank credit risk, measured by non-performing loans and total loans (NPL) 1995-BRICS. This study makes a new contribution to the fintech data series by using principal component analysis to generate fintech indicators [8]. Therefore, for my country, which is in the critical transition period of the economy, the research on the credit risk management of listed companies has not only stayed at the theoretical level, but has more practical significance, which can provide information for the market practice activities of all participants in the market economy. Important guidance [9, 10].

Based on the observation of the current credit risk forecasting abroad, combined with the actual economic situation and conditions in my country, this paper finds that the convolutional neural network in the credit risk forecasting model has the advantages that are most in line with my country's national conditions compared with other models. Therefore, the model was adjusted under the premise of the current economic situation in my country, and the combined algorithm model of CNN and XGBoost was proposed to make it conform to my country's national conditions, which is more practical for my country's commercial banks in terms of credit risk management. It can not only improve the credit risk disclosure ability of my country's commercial banks, but also improve the credit risk management system, and finally achieve the purpose of scientific management of commercial banks.

## **2 Research on Bank Credit Risk Prediction Based on MCDM and CNN**

### **2.1 Convolutional Neural Networks**

Convolutional neural network is inspired by the visual cortex cells of animals. CNN is a kind of forward propagation neural network, which has been widely used in the field of graph and video recognition [11]. The classic CNN model structure has the following parts: a fully connected layer, a pooling layer for pooling operations, an input layer for input data, and a convolution layer for convolution operations. The most important part of the convolutional neural network is its convolutional layer. Between different layers of the network, the convolution kernel plays a crucial role, and the input of the previous layer is formed by the convolution of a set of weights. Input to the next layer. Moreover, the convolutional neural network is more powerful than the traditional filter in that its parameters (weights and thresholds) can be self-trained, modified and optimized according to the algorithm [12, 13].

### **2.2 CNN and XGBoost Combined Algorithm Model**

The algorithm combination model proposed in this paper is built on the basis of convolutional neural network, which is similar to the structure of convolutional neural network,

including input layer, convolutional layer, pooling layer, output layer, and fully connected layer [14]. Since the feature dimension of loan users is small, the combined model adopts two convolutional layers and two pooling layers. The input layer, two convolution layers, two pooling layers, the first fully connected layer and the XGBClassifier() function constitute the combined model of the CNN and XGBoost algorithms. The XGBClassifier() function belongs to the XGBoost algorithm module, and its main function is to predict the loan user's repayment behavior according to the characteristics of the loan user.

XGBClassifier() function: This function belongs to the XGBoost algorithm module. First, the function is fitted between the training data and the label, and then the label is predicted to give the label of the test data, which is mainly used to solve the two-class problem [15].

## 2.3 MCDM Method

### (1) TOPSIS method

The TOPSIS method is a multi-criteria decision-making method under a limited scheme. This method can integrate a variety of decision-making factors, and at the same time consider the decision-maker's preference for each decision-making factor, and evaluate and rank the alternatives on this basis. [16, 17]. TOPSIS is also called the approximation ideal solution sorting method. Its basic idea is very simple. First, an optimal solution and a worst solution are calculated according to the alternative solutions, and then the distance between each solution and the optimal solution and the worst solution is calculated. Score the alternatives, and finally complete the sorting of the solutions. If a solution is closer to the optimal solution, then the solution is better.

Compute the distance:

$$L_i^+ = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^+)^2}, i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (1)$$

$$L_i^- = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^-)^2}, i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (2)$$

where  $L_i^+$  represents the distance from the positive ideal solution, and  $L_i^-$  represents the distance from the negative ideal solution.

To calculate the closeness, the closeness in the TOPSIS method is defined as follows:

$$C_i = \frac{L_i^-}{L_i^+ + L_i^-}, i = 1, 2, \dots, m \quad (3)$$

It can be seen from the above formula that the value range of  $C_i$  is:  $0 \leq C_i \leq 1$ .

When  $L_i^- = 0$ ,  $C_i = 0$ ,  $A_i = A_i^-$  is the negative ideal solution.

When  $L_i^+ = 0$ ,  $C_i = 1$ ,  $A_i = A_i^+$  is a positive ideal solution.

## (2) Analytic Hierarchy Process

AHP can help decision makers choose a plan that satisfies their own decision preferences in complex decision problems. It belongs to the multi-attribute decision-making method in MCDM. The basic idea of analytic hierarchy process to solve the problem is: firstly, stratify the target problem. Consider the key evaluation criteria for the final plan selection, and then decompose the evaluation criteria to establish a hierarchical relationship. The lower-level elements must have a subordinate relationship with the upper-level elements to form a multi-layer analysis structure model. The relative importance of each criterion to the target and the relative importance of each scheme relative to the criterion are determined by mutual comparison, and a paired comparison matrix is formed. After passing the consistency test, the eigenvector corresponding to the eigenvalue of the paired comparison matrix is calculated to determine The weight of each layer of elements; the weight of each layer of elements is comprehensively calculated, and the weight of each scheme of the scheme layer relative to the target layer is obtained, so as to select the appropriate scheme.

## 3 Investigation and Research on Bank Credit Risk Prediction Based on MCDM and CNN

### 3.1 Acquisition of Data

The data in this article is derived from the loan user dataset provided by Bank M, which includes six data tables of user information, bank records, browsing information, bill records, loan time and overdue records. The characteristics of this dataset are: the user's features have been preliminarily numericalized, the features cover a wide range of fields, and the number of features is large. The combination algorithm model constructed in this paper is used to predict whether the loan user's repayment is overdue. In the actual situation, the user to be tested is a small sample data set. In order to better explain it, the `traintestsplitt()` function in the `sklearn` module is used to select 60 users from the data as the users to be predicted, and then use the constructed combination. Algorithmic models predict their repayment behavior.

### 3.2 Feature Selection

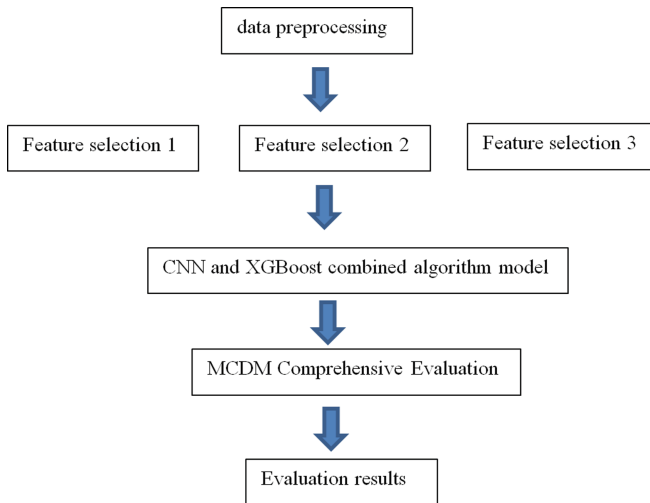
In order to understand the degree of correlation between the characteristics of the loan user's bank record table, draw the number of income ( $n_1$ ), total income ( $n_2$ ), number of expenditures ( $n_3$ ), total expenditure ( $n_4$ ), number of wages ( $n_5$ ), and wages in the bank record table. The heat of the six features of total ( $n_6$ ), in which the correlation between the number of income and total income, the number of expenditures and total expenditure, the number of wages and total wages is between  $[0.99, 1]$ , so the relationship between income, expenditure, wages and total is between  $[0.99, 1]$ . Strong correlation. The correlation between other features and features is below 0.9, so we will not do anything about them for the time being. In order to eliminate the features with greater correlation, this paper constructs three new features, namely average expenditure, average salary, and average income, to replace them. For the original six features, this paper

uses feature selection technology to screen out the important features that affect the repayment behavior of loan users from the user feature set of the correlation analysis, thereby reducing the data dimension. Through feature selection, the feature index of the system can be optimized, and some features with small influence value are eliminated, which can reduce the computational workload and help the credit industry to obtain the effective features of borrowers.

## 4 Analysis and Research on Bank Credit Risk Prediction Based on MCDM and CNN

### 4.1 Multi-space and Multi-criteria Model Comprehensive Evaluation Framework

In this paper, a multi-space multi-criteria model comprehensive evaluation framework is constructed, as shown in Fig. 1.



**Fig. 1.** Comprehensive evaluation framework of multi-space multi-criteria model

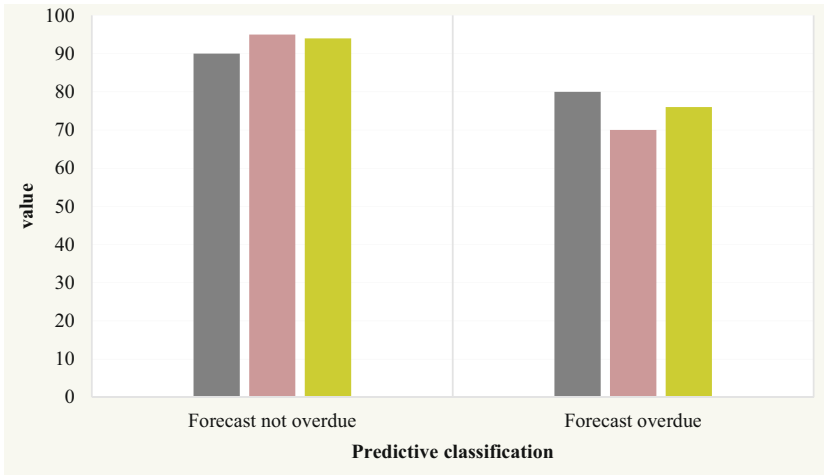
First, the data used in the evaluation process is preprocessed, and then different feature selection methods are used to construct different feature spaces, and then the models are trained and verified in these feature spaces, and the performance of the model on each evaluation criterion is obtained., and finally adopt the MCDM method to set the weight preference for the evaluation criteria according to the problem background of the evaluation model, and synthesize the performance of each model on a single criterion to obtain the final evaluation result of the model.

### 4.2 Application Analysis

According to the predictions of these 60 people, the precision, recall rate, and f1 value are calculated in Table 1.

**Table 1.** Precision, recall, and f1-score values for predicting repayment behavior of users under test

Predictive classification	precision	recall	f1-score
Forecast not overdue	90	95	94
Forecast overdue	80	70	76



**Fig. 2.** Analysis of the effect of the algorithm model

Predicting non-overdue repayments is 90 percent accurate. The accuracy rate of predicting overdue repayment is 80%, indicating that the algorithm model has a better effect on the classification of loan users' repayment behavior, as shown in Fig. 2. The recall rate for predicting non-overdue repayments is 95%, and the combined algorithm works well for actually classifying overdue repayment users. The recall rate of predicted overdue repayments is 70%, which shows that the combined algorithm has a better classification effect on actual overdue users.

## 5 Conclusions

With the rapid development, the bank's credit risk management has not only reformed the original system and institutions, but also put forward new related technical requirements for the identification of risk assessment and the management of credit risk. This paper proposes a customer default risk prediction model based on convolutional neural network for personal credit risk measurement of banks. The customer default risk prediction based on the convolutional neural network model has a high accuracy rate and can accurately evaluate the borrower's credit risk default probability; at the same time, the combined algorithm model of CNN and XGBoost can automatically learn features from the data, which is similar to the artificially designed features. Than can save a lot of time.

**Acknowledgements.** Shaanxi Province Education Science "14th Five-Year Plan" 2021 Project SGH21Y0389; School-level Key Course "Financial Risk Management" Funding Project 2018KC023.

## References

1. Gubareva, M., Borges, M.R.: Rethinking economic capital management through the integrated derivative-based treatment of interest rate and credit risk. *Ann. Oper. Res.* **266**(1–2), 71–100 (2018). <https://doi.org/10.1007/s10479-017-2438-y>
2. Fontes, J.C., Panaretou, A., Peasnell, K.V.: The impact of fair value measurement for bank assets on information asymmetry and the moderating effect of own credit risk gains and losses. *Account. Rev.* **93**(6), 127–147 (2018)
3. Cestau, D.: The political affiliation effect on state credit risk. *Public Choice* **175**(1–2), 135–154 (2018). <https://doi.org/10.1007/s11127-018-0519-3>
4. Lassoued, M.: Comparative study on credit risk in islamic banking institutions: the case of Malaysia. *Q. Rev. Econ. Finan.* **70**(NOV), 267–278 (2018)
5. Dargahi, H., Ghasemi, M., Fatollahi, S.: he impacts of bounced checks on economic growth through the banking credit risk channel emphasis on enforcement of laws: provincial panel approach. *J. Res. Econ. Model.* **10**(40), 7–32 (2020)
6. Yi, W.K., Mahmud, M., Suppiah, T.A.: A study on the determinats of credit risk in Malaysian banking system for the period 1996 to 2017. *Int. J. Psychosoc. Rehabil.* **24**(2), 1012–1034 (2020)
7. Riyazahmed, K., Baranwal, G.: Determinants of credit risk. *Int. J. Account. Finan. Rev.* **6**(1), 53–71 (2021)
8. Okoli, T.T.: Is the relationship between financial technology and credit risk monotonic? Evidence from the BRICS economies. *Asian Econ. Finan. Rev.* **10**(9), 999–1011 (2020)
9. Afaneh, M.: The effect of credit risk indicators on the profitability of banks in the Arab Gulf countries. *J. Econ. Adm. Sci.* **26**(118), 1–11 (2020)
10. Deng, A.S., Rono, L., Sang, J.: Credit risk management and the performance of financial institutions in South Sudan. *Mod. Econ.* **11**(11), 1919–1928 (2020)
11. Zivanovic, V.: The impact of changes in the base and precious metals prices on credit risk factors. *Megatrend Revija* **17**(2), 45–64 (2020)
12. Bourgey, F., Gobet, E., Rey, C.: Metamodel of a large credit risk portfolio in the Gaussian copula model. *SIAM J. Finan. Math.* **11**(4), 1098–1136 (2020)
13. Chandra, D., Kandpal, D.V.: Impact of credit risk management on financial performance of banks-a case of Nainital bank. *Test Eng. Manag.* **83**(March-April 2020), 22429–22446 (2020)
14. Kim, J., Kim, S.: Importance Sampling with Splitting for Portfolio Credit Risk. *Communications for Statistical Applications and Methods* **27**(3), 327–347 (2020)
15. Devi, U., Batra, N.: Exploration of credit risk based on machine learning tools. *J. Crit. Rev.* **7**(19), 4698–4718 (2020)
16. Seo, J.Y.: Machine learning in consumer credit risk analysis: a review. *Int. J. Adv. Trends Comput. Sci. Eng.* **9**(4), 6440–6445 (2020)
17. Lulaj, E., Mazreku, I., Dragusha, B.: Credit risk management and measurement econometric and empirical model in the banking system. *Economica* **16**(3), 155–188 (2020)





# Application of NC Machining Technology in Aircraft Manufacturing

Weiwen Ye<sup>(✉)</sup>

School of Mechanical and Electrical Engineering, Guangdong University of Science and  
Technology, Dongguan 523083, Guangdong, China  
ye49510019@163.com

**Abstract.** With the progress and development of modern science and technology, information technology and mechanical manufacturing technology, China's aerospace manufacturing technology also presents a trend of rapid development. Especially with the support of NC machining technology, the efficiency and quality of aircraft manufacturing have also been greatly improved. The surface quality, thin-wall shape, deep cavity The machining accuracy is more and more suitable for the current advanced aircraft manufacturing requirements. NC machining technology is characterized by its high efficiency and high quality, in recent years, with the rapid development of NC process equipment and NC process technology at home and abroad, NC processing technology has been widely used, especially in the manufacturing of aircraft body parts [1]. Therefore, NC machining technology has become an important technical means in China's aircraft manufacturing process, and it is also an important technical support to improve the development level of China's aerospace industry. Based on this, this paper analyzes the application of NC machining technology in aircraft manufacturing.

**Keywords:** NC Machining Technology · Aircraft Manufacturing · Application

## 1 Introduction

With the support of numerical control technology, some important parts of the aircraft can be accurately cut and polished, so that their accuracy and surface quality can meet the requirements of aircraft assembly. Moreover, the automation and intelligence of numerical control processing also greatly save labor costs, not only improve production quality and efficiency, but also reduce production costs, It ensures the economy and safety of aircraft manufacturing, especially in the manufacturing of some complex components. Therefore, it is necessary to pay attention to the application of NC machining technology in aircraft manufacturing.

At present, China's aviation industry is in a period of rapid development. In order to promote the development of national economic construction and consolidate the needs of national defense construction, we are trying to develop and produce a number of modern aircraft that catch up with and surpass the international advanced level, including the batch production of the third generation military aircraft, the development of the fourth

generation military aircraft, and the R & D of passenger aircraft and Dayun aircraft. The development and production of these advanced modern aircraft must have good advanced technical equipment and advanced manufacturing process methods, and the selection of advanced technical equipment and advanced manufacturing process methods must study and analyze the mechanism characteristics of modern aircraft parts [2].

NC machining technology is the technology of using computer program to control machine tools to process components and realize automatic production. The technology includes hardware and software, in which hardware mainly refers to NC machine tools related to machining and related supporting facilities. NC machine tools mainly input the mold parameters to be manufactured into the sensor inside the NC machine tool through the staff, and the sensor can automatically identify and produce, so as to realize automatic and flexible production of mold manufacturing; Software refers to the computer, program coding and other systems that realize NC machining. With the support of coding program, the relevant dimensions and materials of processed products can be input into the machine tool to complete the automatic machining of machinery. [3] Generally speaking, NC machining technology has the following application characteristics: first, automatic production, which greatly reduces the investment of labor cost, avoids the error caused by manual production, and ensures safe production; Second, centralized production. Compared with the traditional decentralized processing technology, NC processing technology can centrally manage each process, and reduce the piecemeal time between processes and the space occupied by various machines. Therefore, this technology has less space requirements in practical application, and can complete large-scale processing in a small space; The third is flexible production, which mainly refers to the more flexible production line, which can greatly adapt to the market production demand [4].

## **2 Influence of NC Machining Technology on the Application of Aircraft Manufacturing**

### **2.1 Impact on Standard Tooling**

In the process of aircraft manufacturing, the manufacturing department will manufacture a variety of different types of standard samples according to its actual needs and process them according to these standard samples. Due to the large scale of the aircraft, the number of parts is relatively large, and even the sizes and models of various parts are quite different, it requires more time, energy and material resources for the manufacture of standard samples, It is not conducive to the improvement of production efficiency. The application of numerical control technology only requires individual standard tooling to coordinate the installation relationship between various aircraft parts, while the tooling of other aircraft standard parts can realize production and manufacturing through numerical control processing technology, which saves a lot of materials and manpower for aircraft manufacturing enterprises. In general, the production cost is greatly reduced, At the same time, the production efficiency and production quality have been greatly improved.

## 2.2 Influence on Die Line Template

In the past, in the process of aircraft manufacturing, the only original basis was the mold line template, so as to realize the processing of various parts. Therefore, the mold line template has also become a fixed part production and manufacturing template in aircraft manufacturing. The manufacturer only needs to produce the corresponding parts according to the mold line template, but this method has great restrictions on the thinking of design. It is increasingly unable to meet the requirements of current aircraft performance. With the application of numerical control processing technology, the processing and upgrading of various aircraft parts has been realized, and its data-based manufacturing method also advantageously promotes the quality and efficiency of aircraft manufacturing [5].

## 3 NC Machining of Typical Parts

To select whether the product is suitable for NC machining, refer to the following principles:

- (1) The geometry is complex, and the surface of the part is mostly curved, and there are many spatial angles;
- (2) Products that are difficult to process with ordinary machine tools, complicate tooling and increase cost;
- (3) Products with aluminum alloy tensile plate as raw material and large cutting allowance, which can be processed by high-speed cutting;
- (4) Large batch of products suitable for multi spindle NC machining;
- (5) The product has high precision requirements and strict interchangeability requirements;
- (6) The program can complete the processing of opposite parts through the “mirror” function and use plates;
- (7) Products requiring NC machining in subcontracting production and new machine development at home and abroad.

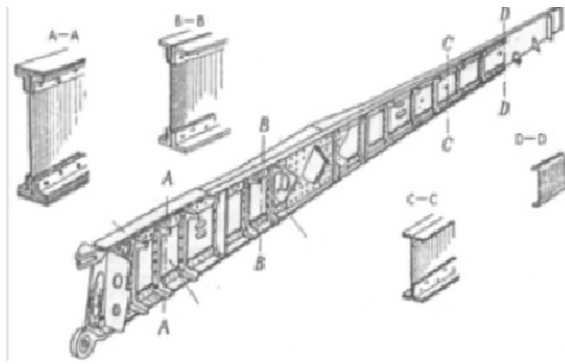
According to the above principles, aircraft integral machined structural parts are one of the main products suitable for NC machining. According to the structure, they can be divided into beams, wall panels, frames and joints. The following describes the common problems of the above types of parts in NC machining [6].

## 4 Application of NC Machining Technology in Aircraft Manufacturing

### 4.1 Beam

The most important standard and requirement of the aircraft during operation is to master the balance. The main function of the beam is to master the fuselage balance so that it can fly more smoothly. In the manufacturing process of beam, load-bearing member

is the key part to ensure the function of beam, so the requirement for its toughness is relatively high. At present, high-strength alloy structure or high-strength aluminum alloy is generally used in the manufacturing of load-bearing member. These materials have high strength and good toughness, which effectively ensure the service performance of the beam, so as to improve the safety factor of the aircraft. The load-bearing components are generally of variable angle profile, the cross-section shape is I-shaped, and equipped with notches, joint holes, etc. (as shown in Fig. 1 below). Therefore, the manufacturing process is relatively complex. In NC machining, the manufacturing process of this kind of parts is generally based on milling. The processing of various holes in individual components needs to be produced in strict accordance with the requirements of their size and location. In addition, the geometric dimension of the beam is relatively large, and there may be deformation problems. Therefore, NC machining technology is required to consider every detail to ensure its manufacturing quality [7].



**Fig. 1.** Schematic diagram of beam

## 4.2 Partition

The diaphragm is mainly used to bear gravity, and can also divide various functional areas of the aircraft. In the processing and manufacturing of the diaphragm, it requires a large area of the processing platform. It can be manufactured only when the working surface is large, which also means that it is necessary to launch the data control facilities with large table surface for operation. The main material of diaphragm is plate, so it is often used for cutting, but the excessive area also causes some difficulties for cutting. At this time, the rigidity and strength of the NC machining fixture can be used to improve the cutting force and cutting speed. The milling method can also avoid the deformation and distortion of the diaphragm in the manufacturing process, and further ensure the quality of the diaphragm [8].

### 4.3 Frame

Frame is an important load-bearing part in aircraft. It is an important part of fuselage shape. Its curved surface characteristics make it difficult to manufacture. In the manufacturing of frame, it is generally die forging, and then it is fine processed. The separate fixture with plane positioning and process hole positioning is adopted to better maintain the flatness and parallelism of the upper and lower parts of the frame. At the same time, the method of repeated processing up and down can be adopted, and the feed amount shall be controlled within 3mm. When approaching the final target size, the feed amount and speed shall be gradually reduced to ensure the integrity and accuracy of cutting. NC machining technology shows a strong automation advantage in the production and manufacturing of frame parts. It can set the feed rate and speed in advance, reduce the error rate of human factors, ensure the effective and rational use of aircraft materials, and reduce the error rate in construction, so as to improve the utilization rate of aircraft manufacturing materials, While reducing the manufacturing cost, it can also ensure that its manufacturing quality meets the requirements of aircraft manufacturing [9].

### 4.4 Joint

The main function of the joint is to connect other parts and bear the role of gravity. The structure of the joint usually has increased differences. In the manufacturing process, the shape of the joint needs to be designed according to different types of parts in the aircraft. Therefore, the general joint has the characteristics of complex geometry and more space angles (as shown in Fig. 2 below), As a result, the process in the manufacturing process is more complex. At present, the materials used in joint parts are generally materials with good strength and rigidity, which are cut by mechanical technology. However, in actual operation, the qualification rate will be relatively low due to the characteristics of the joint. At this time, we can use NC machining technology and adopt cutting tools with different machining depths according to the design drawing paper to give full play to the advantages of NC technology, Use both long and short knives to effectively improve the qualification rate of joint parts [10].

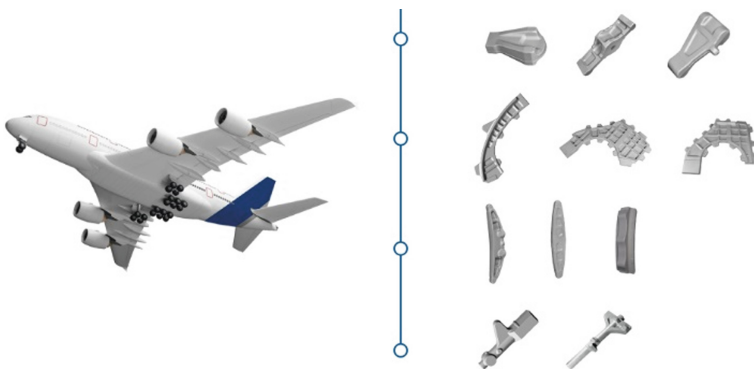


Fig. 2. Schematic diagram of connector parts

## 5 Epilogue

To sum up, the level of aircraft manufacturing is closely related to the development quality of China's aerospace industry. The development of numerical control technology can further improve the efficiency and accuracy of aircraft parts manufacturing, and there is no need for a large number of tooling in the processing and manufacturing process, which greatly shortens the R & D time of aircraft parts, thus continuously strengthening the market competitiveness of the aircraft manufacturing industry. With the progress and development of modern society, the requirements for college processing of parts in aircraft manufacturing are higher and higher. NC technology will also give full play to its application advantages and lay a good foundation for the healthy, stable and sustainable development of aircraft manufacturing industry.

## References

1. Dolganova, I., et al.: Assessment of critical resource use in aircraft manufacturing. *Circular Econ. Sustain.* **2**, 1193–1212 (2022). <https://doi.org/10.1007/s43615-022-00157-x>
2. Zutin, G.C., Barbosa, G.F., de Barros, P.C., Tiburtino, E.B., Kawano, F.L.F., Shiki, S.B.: Readiness levels of Industry 4.0 technologies applied to aircraft manufacturing—a review, challenges and trends. *Int. J. Adv. Manuf. Technol.* **120**(1–2), 927–943 (2022). <https://doi.org/10.1007/s00170-022-08769-1>
3. Lietzau, B., Mönch, L., Biele, A.: A multi-criteria production planning approach for aircraft manufacturing flow lines. *IFAC-PapersOnLine* **55**(2), 144–149 (2022). <https://doi.org/10.1016/j.ifacol.2022.04.184>
4. Eleonore, P., et al.: Global environmental mapping of the aeronautics manufacturing sector. *J. Clean. Prod.* **297**, 126603 (2021). <https://doi.org/10.1016/J.JCLEPRO.2021.126603>
5. Solow, D., Wu, Q., Magri, D.: An application of integer programming to producing aircraft engine parts. *Int. J. Appl. Manag. Sci.* **13**(1), 1 (2021)
6. Nayak, D., et al.: Laser scanning based methodology for on-line detection of inclusion in prepreg based composite aircraft manufacturing. *Mater. Today: Proc.* **24**(2), 591–600 (2020). <https://doi.org/10.1016/j.matpr.2020.04.313>
7. Beuß, F., et al.: Ergonomics simulation in aircraft manufacturing – methods and potentials. *Procedia CIRP* **81**, 742–746 (2019). <https://doi.org/10.1016/j.procir.2019.03.187>
8. Silva, A.V.S., Trabasso, L.G.: Design for automation within the aeronautical domain. *J. Braz. Soc. Mech. Sci. Eng.* **41**(7), 1–17 (2019). <https://doi.org/10.1007/s40430-019-1791-y>
9. Reis, A., et al.: Evolution stages of aircraft manufacturing firms. *Syst. Eng.* **22**(3), 255–270 (2019). <https://doi.org/10.1002/sys.21476>
10. Ivanov, V.K.: Ways to manage small-scale machine-building production. *Russ. Aeronaut.* **62**(1), 134–137 (2019). <https://doi.org/10.3103/S1068799819010185>



# Barriers of Mathematical Language Representation of College Students Based on Data Mining

Haiyan Wu<sup>(✉)</sup>

Zhixing College of Hubei University, Wuhan, Hubei, China  
haiyanshuxue@163.com

**Abstract.** With the innovation and development of computer application technology, colleges have further accelerated the process of information construction, and data mining has gradually become one of the most active research fields in education and teaching big data analysis. As a public basic course, college mathematics has many problems, such as a large number of students, high learning failure rate, and many learning disabilities. Stolyar, a famous mathematical educator of the former Soviet Union, said that mathematical teaching is also the teaching of mathematical language. Therefore, it plays a significant role in mathematics teaching to train students to skillfully use mathematical language to represent mathematical problems and enhance the students' ability of solving problems with mathematical methods. Based on the existing circumstances of students' mathematical language representation obstacles, this thesis deeply discusses the influencing factors of students' performance, constructs a decision tree model of the main factors affecting students' performance according to the decision tree arithmetic in data mining technology, and predicts the level of curriculum obstacles according to the model, so as to help teachers adjust teaching reasonably, and guide students to overcome the obstacles of mathematical language representation.

**Keywords:** Data Mining · Mathematical Language · Representation Barrier

## 1 Introduction

As a new data analysis technology, data mining uses artificial intelligence to analyze the data in the database. At present, it is used on a large scale in many fields. With the innovation and development of computer technology, all colleges and universities in China are implementing campus digital construction, with a large amount of educational data and huge databases. How to scientifically apply data mining to discover the laws in educational data is of great significance and application background for guiding schools to optimize teaching management and improve teachers' teaching work. Literature [1] studied the prediction model of college students' experimental class performance with data mining technology. Educational Data Mining can help learners continuously improve their academic performance through educational technology methods [2, 3]. Literature [4] comprehensively reviewed the educational data mining (EDM) and learning analysis

(LA) in higher education. More and more scholars at home and abroad have introduced data mining algorithms into the field of education, conducted in-depth analysis and mining of students' performance information, and achieved many results [5–9].

Stolyar, a famous mathematical educator of the former Soviet Union, said that mathematical teaching is also the teaching of mathematical language. Therefore, it plays a significant role in mathematics teaching to train students to skillfully use mathematical language to represent mathematical problems and enhance the students' ability of solving problems with mathematical methods. As we all know, college mathematics public courses mainly include advanced mathematics, linear algebra, probability theory and mathematical statistics. They have a wide range of courses, a large number of students and great influence. However, in the process of learning college mathematics, a considerable number of students have learning disabilities, especially mathematical language representation barriers. Mathematical language is mainly divided into graphic language, symbolic language and written language. The research on mathematical language representation barriers mainly focuses on children's enlightenment education and mathematics education in primary and secondary schools, while the research on learning barriers of college students in mathematics curriculum learning is relatively small [10]. Therefore, the further research on the obstacles of mathematical language representation of college students is very significant to improve the learning management of college students and the reform of college mathematics teaching.

In this writing, the decision tree arithmetic in data mining is used to discuss the factors affecting students' mathematical language representation obstacles, and put forward corresponding teaching suggestions to solve the language representation obstacles in mathematical problem, so as to help teachers formulate corresponding improvement measures, and guide students to overcome the mathematical language representation obstacles.

## 2 Model Preparation

The decision tree algorithm model can be established through data mining technology. On the one hand, it can grasp in advance the types and current situation of language representation barriers in mathematical problem solving of college students, as well as the degree of mathematical language representation barriers of students in different majors (science and engineering, economics and management, etc.). On the other hand, it reveals the important factors that affect students' mathematical language representation barriers, so as to help students eliminate learning barriers, and also provide an effective reference for college mathematics teaching. The research process generally includes three stages: data collection, model building, analysis and decision-making. First, the collection and preprocessing of educational data are completed to generate high-quality data. Then, analyze and mine the data, explore the key data variables, build the model under the decision tree rules and test the prediction effect. Finally, the conclusion of the model is refined and formed into a decision to solve the problem. Each stage interacts with each other and iterates circularly.

C 4.5 algorithm is mainly used to establish decision tree classification rules. It can not only process attributes with continuous values, but also process training samples



with vacant attributes. Based on the decision tree node selection criteria to maximize the information gain rate, complete the processing of different levels of data, and complete the overall construction of the decision tree from root to branch and from branch to leaf. The procedures of using C 4.5 algorithm to generate a decision tree are as below.

- 1) The information gain rate of each attribute of the training sample is obtained.
- 2) The attribute with the largest information gain rate is selected as the root node of the decision tree.
- 3) Divide the data set according to the root node attribute values, and recursively perform step 2 on the subset until the observation data in the subset is consistent with the values in the classification attributes.
- 4) According to the established decision tree, the classification rules are extracted and the decision tree is constructed.

The procedure of generating the decision tree model is shown in Fig. 1.

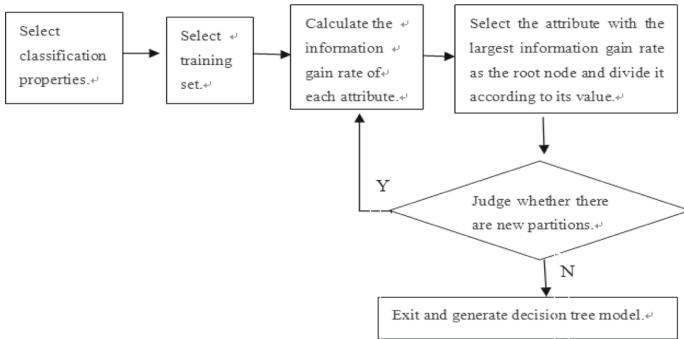


Fig. 1. The procedure of generating the decision tree model

### 3 Model Application

#### 3.1 Data Collection

In order to study the types of language representation barriers in students' mathematical problem solving, the test paper takes the knowledge points of higher mathematics courses as the test content, which involves the understanding, transformation and expression of mathematical language, and can accurately reflect the current circumstances of language representation barriers in students' mathematical problem solving. In this thesis, we collected the personal information of the students in Zhixing College of Hubei University in the form of online questionnaires, and finally collected 1016 questionnaires.

**Table 1.** Language representation barrier level assignment table

Levels of language representation barriers	NO	Mild	Moderate	Severe
Assignment	3	2	1	0

The results of the test paper are assigned according to the levels of language representation barriers in students' mathematical problem solving, as shown in Table 1. The mathematical language representation disorder subscale adopts a 4-level scoring system, including none, mild, moderate and severe. The more serious the obstacle is, the lower the score is, and vice versa.

In order to facilitate later data mining, students are classified into three categories according to their professional nature: science and engineering, economics and management, and humanities. According to the category of college entrance examination, they are divided into two categories: ordinary college entrance examination and skill college entrance examination. The data of 1016 students will be randomly divided into training set ( $n = 711$ ) and test set ( $n = 305$ ) in line with the ratio of 7:3. Based on the training set, the C4.5 algorithm is used to generate a prediction model, which is used to predict the test set data, and compare the prediction results with the actual results to test the model effect. The variables in the data are displayed in Table 2.

**Table 2.** Description of variables in data

Variable	Name	Value
Student gender	STUSEX	M(male), F(female)
College entrance examination category	GKLB	PT(Ordinary college entrance examination) JN(Skill college entrance examination)
Major classification	ZYFL	LG(science and engineering), JG(economics and management), RW(humanities)
Weekly time for self-study of mathematics after class	ZXSJ	1 (less than 2 h), 2 (2 to 5 h), 3 (more than 5 h)

### 3.2 Model Building

Decision tree algorithm is used to mine and analyze the college entrance examination category, student gender, major classification and other items in student data, find out the key factors that cause students' mathematical language recognition obstacles, help teachers better predict students' grades, timely find students' problems in learning, and make suggestions for students to learn better. The sample data is calculated with the algorithm C4.5. The decision tree after pruning is shown in Fig. 2.

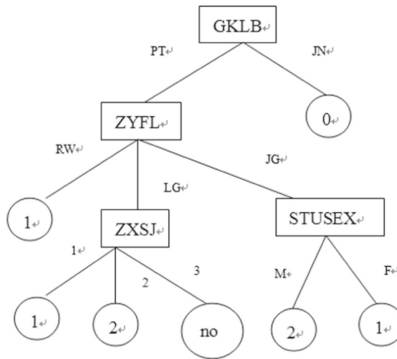


Fig. 2. The decision tree after pruning

### 3.3 Analysis of Resolution

This study finds that the key factor that leads to the obstacle of mathematical language representation is the category of college entrance examination. Generally speaking, the mathematics foundation of the skill college entrance examination students is not solid enough, and the mathematics subject needs long-term accumulation. The candidates with a weak foundation also encounter more obstacles in mathematics learning. This is consistent with the judgment drawn from daily experience. For ordinary college entrance examination candidates, the professional category factor plays a key role in the test scores. The scores of science and engineering students are generally better than those of economic students. The decision tree model combines the student data content to complete the evaluation, and the prediction accuracy can reach 81.5%. The above experimental results indicate that the prediction of the model has certain referential value. Its specific practical application value lies in: at the beginning of the class, teachers predict the students with serious obstacles in mathematics learning through the model, and issue academic early warning in advance, so as to take targeted preventive measures to help students improve their mathematics learning level.

Based on the test data analysis, the following conclusions are obtained.

- 1) Compared with written language and graphic language, symbolic language is the most difficult to understand, which is determined by the simplicity and abstraction of symbolic language itself. Therefore, symbolic language should be paid more attention to in the understanding of mathematical language.
- 2) Obstacles are most likely to occur in the transformation from graphic language to written language and symbolic language. Although the graphic language is intuitive, it needs to find out the details. It is difficult for students to accurately convert the graphic language into written language and symbolic language.
- 3) It is easier to have obstacles when using mathematical language to express. Students are especially used to reasoning or calculation in symbolic language. They not only pay insufficient attention to written language, but also lack the training of mathematical expression in written language.

## 4 Conclusion

In the epoch of big data, we make the best of information technology to mine education data, which provides an important scientific basis for teaching reform. This paper fully conforms to the demands of the processing of the information age, uses the decision tree algorithm to mine student data information, finds out the key factors that affect the obstacles of students' mathematical language representation, and realizes the analysis and prediction of the obstacles of students' mathematical language representation, so as to help teachers adjust their teaching work reasonably, and guide students to overcome the obstacles of mathematical language representation.

**Acknowledgements.** This work was supported by the 2021 general project of Hubei Educational Science Planning (No. 2021GB086 Research on Language Representation Barriers and Countermeasures of Students in Applied Undergraduate Colleges in College Mathematics Learning).

## References

1. Pan, D., Wang, S.H., Jin, C.: Research on student achievement prediction based on BP neural network method. *Adv. Artif. Syst. Med. Educ.* **75**(5), 15–23 (2021)
2. Li, M.: Study on the influence of non-intelligence factors on college students' English learning achievement based on C4.5 algorithm of decision tree. *Wireless Pers. Commun.* **2**(5), 1213–1222 (2018). <https://doi.org/10.1007/s11277-017-5177-0>
3. Adekitan, A.I., Salau, O.: The impact of engineering students' performance in the first three years on their graduation result using educational data mining. *Heliyon* **5**(2), e01250 (2019). <https://doi.org/10.1016/j.heliyon.2019.e01250>
4. Aldowah, H., Al-Samarraie, H., Fauzy, W.M.: Educational data mining and learning analytics for 21st century higher education: a review and synthesis. *Telematics Inform.* **37**, 13–49 (2019)
5. Pelaez, K., Guarcello, M., Fan, J., et al.: Using a latent class forest to identify at-risk students in higher education. *J. Educ. Data Min.* **11**, 18–46 (2019)
6. Hussain, S., et al.: Regression analysis of student academic performance using deep learning. *Educ. Inf. Technol.* **26**, 1–16 (2020)
7. Adekitan, A.I., Salau, O.: Toward an improved learning process: the relevance of ethnicity to data mining prediction of students' performance. *SN Appl. Sci.* **2**(9), 266–281 (2019)
8. Guarin, C.E.L., Guzman, E.L., Gonzales, F.A.: A model to predict low academic performance at a specific enrollment using data mining. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje* **10**(3), 119–126 (2015)
9. Yang, F., Li, F.W.B.: Study on student performance estimation, student progress analysis, and student potential prediction based on data mining. *Comput. Educ.* **123**, 97–108 (2018)
10. Toll, S.W.M., Van Der Ven, S.H.G., Kroesbergen, E.H., Van Luit, J.E.H.: Executive functions as predictors of math learning disabilities. *J. Learn. Disabil.* **44**(6), 521–532 (2011)



# Mathematical Model (MM) and Signal Processing (SP) in Computer Simulation (CS) on Account of Improved Genetic Algorithm (IGA)

Bo Wang<sup>(✉)</sup>

School of Science, Dalian Jiaotong University, Dalian 116028, Liaoning, China  
rvewang1998@163.com

**Abstract.** MM and SP a computational feature in which a computer can simulate the state of a system at high speed and accurately while eliminating machine faults. Over the years, computer MM and SP have slowly been accepted and accepted by most people. Computers should consider all aspects of how to lay out MM and SP modules. Whether this system layout can meet the application requirements can be tested to prove. In this text, a CS patterning and SP module are set up, and computer MM and SP applications are processed. This text studies the MM and SP in CS on account of IGA, introduces the professional knowledge of MM and signal in CS, and expounds its operation principle. The data test shows that the MM and SP in CS on account of IGA can improve the accuracy of MM and signal efficiently.

**Keywords:** IGA · Computer Simulation · Mathematical patterning · Signal Processing

## 1 Introduction

The principle of wavelet analysis mode is to extract the characteristics and data of partial discharge signal and electric wave signal through partial discharge. On the basis of repeated investigation and familiar with pd SP and experiment at home and abroad, ciah used wavelet analysis algorithm to analyze the characteristics of two kinds of signals, part data, frequency data, and Fisher discriminant mode to select the most characteristic quantity for identification processing. This operation can greatly improve the accuracy of simulation data patterning and SP accuracy. MM and SP in CS on account of IGA is beneficial to the progress of MM and signal simulation technology.

Domestic and foreign scholars have studied the IGA. In foreign studies, KumariR proposed that in the process of image segmentation, the subset of the image should be found according to the gray value of the pixel or the position of the pixel. Image segmentation using genetic algorithm is an NP-complete problem, and finding the solution of this kind of problem is a computationally difficult task [1]. In the new genetic algorithm proposed by AcunaY, the global grabbling power is realized by an improved island

pattern, while the part grabbing power is realized by a new accommodative disparate outburst trimming program. Part radical grabbing mode is adopted to handle raw concentration [2]. Al-obaidima proposed an optimization framework on account of species conservation Genetic Algorithm (SCGA) to optimize process design and operational parameters. In order to enable readers to have a deeper understanding of the process, the effects of membrane design parameters on the dimethylphenol interception rate, water recovery rate and specific energy consumption level of the two different processes were studied [3].

Neurons complete the construction of the algorithm network by means of synaptic connectivity, which can collect data information and collect electrical characteristics [4, 5]. It is very important for medical application to study the data and transmission mechanism of algorithmic network. The neural network has nonlinear complex feedback, and it is difficult to conduct experiments from the internal structure [1, 6]. On the premise of in-depth study of neural network structure, the application field can be simulated, so that the internal structure and attributes can be obtained. The MM and SP in CS on account of IGA is beneficial to improve the simulation effect and the quality of CS patterning and SP.

## 2 Regrabbling on the Design of MM and SP in CS on Account of IGA

### 2.1 IGA

Genetic algorithm refers to the global grabbing and optimization of targets in a parallel way [7, 8]. Simulated biological outburst was used to solve the optimal solution.

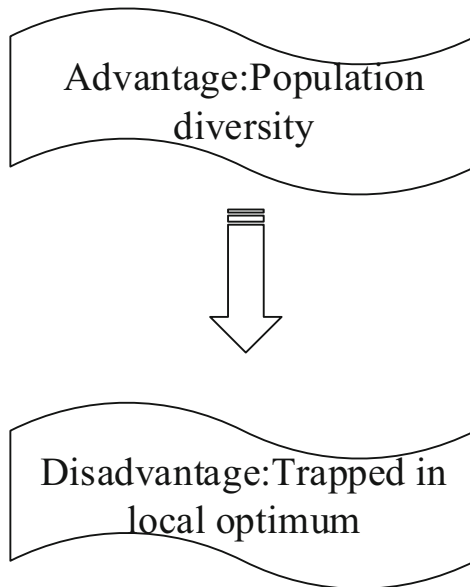


Fig. 1. Advantages and disadvantages of genetic algorithms

Compared with the part grabbing algorithm, this algorithm has more advantages. The algorithm can grab a wider solution space, which is an effective technique for solving.

One of the advantages of the algorithm is to process the data in the way of population. At this time, the diversity of population can be inherited and the advantages of population can be borrowed into the algorithm. A population is essentially a viable solution space, so a single individual in a viable solution space is a viable solution. The efficiency of the algorithm can be improved by using population theory to grab feasible solutions. In addition, the defects of genetic algorithm also exist. If individuals in the feasible solution space are competing with each other, survival of the fittest will be carried out, so that similar individuals will remain. In this mechanism, there are all similar individuals, the original diversity of the population will be broken, so that the algorithm will fall into part optimum more early. In this case, the global effect of the crossover is reduced, so that the next generation is not more easily produced. Population outburst adopts part variation for processing, and the solution is limited to the domain scope for solving, so the global optimal solution cannot be achieved, as shown in Fig. 1.

Crossover operation is a mainstream computing mode in genetic algorithm, which influences the efficiency of the pattern. So, the key to algorithm success is to have a perfect crossover operation. The algorithm steps of the crossover pattern are as below:

Step1: calculate the Euclidean distance between any two elements in the population and generate the adjacency matrix D;

Step2: Use the algorithm pattern to process the adjacency matrix D and generate the minimum spanning tree T;

Step3: solve the correlation value of T, its average weight QUOTE and select the maximum weight value ( $0 < \delta < 1$ ) in T where V is not greater than  $\delta \times$ ;

Step4: read tree T one by one, grabbing all boundaries greater than V and then partition to obtain sub-connected graphs;

Step5: grabbing the sub-connected graph one by one, obtain subclasses, and then perform sequence arrangement;

Step6: Roulette picks out any single element X1, obtains the sequence I of the subpopulation below the single element, and obtains an optimal individual X2 from the subpopulation I;

Step7: Obtain the subpopulation J nearest to I, and arbitrarily select a single individual Y of subpopulation J;

Step8: x1 and X2 perform crossover operation, and the set obtained is X;

Step9: Select individuals farther from individual Y, assume that individual X2 and individual Y operate cross, and obtain set Y;

Step10: Use the more excellent individuals in the two sets of algorithms to change the individuals into offspring individuals.

## 2.2 MM and SP in CS on Account of IGA

### 1) CS MM

After neurons are connected to all equivalent circuits, professional theoretical knowledge and neuron knowledge are used to construct mathematical algorithm

formulas [9, 10]. The electrical properties are shown by the membrane potential of neurons.

Numerical representation adopts individual elements in the algorithm formula under neurons, and such representation or description of data is called numerical representation of objects [11, 12]. The data value representation can have the following specific advantages:

- (1) Data sharing within the network [13–14]. Each neuron in a neural algorithm can be composed of basic components, and one of its characteristics is that it can be shared.
- (2) Network mutual data reuse. Algorithm data assembly, using the common constituent in the algorithm.
- (3) Data files can be generated and edited [15–16]. Basic object files are usually small and in a fixed form. In this case, you can generate files for each basic object configuration interface.
- (4) Object-oriented programming. Data description from the basic object from the hair, its link to the corresponding class, can be object-oriented programming. Its advantage is that it can ensure data security and is easier to expand and process.
- (5) Optional parameters [17]. Parameters are divided into files, which have corresponding algorithm formulas and specific definitions, and can have value ranges.

2) SP

The selection of appropriate wavelet is very important for PD signal analysis. When selecting the optimal wavelet basis in multi-scale analysis, in addition to considering the orthogonality, support set and symmetry of the wavelet, it is necessary to ensure that the optimal wavelet can best reflect the frequency characteristics of PD signal at different layers. For a given signal, the optimal wavelet has the largest wavelet coefficients in the time-scale domain.

### 3 Regrabbling on MM and SP Effect in CS on Account of IGA

MM and SP in CS on account of IGA, the membrane potential  $V_i$  of each neuron is expressed by dynamic equation:

$$\frac{dV_i}{dt} = \frac{sI_{ex} - \sum_{j=1}^m I_{vdij} - \sum_{k=1}^n I_{eS_{ik}} - \sum_{k=1}^n \sum_{l=1}^p I_{cS_{ikl}}}{C_{M_i}} \tag{1}$$

where,  $I$  corresponds to the subscript of neuron membrane potential  $V$  in the network;  $J$  corresponds to the subscript of voltage dependent conductance current QUOTE ;  $K$  corresponds to the subscript of electrical synaptic current QUOTE ;  $L$  and subscripts corresponding to chemical synaptic currents QUOTE ;  $N$  is the number of neurons and electrical synaptic currents in the network;  $M$  is the number of ion current in each neuron;  $P$  is the number of chemical synaptic currents corresponding to each electrical synapse.



SP. The simulation system has 4 external terminals: display Monitor, data, printing interface, Postscript format file. Postscript doesn't have a high settlement. The external algorithm will process according to the settlement and automatically filter the complex and useless information under certain precision conditions. Simulation is internal to the system and requires very little complexity. Its main principles are described as below:

- (1) According to different yield media (Monitor, printer, file, and Postscript file), obtain its range (width, height);
- (2) Set the yield optimization factor (Rx, RY) according to different media. The larger the coefficient, the taller the corresponding yield settlement. The values of optimization factors for each yield media are as below.
- (3) According to the media yield range and optimization factor, establish the corresponding optimization rules

$$\text{Int.}[\text{curX} * \text{rx}] < \text{width} \quad (2)$$

$$\text{Int.}[\text{curY} * \text{ry}] < \text{height} \quad (3)$$

$$\text{Int.}[(\text{preX} - \text{curX}) * \text{rx}] \neq 0 \quad (4)$$

$$\text{Int.}[(\text{preY} - \text{curY}) * \text{ry}] \neq 0 \quad (5)$$

Int.[] indicates the integer. (preX,preY) denotes the original value point; (curX,curY) denotes the current numerical point;

- (4) Apply rules and accommodatively optimize the yield data.

- (a) The original numerical point (preX,preY) is cleared;
- (b) Calculate the current numerical point (curX,curY)
- (c) If rules I and II are met, go to (d); Otherwise, go to (e)
- (d) yield the current point if rule three or four is met; Otherwise, go to (f)
- (e) Assign the current value point to the previous value point;
- (f) If not the last numerical point, go to (b); Otherwise, end.

#### 4 Investigation and Analysis of MM and SP in CS on Account of IGA

The data patterning and SP effects of the pattern in this text were detected by binary coding and real coding. The algorithm patterns were traditional algorithm Harr+ Adaboost, DPM and IGA in this text. 23 functions were selected to compare the algorithm effects. Parameter reference of algorithm pattern. Configuration environment: The environment mirror system is Windows8, the processor is Inter(R)Core(TM) I3-8556, CPU@4.20GHz, the memory is 12GB, and VS2008 software is used to run the algorithm.

During the test steps, the computer is used to extract data from all test points, such as system operation data, channel counting data, bus storage data, etc. If the detection

**Table 1.** Test parameter list

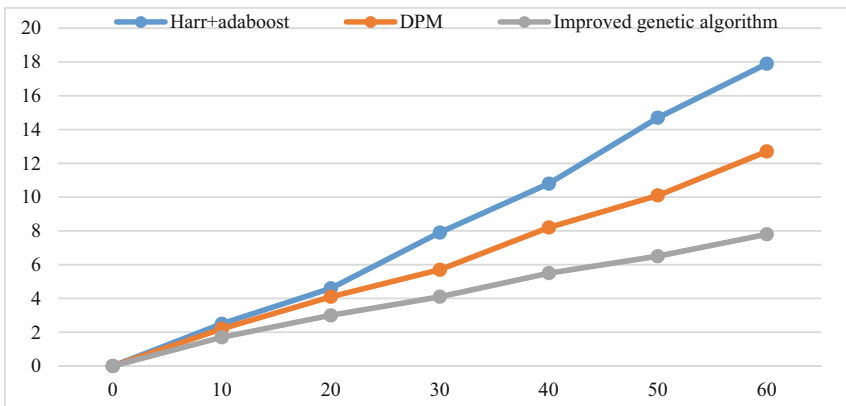
Algorithm	Gen	Time(s)
Harr+ adaboost	18	60
DPM	13	60
IGA	8	60

bit is faulty, you can query fault parameters, such as the number of channels, fault classification, fault category, and discrete data. The time required by Harr+ Adaboost, DPM and the IGA in this text in testing, and the algebraic Gen of population outburst are shown in Table 1.

The three items in the first row of Table 1 are Algorithm, POPULATION outburst algebra (Gen) and total running Time respectively. The first column contains three algorithms, namely Harr+ Adaboost, DPM and the IGA in this text. In Table 1, the maximum outburstary algebras of the three algorithms in the test are 18,13,8 respectively. The total running time is 60 s.

In Fig. 2, the gray line represents the IGA in this text, the orange line represents DPM and the blue line represents Harr+ Adaboost. As can be revealed from the figure, the IGA in this text has the lowest line bit and the least population outburst algebra. It can be revealed that the IGA has the highest effect.

The data show that the MM and SP in CS on account of IGA has remarkable effect in the field of MM and SP in CS.



**Fig. 2.** Algorithm operation effect diagram

## 5 Conclusions

Many test data and practical performance can prove that the algorithm pattern is very accurate, the data description is correct, the signal calculation is reasonable, and the

variable dynamic processing and parallel operation are carried out. This simulation system uses IGA pattern, computer tools, programming language is Java. The simulation system can run on all Java language platforms. The system uses the image interface, which has stored many case scenes. The MM and SP in CS on account of IGA is beneficial to the progress of CS MM and SP technology.

## References

1. Kumari, R., Gupta, N., Kumar, N.: Image segmentation using IGA. *Int. J. Eng. Adv. Technol.* **9**(1), 1784–1792 (2020)
2. Acuna, Y., Sun, Y.: An efficiency-IGA and its application on multimodal functions and a 2D common reflection surface stacking problem. *Geophys. Prospect.* **68**(4), 1189–1210 (2020)
3. Al-Obaidi, M.A., Ruiz-García, A., Hassan, G., et al.: Pattern based simulation and genetic algorithm based optimisation of spiral wound membrane RO process for improved dimethylphenol rejection from wastewater. *Membranes* **11**(8), 595–595 (2021)
4. Niazy, N., El-Sawy, A., Gadallah, M.: Solving capacitated vehicle routing problem using chicken swarm optimization with genetic algorithm. *Int. J. Intell. Eng. Syst.* **13**(5), 502–513 (2020)
5. Kukker, A., Sharma, R.: Genetic algorithm-optimized fuzzy lyapunov reinforcement learning for nonlinear systems. *Arab. J. Sci. Eng.* **45**(3), 1629–1638 (2020). <https://doi.org/10.1007/s13369-019-04126-9>
6. Mirzaie, K.: An improved memetic genetic algorithm on account of a complex network as a solution to the traveling salesman problem. *Turk. J. Electr. Eng. Comput. Sci.* **28**(5), 2910–2925 (2020)
7. Ongcunarak, W., Ongkunaruk, P., Janssens, G.K.: Genetic algorithm for a delivery problem with mixed time windows. *Comput. Ind. Eng.* **159**(1), 107478 (2021)
8. Mekki, B.S., Langer, J., Lynch, S.: Genetic algorithm based topology optimization of heat exchanger fins used in aerospace applications. *Int. J. Heat Mass Transf.* **170**(2), 121002 (2021)
9. Shirmohammady, N., Izadkhah, H., Isazadeh, A.: PPI-GA: a novel clustering algorithm to identify protein complexes within protein-protein interaction networks using genetic algorithm. *Complexity* **2021**(7), 1–14 (2021)
10. Viana, M.S., Junior, O.M., Contreras, R.C.: An Improved part grabbling genetic algorithm with a new mapped accommodative operator applied to pseudo-coloring problem. *Symmetry* **12**(10), 1684 (2020)
11. Mahrooghi, A., Lakzian, E.: Optimization of Wells turbine performance using a hybrid artificial neural fuzzy inference system (ANFIS) - Genetic algorithm (GA). *Ocean Eng.* **226**(1), 108861 (2021)
12. Jha, S., Mehta, A.K.: A hybrid approach using the fuzzy logic system and the modified genetic algorithm for prediction of skin cancer. *Neural Process. Lett.* **54**(2), 751–784 (2021). <https://doi.org/10.1007/s11063-021-10656-x>
13. Iqbal, F., Rehman, S.U., Malik, K.I.: Optimal load scheduling using genetic algorithm to improve the load profile. *Afr. Dev. Resour. Regrabbling Inst. (ADRR) J.* **4**(9(3)), 1–15 (2021)
14. Kolouri, S., Yin, X., Rohde, G.K.: Neural networks, hypersurfaces, and the generalized radon transform [lecture notes]. *IEEE SP Mag.* **37**(4), 123–133 (2020)
15. Samaali, H., Najjar, F., Chaalane, A.: Patterning and design of an ultra low-power NEMS relays: application to logic gate inverters. *Analog Integr. Circuits SP* **104**(1), 17–26 (2020). <https://doi.org/10.1007/s10470-020-01658-1>
16. Radzhapov, S., Rakhimov, R., Radzhapov, B., et al.: Calculation of stages of the technological process of manufacture of PPD detectors using computer MM and production of alpha radiometer on their basis. *Comput. Nanotechnol.* **7**(2), 21–28 (2020)

17. Wyrwal, J., Pawelczyk, M., Liu, L., et al.: Double-panel active noise reducing casing with noise source enclosed inside – patternling and simulation study. *Mech. Syst. SP* **152**(2), 107371 (2021)



# English Translation Intelligent Recognition Model Based on Reverse Maximum Matching Segmentation Algorithm

Ruichao Li<sup>(✉)</sup>

School of Translation Studies, Xi'an Fanyi University, Xi'an, Shaanxi, China  
280663560@qq.com

**Abstract.** Since there is no obvious delimiter in Chinese, the problem of word segmentation that does not exist in English is a very important step in Chinese. This paper mainly studies the design of English translation intelligent recognition model based on reverse maximum matching segmentation algorithm. In this paper, we first propose the inverse maximum matching segmentation algorithm, and propose the NMT model based on dual learning and mask language model (MLM). The dual learning based NMT model uses one encoder and two decoders (including a source language decoder and a target language decoder). The experimental results show that the method proposed in this paper has a good effect in identifying new words in the article.

**Keywords:** Reverse Maximum Matching · Word Segmentation Algorithm · English Translation · Intelligent Model

## 1 Introduction

Due to the diversity and complexity of languages and the lack of computer performance, there has not been much progress in machine translation for a long time after the concept was proposed. However, in the last 30 to 40 years, with the vigorous development of deep learning of artificial intelligence, the development of machine translation has a new direction. As we all know, the translation between languages not only involves a variety of rich vocabulary, but also needs to consider the complex correspondence between grammar and sentence patterns. Traditional machine translation methods generally design templates for the translation between two languages to match, which is undoubtedly inefficient and inaccurate. In the research of machine translation, natural language processing is an extremely important part [1, 2]. In the field of Natural Language Processing (NLP), Chinese is more difficult to process than English, Portuguese and other western languages. One important reason is that in the Chinese corpus, words as the basic unit of a sentence do not have formal delimiters (while Spanish words have space marks as delimiters). Therefore, word segmentation, which does not exist in Spanish, is a very basic and important step in Chinese processing [3]. The correct segmentation of Chinese sentences will be of great benefit to the downstream tasks of

natural language processing, such as emotion classification and topic analysis. However, the existing word segmentation tools are more or less based on the dictionary. Research shows that out of vocabulary (words that do not exist in the dictionary) is an important factor affecting the accuracy of word segmentation. More than 60% of Chinese word segmentation errors are caused by the existence of new words. Relying on manpower to build a thesaurus requires considerable effort. With the rapid growth of the number of new words, the cost of maintaining and updating the thesaurus also increases [4]. In this context, the technology of automatically mining new words from text by computer came into being.

Chinese word segmentation was proposed in the 1980s. After nearly 40 years of development, it is now facing two major technical problems: segmentation of ambiguous words and unlisted words. Ambiguous words are caused by two or more segmentation results in a Chinese sentence. Common segmentation ambiguities are intersection type segmentation ambiguities, combination type segmentation ambiguities and the ambiguities of the Chinese language itself; Among them, unknown words are also called new words, which refer to words that are not recorded in the dictionary, such as network words, person names, place names, organization names, and professional terms in all walks of life. Therefore, whether ambiguous words and unknown words can be effectively resolved is an important factor to judge the performance of Chinese word segmentation methods. Chinese word segmentation methods can be roughly divided into three categories: first, dictionary based word segmentation methods, which match the Chinese sentences to be segmented one by one with the dictionary prepared in advance. If words with the same segmentation fields in Chinese sentences are found in the dictionary, the matching is successful; otherwise, the new Chinese character string obtained by subtracting characters is matched again [5]. The second is a word segmentation method based on statistics. It uses the frequency of adjacent Chinese character strings appearing in the corpus as a basis to judge whether adjacent Chinese character strings can be combined into a new word. Although this method solves the problem caused by the imperfect dictionary, it will also segment some wrong high-frequency Chinese character strings as an independent whole, such as “my”, “ate” “Why” and other meaningless words [6]. However, it is difficult to convert syntactic and semantic information into a recognizable way for computers, so this method is still in the exploration and development stage [7].

With the continuous development of the times, the world is accelerating its integration. More and more fields require cooperation and communication between different countries to work effectively, and language communication is the most basic requirement for cooperation, which means that machine translation has a decisive strategic significance in both academia and industry.

## **2 Translation Model Based on Reverse Maximum Matching Algorithm**

### **2.1 Reverse Maximum Matching Segmentation Algorithm**

The word segmentation algorithm based on string matching is also called dictionary based word segmentation algorithm or mechanical word segmentation algorithm [8].

This algorithm first requires a computer word segmentation dictionary with sufficient coverage, and then matches the Chinese string to be segmented with the entries in the word segmentation dictionary according to certain rules. If the word segmentation dictionary contains the characters or character strings in the Chinese string, the matching is successful, and the character or string is regarded as a recognized word [9, 10].

The advantages and disadvantages of the word segmentation algorithm based on string matching are as follows:

Advantages: simple algorithm, easy implementation and high efficiency;

Disadvantages: low efficiency, slow speed, difficulty in handling ambiguity, inability to learn independently, and unsuitable for handling Chinese texts of Jiaotong University.

This segmentation algorithm can have different classifications. According to the starting position and direction of scanning, it can be divided into forward matching algorithm and reverse matching algorithm.

Because there are many positive phrases in Chinese, if we match from back to front, we can do reverse matching to reduce the error rate.

The basic idea of Reverse Maximum Matching Method (RMM) is as follows:

Assume that the number of characters in the entry with the largest number of characters in the word segmentation dictionary is MaxLen, and set the Chinese string to be segmented as S1;

Judge whether S1 is empty. If it is empty, go to step 7. If not, go to the next step;

Take no more than MaxLen characters from the string S1 from left to right as the matching string W;

Find the word segmentation dictionary. If the character string W exists in the word segmentation dictionary, the matching succeeds. Go to the second step. If the character string W does not exist, the matching fails. Go to the next step;

Remove the leftmost word of W as the new W;

Determine whether W is a single word. If it is a single word, go to step 7. If it is not a single word, go to step 4:

Output results.

## 2.2 English Translation Intelligence Model

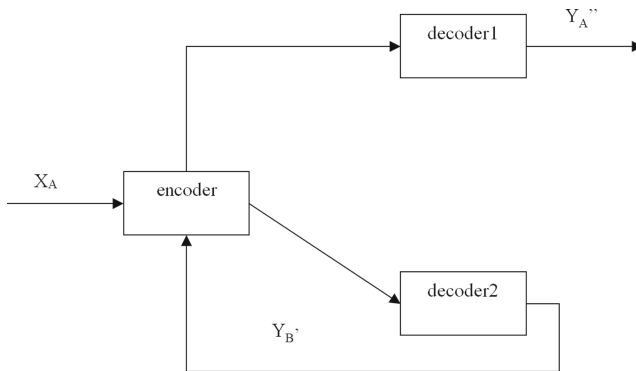
NMT semi supervised machine translation model can use monolingual corpus to improve the performance of the pre training model. Even if only monolingual corpora are available, dual NMT models can obtain high-quality feedback through forward translation and back translation.

For the sake of brevity of language expression, this paper uses (E, F, G) to represent a message of language A (which can be a sentence or a piece of text). Agent A only understands language A, and sends a message (E, F, G) to Agent B through a noisy channel 1. Channel 1 uses a translation model to convert the information into messages (x, y, z) of language B. Proxy B only understands language B. After receiving the message (x, y, z), it checks the information and tells proxy A whether the message is a natural statement in language B. Then agent B sends the received message (x, y, z) to agent A through another noisy channel 2, and another translation model in channel 2 can convert this information into language A messages (H, I, J). After Agent A receives the message (H, I, J), it checks the message and determines whether it is consistent with the

original message (E, F, G) sent by Agent A, and then feeds it back to Agent B. Through the feedback mechanism, the two agents can know whether the two translation models perform well, and then make improvements to the translation models. The above process can also start from agent B by sending the original information of language B.

Given a corpus DA of language A and a corpus DB of language B. The two corpora are not strictly aligned. There are two translation models that can translate language A into language B, or translate language B into language A. Starting from a statement in corpus DA, the model translates it into language B, and then back to language A. By evaluating the original sentences and the back translation results, we can know the translation quality of the two translation models, and then improve the model parameters respectively.

The translation model encoder used in this paper is composed of Transformer, which acts as a shared encoder. The two decoders also use Transformer, which contains a source language decoder and a target language decoder. The two decoders have the same structure, but different model parameters. As shown in Fig. 1, the source language is language A and the target language is language B. Decoder 1 is the source language decoder, and decoder 2 is the target language decoder. First, we can train a dual NMT model to translate sentence  $X_A$  (sentences from monolingual corpus of language A) into language B.  $Y_A$  and  $Y_B$  represent the output statements of language A and language B. We input the statement  $X_A$  into the encoder. After a dual translation process, decoder 2 can output a statement  $Y_B'$ , and then  $Y_B'$  is used as the input of the encoder. Finally, decoder 1 outputs a sentence  $Y_A''$ . The dual NMT model can be improved by evaluating  $X_A$  and  $Y_A''$ .



**Fig. 1.** Example of dual NMT model

In document level translation tasks, the standard training goal is to maximize the log likelihood estimation of training data:

$$\hat{\theta} = \arg \max_{\theta} \left\{ \sum_{(X,Y) \in D_d} \log P(Y|X; \theta) \right\} \quad (1)$$

among  $\theta$  It is a set of model parameters. However, the number of large-scale document level parallel corpora is limited. Under this condition, the translation performance of the



translation model for document level corpora is often worse than that for sentence level corpora, which is due to the insufficient estimation of the low probability in the text. In the past, some researches used the “two-step” training strategy to train document level translation models using large-scale sentence level parallel corpora.

### 3 Translation Model Word Segmentation Experiment

#### 3.1 Experimental Data

In the field of Chinese word segmentation, there are four commonly used data sets: MSRA data set provided by Microsoft Research Asia and PKU data set provided by Peking University, XITYU data set provided by City University of Hong Kong and CKIP data set provided by Chinese Knowledge and Information Processing Laboratory of Taiwan Central Research Institute. The first two are simplified Chinese, while the last two are traditional Chinese. This paper uses the first two datasets.

This paper first uses the 600000 news texts obtained by crawling to train, and then uses the standard test corpus of Peking University (PKU) and Microsoft Research Institute (MSR) to evaluate the reverse maximum matching segmentation method introduced in this paper. The size of the test corpus of Peking University is 498KB, and the size of the test corpus of Microsoft Research Institute is 548KB. Finally, the performance of the segmentation algorithm proposed in this paper is compared with the current four segmentation tools.

#### 3.2 Evaluation Index

In this experiment, the performance of Chinese word segmentation algorithm is compared by using three evaluation indicators: accuracy rate, recall rate and F1 (F1 measure) value. The three evaluation indicators are defined as follows.

The accuracy rate  $P$  refers to the probability that all the predicted positive samples are actually positive samples, and its formula is:

$$precise = \frac{TP}{TP + FP} \quad (2)$$

Recall rate  $R$  refers to the probability that positive samples are predicted to be positive samples, and its formula is:

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

Accuracy and recall are contradictory to some extent. For example, blindly pursuing accuracy will lead to a decline in recall, while deliberately pursuing recall will also lead to a decline in accuracy. Therefore, the F1 value is a comprehensive evaluation index that can simultaneously consider the accuracy rate and recall rate, is the weighted harmonic average of  $P$  and  $R$ , and its formula is:

$$F1 = \frac{2 \times precise \times recall}{precise + recall} \quad (4)$$

## 4 Analysis of Word Segmentation Results

### 4.1 PKU Test Set

In this paper, PKU and MSRA are used to compare the effect of reverse maximum matching segmentation algorithm. Among them, jieba, NLPIR, LTP and THULAC are the three evaluation indicators. The data comes from THULAC official website.

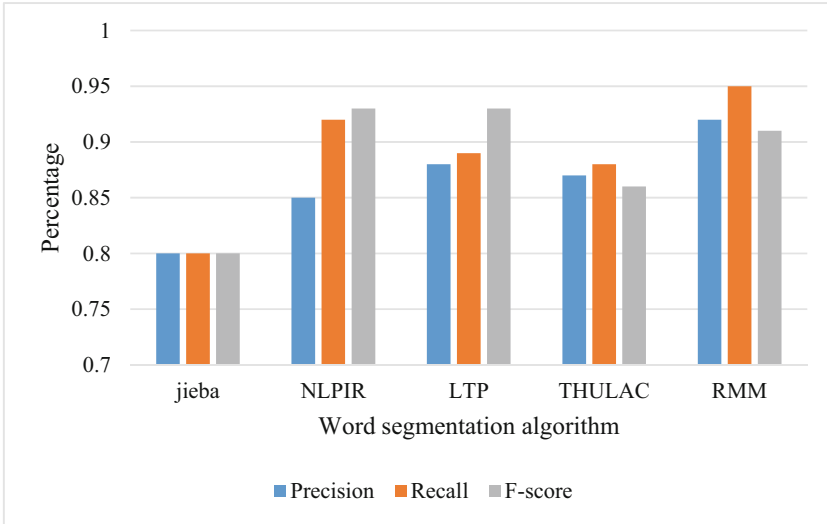
**Table 1.** Word segmentation results of pku test set

	Precision	Recall	F-score
jieba	0.83	0.75	0.80
NLPIR	0.91	0.91	0.91
LTP	0.97	0.94	0.94
THULAC	0.92	0.90	0.91
RMM	0.96	0.96	0.94

It can be seen from Table 1 that on the pku test set, the accuracy of the reverse maximum matching algorithm proposed in this paper is as high as 96%. Compared with the current four word segmentation tools, it is 13% higher than the jieba word segmentation tool, 5% higher than NLPIR and THULAC, and slightly lower than the LTP word segmentation tool; In terms of recall rate and F1 value, it is 21% and 14% higher than that of jieba word breaker, respectively, and is on a par with NLPIF, LTP and THULAC word breakers.

### 4.2 MSR Test Set

It can be seen from Fig. 2 that in the msr test set, the reverse maximum matching word segmentation algorithm proposed in this paper has better ambiguity word processing results. Compared with the four word segmentation devices, the accuracy rate and recall rate have improved by varying degrees, respectively 15% and 14% higher than the jieba word segmentation device; F1 value is slightly lower than NLP and LTP word breakers.



**Fig. 2.** Word segmentation results of msr test set

## 5 Conclusions

In the field of natural language processing, especially in Chinese natural language processing, word segmentation is an important benchmark task, and the performance of its results will directly affect the final performance of subsequent machine translation, information retrieval and other high-level pragmatic tasks. The powerful feature self-learning ability of deep learning frees researchers from complex feature design, and its excellent generalization ability makes it a hot spot in the field of natural language processing. The research goal of this paper is to complete the new word recognition and disambiguation, which is slightly better than the current mainstream word segmentation tools such as jieba, nlp, etc., but there are still some improvements, and the performance of word segmentation needs to be improved. The next research work can be carried out from the following two aspects: we can consider the fusion of new word recognition algorithm and ambiguity word disambiguation algorithm, that is, a word segmentation algorithm has the ability to process ambiguity words and recognize new words.

**Acknowledgements.** This work was supported by the teaching reform project of Xi'an Fanyi University (Number: J20B47).

## References

1. Doval, Y., Gómez-Rodríguez, C.: Comparing neural-and N-gram-based language models for word segmentation. *J. Am. Soc. Inf. Sci.* **70**(2), 187–197 (2019)
2. Bernard, M., et al.: WordSeg: standardizing unsupervised word form segmentation from text. *Behav. Res. Methods* **52**(1), 264–278 (2019). <https://doi.org/10.3758/s13428-019-01223-3>

3. Malakar, S., Ghosh, M., Bhowmik, S., et al.: A GA based hierarchical feature selection approach for handwritten word recognition. *Neural Comput. Appl.* **32**(7), 2533–2552 (2020)
4. Asgari, E., McHardy, A.C., Mofrad, M.R.K.: Probabilistic variable-length segmentation of protein sequences for discriminative motif discovery (DiMotif) and sequence embedding (ProtVecX). *Sci. Rep.* **9**(1), 1–16 (2019)
5. Naseem, U., Razzak, I., Khan, S.K., et al.: A comprehensive survey on word representation models: From classical to state-of-the-art word representation language models. *Trans. Asian Low-Resour. Lang. Inf. Process.* **20**(5), 1–35 (2021)
6. Perruchet, P.: What mechanisms underlie implicit statistical learning? Transitional probabilities versus chunks in language learning. *Top. Cogn. Sci.* **11**(3), 520–535 (2019)
7. Owens, C.A., Peterson, C.B., Tang, C., et al.: Lung tumor segmentation methods: impact on the uncertainty of radiomics features for non-small cell lung cancer. *PLoS ONE* **13**(10), e0205003 (2018)
8. Nararatwong, R., Kertkeidkachorn, N., Cooharajanone, N., et al.: Improving Thai word and sentence segmentation using linguistic knowledge. *IEICE Trans. Inf. Syst.* **101**(12), 3218–3225 (2018)
9. Jeong, H., Park, S., Jung, J.H., et al.: Segmented differential power processing converter unit and control algorithm for photovoltaic systems. *IEEE Trans. Power Electron.* **36**(7), 7797–7809 (2020)
10. Khattak, A., Asghar, M.Z., Saeed, A., et al.: A survey on sentiment analysis in Urdu: a resource-poor language. *Egypt. Inform. J.* **22**(1), 53–74 (2021)



# A New Retraceable Clustering Algorithm

Yefan Liu<sup>1</sup>(✉), Yijing Liang<sup>2</sup>, and Yina Yin<sup>1</sup>

<sup>1</sup> Information Center, Liaoning Earthquake Agency, Shenyang, Liaoning, China  
zhanxia03296528@163.com

<sup>2</sup> Liaoning Earthquake Station, Liaoning Earthquake Agency, Shenyang, Liaoning, China

**Abstract.** With the advent of the era of Internet of everything, building a smart city requires us to set off a technological revolution in various industries. Facing the current trend of geographic information data continuously developing to big data, this paper studies the suitable clustering algorithm for the characteristics of geographic information data, an efficient retraceable hierarchical clustering algorithm is proposed for two-dimensional spatial data, firstly, the region map is divided by a method similar to grid indexing, and the precision of Geohash is used as the grid size division criterion. The data density of each grid under different Geohash precision is calculated, and a suitable precision is selected as the basis of clustering, and the regions are classified as locally dense, locally uniform and locally dispersed according to the data density under this precision, and reasonable clustering is performed according to the region division. Subsequently, the clustering is divided into two parts: splitting phase and merging phase, adding markers to subclasses at the time of splitting so that they can be traced back to the clusters before splitting, and then testing whether the splitting results need to be merged at the end of splitting. Compared with coalescence or splitting methods, it can better delineate clusters of clusters. The marker detection method makes the splitting results traceable back, improves the fault tolerance, and the merging phase is less computationally intensive.

**Keywords:** Spatial index · Hierarchical clustering · Geohash

## 1 Introduction

Spatial indexing is the cornerstone in GIS and spatial databases, and exploring new ways to organize data has become an important way to meet the growing demand for data retrieval and query. To meet the needs of users' preferences for infrastructure near the target listing, in this paper, we propose a new hierarchical clustering algorithm based on CLUBS+ clustering by introducing Geohash precision and adding markers to subclasses, called New Retraceable Clustering Algorithm (NRCA).

## 2 Related Work

Scholars from various countries have never stopped their research on spatial indexing, but many of them are not applicable to GIS (Geographic Information System) systems

due to various reasons. Xu et al. [1] proposed a secondary index that supports two-dimensional data indexing and k-nearest neighbor algorithm query based on LevelDB and R-tree to solve the key-value database itself does not have the problem of the method of querying keys from their values and the problem of ensuring both the consistency of data and the timeliness of spatial data. However, its centroids need to set the number in advance, which is not applicable to geographic information data. xu et al. [2] constructed a hybrid index of hilbert information retrieval tree by introducing keyword inverted documents in nodes and integrating it with LDA topic model, although the query efficiency of semantically related texts was improved by topic classification, but the query efficiency for geographic information was not greatly improved. Based on k-means++, Zou et al. proposed a dynamic tree TR tree [3], which improved the indexing efficiency and inter-node correlation compared with the R tree based on location clustering, however, their research lacked the exploration of query efficiency improvement. Liu et al. [4] proposed a three-dimensional grid-R tree hybrid index structure, however, the characteristics of geographic information data were not considered, thus it could not However, the characteristics of geographic information data are not taken into account, and thus cannot be queried efficiently.

Clustering refers to the division of data into multiple classes or clusters. These classes or clusters are highly similar to each other with objects in the same cluster, while objects in different clusters are more different from each other [5]. It is divided into various ways [6], and the algorithm in this paper uses hierarchical based clustering. Hierarchical clustering refers to decomposing a given data set so that it is hierarchical until certain requirements are met [7]. Representative algorithms are BIRCH [8], CLUBS [9], CHAMELEON [10], etc.

### 3 New Retraceable Hierarchical Clustering Algorithm

#### 3.1 Splitting Strategy Based on Statistical Sample Distribution

The distribution of data within the class in each dimension is counted during the splitting process to determine the best splitting position. Considering the characteristics of GIS data, the method based on statistical data distribution does not require labels and training data and is more suitable for GIS database use. The suitable split location should satisfy the following two conditions:

- (1) The fluctuation of the line graph of the data distribution of the dimension must be above  $k(k < 1)$  times the fluctuation of the initial line graph.
- (2) The number of data within the larger data grouping after the split is greater than the mean number of data before the split for any dimension corresponding to all subclasses.

For condition (1), the fluctuation characteristics in the line graph can reflect the divisibility of the current class of data. The data fluctuation calculation process should be implemented by the standard values of the line graph. If the above two conditions are met, then the splitting position is the one with the least number of data, dividing the data set into two sub-data sets, and iterating such a process.

The splitting process is expressed by the following equation:

$$C_1 = \{P_j | P_j \in D \wedge p_i^j \leq h, i = 1 \sim d, j = 1 \sim n\} \quad (1)$$

$$C_2 = \{P_j | P_j \in D \wedge p_i^j > h, i = 1 \sim d, j = 1 \sim n\} \quad (2)$$

### 3.2 Merging Strategy with Additional Marker Detection

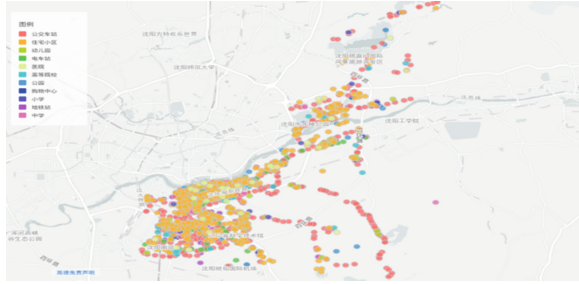
The algorithm has added a marker to each data during the splitting process, the main reason for this is to identify the source of the subclass splitting and to clarify the possibility of splitting the same data set. During the merge, the last layer of subclasses is first detected, in a bottom-up manner, all the way to the first layer of subclasses. That is, starting from the current digit marker to the end of the merge with digit marker 1. In the detection of each layer, it is sufficient to detect only the last  $r$  bits of all subclass markers ( $r$  is the number of bits of the hierarchical marker). The subclasses with the largest marker bits are considered to be the subclasses that should be detected immediately, and these subclasses must be obtained by splitting from the same data set. In addition, if there are subclasses that are not derived from the same class division at some levels, their merging at those levels can be performed in parallel, which greatly reduces the computational effort. The merging of subclasses requires the following two conditions:

- (1) The subclasses are not “disconnected” in any dimension.
- (2) After the merging process, the intra-class and inter-class similarities become higher and lower, respectively.

Accordingly, the measure of “disconnected” is also analyzed here using a line graph of the sample distribution. If two subclasses have no samples in the line graph of a certain dimension, they are disconnected in this dimension and will not be merged.

## 4 Algorithm Analysis and Experiments

This section shows the process of clustering with examples. The experimental data set is the spatial geographic information data of Hunnan District, Shenyang City, which contains location names, addresses, latitude and longitude coordinates, telephone and other information, and the location names, latitude and longitude are retained after cleaning for experiments. The data are visualized as shown in Fig. 1. It can be seen that the data set has different regions and different data distribution characteristics, such as locally dense, locally scattered, and locally uniform.



**Fig. 1.** Data Visualization

The clustering method in this paper uses the same evaluation index as the CLUBS + algorithm. That is, F-M value and ARI (Adjusted Rand Index). The greater the value of these two indexes, the better the clustering effect.

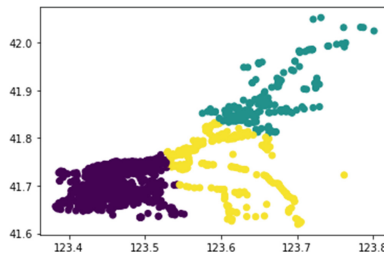
This algorithm uses a method similar to grid indexing to divide the region map, and uses the precision of Geohash as the grid size division criterion. The data density of each grid under different Geohash precision is calculated, and a suitable precision is selected as the basis of clustering, and the regions are classified as locally dense, locally uniform and locally dispersed according to the data density under this precision, and reasonable clustering is performed according to the region division.

As in Fig. 2, when the precision level of Geohash is selected as 4, this dataset is divided into 3 clusters with low intra-cluster similarity, and there are sparse points (e.g., top right corner) and dense points grouped into one cluster. This is due to the large range and small number of grids at precision of 4. And the grid range at precision of 4 is far beyond the boundary of the Huntington area.

As shown in Fig. 3. When the precision level of Geohash is selected as 5, this dataset is divided into 7 clusters with high intra-cluster similarity, and sparse points (e.g., top right corner) are separated from dense points. The grid division is more reasonable.

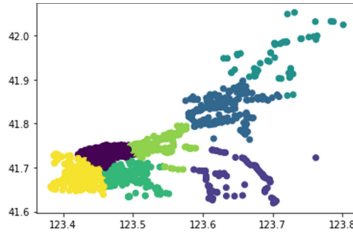
The accuracy level of Geohash is 6, the grid area is too small, the computational effort increases dramatically, the time consumption increases significantly, and it is difficult to distinguish dense and sparse regions with this accuracy, which loses the value of clustering.

Therefore, the most appropriate Geohash precision level for this dataset is 5.



**Fig. 2.** Clustering results at precision level 4





**Fig. 3.** Clustering results at precision level 5

For F-M and ARI values, CLUBS+, BIRCH, and k-means are used as comparison algorithms in this paper, and the comparison algorithm data are obtained from the experimental results of the original CLUBS+. Due to the different devices and programming languages, there is no comparison of running time in this paper, and the specific results are shown in Table 1.

As can be seen from the Table 1, in terms of F-M values, the F-M of the NRCA clustering algorithm decreases slightly as the number of clusters increases, but it is still slightly higher than the CLUBS+ and Boolean clustering algorithms, while the advantage is more obvious compared to the other comparison algorithms.

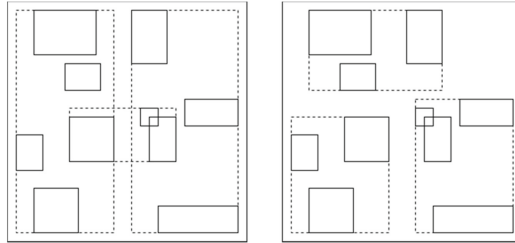
**Table 1.** Comparison with classical clustering algorithm at different number of classes

Methods	2 Clusters F-M	4 Clusters F-M	8 Clusters F-M	2 Clusters ARI	4 Clusters ARI	8 Clusters ARI
NRCA	0.994	0.992	0.986	0.993	0.991	0.984
CLUBS+	0.990	0.987	0.985	0.990	0.985	0.982
BOOL	0.987	0.984	0.976	0.985	0.980	0.976
BIRCH	0.972	0.905	0.808	0.951	0.845	0.742
k-means	0.934	0.909	0.895	0.904	0.826	0.745

For the ARI value, with the increase in the number of clusters, the NRCA clustering algorithm gradually becomes equal to the CLUBS+ algorithm, but still higher than the other comparative algorithms.

Compared with the classical clustering algorithm, the time complexity of this algorithm is not large. The time complexity of this algorithm is divided into the split phase time complexity and the merge phase time complexity, which are  $O(d \cdot n \log n)$ ,  $O(n)$ , respectively.  $n$  denotes the number of data and  $d$  denotes the data dimensionality. The overall time complexity of the algorithm is determined by the part with the highest time complexity, so the overall time complexity of this algorithm is  $O(n \log n)$ . According to Fig. 4, it can be found that the leaf nodes obtained by using the NRCA clustering

algorithm are more compact and have a small dead space area compared to the Hilbert R-tree. Based on the above analysis, a comparison of the characteristics of the present algorithm (New Retraceable Clustering Algorithm, NRCA) and common classical clustering algorithms is shown in Fig. 5.



**Fig. 4.** Delineation of leaf nodes of Hilbert R-tree before and after clustering with NRCA

Cluster \ Features	Parameters	Scalability	Data Shape	Data Properties	Time Complexity
NRCA	GEOHASH Accuracy	Higher	Arbitrary	Arbitrary	$O(n \log n)$
CLUBS+	—	General	Arbitrary	Numerical	$O(n * k)$
k-means	Number of classes	General	Convex	Numerical	$O(n)$
DBSCAN	Neighborhood radius	General	Arbitrary	Numerical	$O(n \log n)$
BIRCH	Radius threshold of the largest sample	Higher	Convex	Numerical	$O(n * k)$

**Fig. 5.** Comparison of NRCA and classical clustering algorithm

## 5 Conclusions

The NRCA clustering algorithm proposed in this paper improves on the CLUBS+ clustering algorithm. The statistical data distribution is utilized in the splitting phase to specify the best splitting position and obtain much more subclasses than the actual number of clusters. When entering the merging phase, the subclasses that were overdivided in the previous process are merged while making them into more correct classes. Compared with the CLUBS+ algorithm, this algorithm is relatively more accurate and effective, while breaking the limitation that the intermediate results of hierarchical clustering are not retraceable.

## References

1. Xu, R., Liu, Z., Hu, H., Qian, W., Zhou, A.: An efficient secondary index for spatial data based on LevelDB. In: Nah, Y., Cui, B., Lee, S.-W., Yu, J.X., Moon, Y.-S., Whang, S.E. (eds.) DASFAA 2020. LNCS, vol. 12114, pp. 750–754. Springer, Cham (2020). [https://doi.org/10.1007/978-3-030-59419-0\\_50](https://doi.org/10.1007/978-3-030-59419-0_50)

2. Yidan, X., Jinyu, H.: Research on hybrid mechanism based on Hilbert R-tree and LDA. *Comput. Simul.* **36**(012), 415–420 (2019)
3. Zou, C.W., Qin, C.: A k-means++ based method for dynamic construction of spatial topic R-trees. *Computer Applications* (2021)
4. Liu, Y.S., Gong, X., Kong, D.H., et al.: Research on hybrid indexing method based on 3D grid R-tree. *J. Yanshan Univ.* **44**, 48–59 (2020)
5. Hinneburg, A., Keim, D.A.: Clustering techniques for large data sets. In: *Tutorial Notes of the Fifth Acm Sigkdd International Conference on Knowledge Discovery & Data Mining* (2000)
6. Dong, P.: Efficient query and analysis system for distributed spatial information. Ph.D. thesis, Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing (2003)
7. Ma, D., Zhang, A.: An adaptive density-based clustering algorithm for spatial database with noise. In: *IEEE International Conference on Data Mining*. IEEE (2004)
8. Zhang, T., Ramakrishnan, R., Livny, M.: BIRCH: an efficient data clustering method for very large databases. *ACM SIGMOD Rec.* **25**(2), 103–114 (1996)
9. Mazzeo, G.M., Masciari, E., Zaniolo, C.: A fast and accurate algorithm for unsupervised clustering around centroids. *Inf. Sci.* **400**, 63–90 (2017)
10. Karypis, G., Kumar, V.: Chameleon: A hierarchical clustering algorithm using dynamic modeling. *IEEE Trans. Comput.* **32**(8), 68–75 (1999)



# Digital Transformation of the Smart Home Industry

Haijie Shen<sup>1,2</sup>(✉)

<sup>1</sup> College of Electronic Information Engineering, Xi'an Siyuan University, Xi'an 710038, Shaanxi, China

shenhaijie1981@163.com

<sup>2</sup> Mapúa University, 1002 Manila, Philippines

**Abstract.** In recent years, China has initially formed a smart home ecosystem. With the development of new generation information technology, the smart home industry has also entered a period of rapid development. Smart home can be used as an entry point for smart city construction, and the digital transformation of smart home industry is of great significance to promote the digital transformation of cities. However, there are currently problems in the smart home industry, such as the digital ecosystem is not yet complete, device brands isolated from each other, inability to realize deeply customized smart needs, possible leakage of user privacy, high threshold for elderly people to use, and failure to lead the digital transformation of cities. This paper proposes to solve the above problems by improving digital ecology and enterprise structure, creating a unified platform for industrial integration, strengthening infrastructure construction and data sharing, establishing a multi-party privacy protection credit mechanism, integrating smart home and smart community, and exploring humanized voice control terminals. We hope to find out the smart home industry development positioning, seize the opportunity, lead the new direction of the industry.

**Keywords:** Digital Ecosystem · Digital Transformation · Smart Home

## 1 Introduction

Smart home, is a systematic product that integrates automatic control, computer, Internet of Things, artificial intelligence and other technologies to achieve centralized management of functions such as home appliance control, environmental monitoring, video and audio entertainment, and information management with the user as the carrier and content as the center. Smart home industry is the product of digital transformation of home industry, but the digital transformation of smart home industry has just begun [1].

### 1.1 Background of Research

Markets estimates that the global smart home market will reach 122 billion US dollars in 2022, with an average annual growth rate of 14%. From the point of view of

demand analysis, the scale of China's smart home market will present a situation of rapid growth. Internationally, Google, Apple, Amazon and other global technology giants have arranged the smart home industry [2]. China's domestic Internet and home appliance giants such as Alibaba, Baidu, Tencent, Huawei, Xiaomi, Haier and Midea are also scrambling to seize the market, and large-scale market applications are landing rapidly [3].

5G, voice recognition, computer vision, cloud services and other next-generation information technologies are the core technologies leading the development of smart home. As the new generation of information technology continues to integrate into all areas of urban development, smart city is becoming an important symbol of world-class cities. In the last two years, the construction of smart cities has paid more and more attention to urban governance [4]. For the government, digitalization is the only way to realize the fine management of cities. Smart homes can serve as an entry point for smart cities.

## 1.2 Problems in the Digital Transformation of Smart Home

Smart home has gone through more than 10 years, but the digital transformation of smart home industry has just begun. At present, the common problems in the digital transformation of smart home mainly include the following.

- (1) Low market concentration, cross-brand or cross-category household products are difficult to truly achieve connectivity, is the pain point facing the development of smart home industry [5].
- (2) As the operation of some interactive devices in smart home is too complicated and not humanized, there are obstacles for the elderly to use smart home and they cannot enjoy the convenience of life brought by digitization.
- (3) With more and more user data being acquired, shared and applied, the problem of smart home security governance becomes prominent. Due to the lag in the construction of relevant laws and policies, there are security problems in the smart home field [6].
- (4) Smart home is an important booster of cities digital transformation [7]. Due to the lack of unified planning and coordination by the government, the development of smart home is chaotic, and it cannot exert the due effect of boosting the construction of digital community and forming a smart city.

## 1.3 Research Program

This paper proposes to solve the above problems by improving digital ecology and enterprise structure, creating a unified platform for industrial integration, strengthening infrastructure construction and data sharing, establishing a multi-party privacy protection credit mechanism, integrating smart home and smart community, and exploring humanized voice control terminals [8].

## 2 Research Content

### 2.1 Improving the Digital Ecosystem and Enterprise Structure

China has initially formed a smart home ecosystem. From the perspective of the industrial chain, smart home can be divided into upstream technology layer, midstream system layer and downstream single product layer, as shown in Fig. 1 [9].



Fig. 1. The Smart home ecosystem

The construction of the smart home digital ecosystem requires the joint efforts of various stakeholders. In addition, the complete smart home digital ecosystem also includes the following stakeholders:

(1) Government

Relevant government departments should complete the planning and construction of smart home standardization system as soon as possible. It includes overall standards, business standards, application standards, cloud platform technology standards, data management standards, terminal product standards, security standards, etc. At the same time, strengthen enterprise management, standardize the market, for the development of smart home to create a good social and economic environment.

(2) Industry associations

Smart home related industry associations should explore the connectivity between different brands of equipment, and strive to promote the formulation and implementation of relevant standards. Industry associations provide unified quality certification, safety certification also, to build a unified standard, good compatibility and testing evaluation, market supervision specification of smart home control system.

(3) Smart home enterprises

Manufacturers can control the cost of smart home system through information system integration, reduce the phenomenon of repeated investment, so that it can meet the consumption level of the majority of consumers, and promote the popularity and generalization of smart home products in China.

Smart home manufacturing enterprise should improve product and simplify the operation of the management system to better meet the actual needs of people's lives.

Third-party service institutions provide scheme customization and consulting services. Through the establishment of whole-house intelligent customization service company, help users to develop their personalized whole-house intelligent plan.

(4) ICT companies

When ICT companies provide data storage services, they need to ensure the independence, stability and security of data. ICT companies can also build a public IoT platform and provide IoT technical support.

(5) The users

Meeting users' humanized and personalized needs is the ultimate goal of various smart home products.

## 2.2 Standardization of Platforms

In the era of AIoT, smart home carries out big data analysis on the basis of collecting a large amount of data at the device end and connecting the bottom layer with the cloud, so as to realize active intelligence and whole-house intelligence.

Considering the large amount of sensitive data in the industry, the government can standardize the service and application of smart home products by formulating unified smart home device access standard agreement, and integrate smart home cloud platform into unified service management framework to deepen management, service and interaction [10]. Smart home system localization, cloud and marginalization are combined to improve the overall operation efficiency of smart home and bring better product experience to consumers.

The government can rely on the big data resource platform to establish and improve the subject database for smart home, promote the two-way sharing mechanism between smart home enterprises and the government, and provide support for the refined governance and services of urban communities.

In addition to unify the standard protocols of various smart home hardware at the policy level, it is necessary to accelerate the digital transformation of traditional home manufacturers and integrate the supply chain of the smart home industry.

## 2.3 Boost the Development of Smart Cities

The elderly who are accustomed to a traditional way of life often encounter “digital divide” in the process of digital transformation of urbanization. Smart home applications based on voice control can be developed to create smart home products with low learning costs and temperature, so that the aging society can keep pace with the digital era.

With the home intelligent control panel as the unified entrance, it focuses on application scenarios such as home network, online education, pension and health, building intercom, intelligent terminal distribution, intelligent security, parking management, and track capture of community activities, etc., to grasp the pain points of consumer demand and create an immersive terminal covering People’s Daily life. Urban digital transformation is a major strategy concerning the overall situation and the long-term. It is necessary to take digitalization as the lead, focus on “digitalization + human settlements” and “digitalization + community”, realize the penetration of technology, services and solutions into the residential space, and boost the construction of digital economy innovation and development highland.

## 3 Conclusions

This paper takes the digital transformation of smart home industry as the topic, explores the problems in digital transformation of smart home industry. And proposes to solve the above problems by improving digital ecology and enterprise structure, creating a unified platform for industrial integration, strengthening infrastructure construction and data sharing, establishing a multi-party privacy protection credit mechanism, integrating smart home and smart community, and exploring humanized voice control terminals.



We hope to find out the smart home industry development positioning, seize the opportunity, lead the new direction of the industry.

**Acknowledgement.** This work is supported by the first-class undergraduate project "Linux Operating System" in Shaanxi Province. Thank Dr. Torres William as my teacher led me to learn the course of "DIGITAL ECOSYSTEM AND DIGITAL TRANSFORMATION", his rigorous work attitude and high-level academic accomplishment let me deeply admire!

## References

1. Boley, H., Chang, E.: Digital ecosystems: principles and semantics. In: Proceedings of the 2007 Inaugural IEEE-IES Digital EcoSystems and Technologies Conference, pp. 398–403. IEEE (2007)
2. Jacobides, M.G., Sundararajan, A., Van Alstyne, M.: Platforms and ecosystems: enabling the digital economy. In: World Economic Forum Briefing Paper. World Economic Forum, Switzerland (2019)
3. Aulkemeier, F., Iacob, M.-E., van Hillegersberg, J.: Platform-based collaboration in digital ecosystems. *Electron. Mark.* **29**(4), 597–608 (2019). <https://doi.org/10.1007/s12525-019-00341-2>
4. Cozzolino, A., Corbo, L., Aversa, P.: Digital platform-based ecosystems: the evolution of collaboration and competition between incumbent producers and entrant platforms. *J. Bus. Res.* **126**, 385–400 (2021)
5. Ning, X.: Analysis on the characteristics and development trend of modernization of living quality—taking the popularization of smart home as an example. In: Proceedings of the 3rd International Modernization Forum, pp. 399–407 (2019)
6. Yang, Z., Cho, J.H.: Application and development trend of smart home in residential interior design. In: Proceedings of 4th International Conference on Control Engineering and Artificial Intelligence (CCEAI 2020), pp. 219–224 (2020)
7. Xie, Z., Qi, H., Wu, J.: Application of mobile monitoring system of smart home based on Wifi. In: Proceedings of 2019 4th International Industrial Informatics and Computer Engineering Conference (IIIEEC 2019), pp. 267–271 (2019)
8. Zhang, B., Li, J.: Research on interaction design of aging smart home products based on QFD/TRIZ. In: Proceedings of 2019 International Conference on Humanities, Cultures, Arts and Design (IHCAD 2019), pp. 294–299 (2019)
9. Souphamith, P., Yu, L.: The application of virtual reality technology in smart hom. In: Proceedings of 2019 International Conference on Big Data, Electronics and Communication Engineering (BDECE 2019), pp. 174–178 (2019)
10. Zhang, G., Zhang, S., Jin, J.: BCI for lighting control of smart home based on alpha-block. In: Proceedings of the 2nd International Conference on Robotics, Control & Automation Engineering (RCAE 2019), pp. 71–74 (2019)



# Identity Authentication Method of IoT Equipment Based on Alliance Chain

Zhongzheng Xiang<sup>1</sup>, Yanjun Hu<sup>2</sup>, Xiaokang Chai<sup>1</sup>, Kai Zheng<sup>2</sup>(✉),  
and Xuemei Shen<sup>2</sup>

<sup>1</sup> State Grid Zhejiang Electric Power Co., Ltd., Hangzhou 310007, China

<sup>2</sup> State Grid Zhejiang Electric Power Co., Ltd., Hangzhou Electric Power Company,  
Hangzhou 310001, China  
614658757@qq.com

**Abstract.** IoT applications play an important role in today's society, where devices from different domains collaborate on the same task, and device authentication technologies are urgently needed to ensure the security of IoT applications. In addition, due to the wide distribution of IoT devices, complex application scenarios, and large differences in node performance, their identity authentication efficiency is low. The existing authentication methods have problems such as high key management overhead and reliance on trusted third parties. Therefore, an overall architecture of distributed storage for IoT device identity authentication is proposed. It mainly includes the distributed IoT device identity authentication mechanism and the detailed process of device registration consortium chain and consortium chain verification device. The performance analysis of the identity authentication method shows that the block creation time of different devices is less than 0.55 s, the IoT devices can join the alliance chain, and the identity authentication mechanism is real and effective.

**Keywords:** Blockchain · Authentication · Internet of Things · Smart Contract

## 1 Introduction

With the development of the Internet of Things technology, from the access of scattered devices to the Internet of Everything, the changes in the scope, mode, channel, and efficiency of social communication are realized. A key application scenario of IoT is the combination of industrial scenarios, which is also known as Industry 4.0[1]. However, while the Internet of Things brings convenience to people's work and life; it also brings about a series of new technical and social problems [2]. Ubiquitous connections will lead to severe security risks, so the authentication of IoT devices is extremely important [3]. However, with the further growth of IoT interconnectivity, while bringing more opportunities to production and manufacturing, it also creates an environment for various types of cyber-attacks. Therefore, it is necessary to authenticate the special data collected by the IoT device layer in order to maintain IoT security efficiently and safely. In this paper, we design a consortium chain consisting of the main body of the consortium

chain, low-level device nodes, and high-level device nodes. Low-level device nodes are IoT device nodes, and high-level device nodes are nodes represented by institutions or organizations. The alliance chain provides the decentralized storage of this node information and the storage of identity authentication results. Each organization uses smart contracts to authenticate the identity of IoT nodes and the use of IoT devices. The alliance chain is obliged to manage IoT node information. The networking node applies to multiple organizations to join the alliance chain to complete the initialization. IoT nodes must complete IoT device registration in the alliance chain, upload IoT node information, and submit identity information to the alliance chain when they apply to join the network again. When joining the network for the second time, the main authentication information needs to be submitted to the constructed alliance chain, and the high-level nodes need to accept the request and then create a wallet. The low-level device node also associates the device to be authenticated with the authenticated device node to facilitate subsequent operations.

## 2 Related Work

In recent years, identity authentication technology has solved some of the security problems of the Internet of Things, providing an effective idea for ensuring the security of the Internet of Things. Compared with the traditional identification technology, the identity authentication strategy based on biometrics has the advantages of high security and unique features, but this method cannot perform M2M authentication and is not suitable for general industrial environments; Li et al. [4, 5] proposed a two-factor identification method that combines biometric information and password; Deebak et al. [6] proposed to combine smart card and biometric information for authentication, but this method requires some special hardware to achieve the purpose of identification and authentication; Beltran et al. et al. [7] proposed a token-based federated identity authentication method to solve the energy limitation problem of low-power devices in some scenarios of the Internet of Things.

The above research results are all based on traditional identity authentication methods, and their centralized structure is difficult to meet in today's Internet of Things system. As a decentralized technology, blockchain has the characteristics of non-tampering, traceability, openness, and transparency, which can be used as a guarantee for network security. Bartolomeu et al. [8] proposed a self-sovereign identity (SSI) technology based on the Internet of Things, which provides a decentralized identity identifier (DID) based on blockchain technology to build a model that does not require centralized trust; Yu et al. [9] proposed a selective revocation anonymous authentication method for smart industrial applications based on blockchain to solve the problem of credential revocation in smart industrial environments; Shen et al. [10] proposed a blockchain-assisted security device authentication mechanism BASA, to solve the authentication problem of devices between different domains; Zhang et al. [11] constructed a large-scale heterogeneous WSNs collaborative authentication protocol to ensure the reliability of data sources in the IoT environment; Wang et al. [12] designed a collaborative authentication protocol, including a memory-efficient and fast validator locator, a trust model to assess device trustworthiness, and a protocol suite for dynamic update and revocation verification to efficiently authenticate IoT devices.

Although the above methods solve the reliability problem of IoT terminal devices to a certain extent, the centralized identity authentication method has the problem that if the server is attacked, the whole system will be unusable. Blockchain-based identity authentication the method has a limited scope of application. Since most of the alliance chains are based on the credit generated by the collective endorsement of multiple organizations or institutions, traditional blockchains' efficiency and cost advantages can be greatly improved [13]. Therefore, this paper proposes an IoT device identity authentication method based on the alliance chain, which solves the problem of mutual distrust among multiple nodes and is of great significance for promoting the large-scale security application of IoT technology [14].

Considering the above problems, this paper designs an IoT device identity authentication method based on the consortium chain. The main purpose is to store the authentication result in the consortium chain while realizing the identity authentication of the IoT device [15]. When an IoT device needs to be used, the owner of the device can store user information such as username and password on the server. The server is generally represented by a centralized organization. Once a user or device needs to log in, the user's stored information can be used for comparison through the centralized server. If the comparison is successful, the user device is allowed to log in. At the same time, the information of these IoT nodes is identified by an asymmetric encryption algorithm using a public key, because, during the specific communication process, the identity information of IoT devices and the security process of blockchain node information verification need to be protected [16].

### 3 The Specific Method of Identity Authentication

When the first request of the IoT node is sent to the consortium chain, the consortium chain processes the sent information, and the consensus generates a new block according to the algorithm. The alliance chain uses the information processing module to process the information of the IoT nodes, and at the same time sends the processing results to the external interface layer. If the device does not request authentication for the first time, the consortium chain will verify the node with existing data, mainly including locally stored or previously maintained data. Figure 1 illustrates the overall architecture of the system.

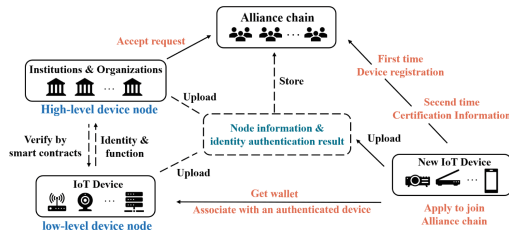
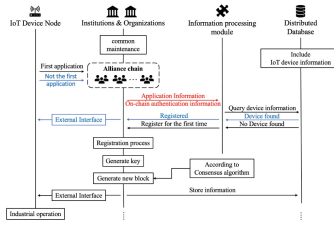


Fig. 1. Overall architecture of the system



**Fig. 2.** Overall Process Flow Chart

Among them, the main functions of each module are as follows:

- 1) Institution or organization: As the target object of the alliance chain, institutions and organizations mainly select a node for accounting operations. At the same time, it can query the information it wants to know through the smart contract, so as to carry out subsequent operations.
- 2) Alliance chain: As a variant application of blockchain, the alliance chain mainly realizes the functions of creating IoT device identity, uploading IoT device data, and IoT device identity authentication.
- 3) IoT devices: IoT devices are the foundation of the entire consortium chain, which have certain computing and storage capabilities and meet network transmission functions. In this method, the IoT device is mainly responsible for the storage of the private key and the functions of information exchange and transmission with the consortium chain.

The most important node in the alliance chain is the low-level device, that is, the IoT device, because the information provided by the IoT device is the core element that constitutes this network. IoT devices exchange and transmit information with alliance chains through various communication protocols, such as WiFi, 4G, Lorea, and NB-ITT. When initializing the device, you first need to initiate a registration request, and add the device to the consortium chain network by calling the API provided by the organization. Once the device registration is completed, the alliance chain will create a wallet for the device, and the private key of the wallet will be stored on the device. The overall flow of the process is shown in Fig. 2.

The equipment registration alliance chain process is as follows:

- Step 1: Initialize the device, and then the IoT device initiates a registration request to the alliance chain;
- Step 2: The alliance chain obtains unique identification information such as IoT device ID by using the information requested by the device to register;
- Step 3: The alliance chain invokes the security control function to issue a registration application to the institution;
- Step 4: If the registration is successful, the institution or organization obtains the device information to obtain the corresponding permissions;

Step 5: After successful registration, the alliance chain will create a wallet for the node and send the private key to the node after creation.

The specific registration process is shown in Fig. 3.

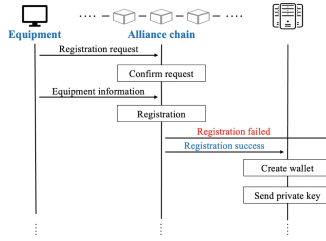


Fig. 3. Equipment Registration Process Chart

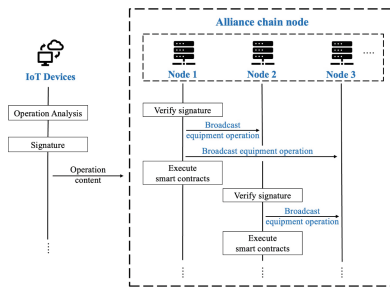


Fig. 4. Equipment Certification Process Diagram

The authentication process of the IoT device identity authentication function is mainly composed of 4 steps: first, the public key is converted into its own special ID through the SHA 256 algorithm; when the ID is generated, the physical device sends a request to join the alliance chain; The public key of the node’s identity ID is stored in the block, and the information is stored in the current blockchain network; finally, the specific operation process and operation authority of the low-level node are set using smart contract technology. The user of this device can then remotely operate the device in the form of blockchain “transactions”, query device information, add IoT devices to the network, etc.

In terms of security, in order to improve the security of the entire blockchain network, the system will pre-check the operation of IoT devices. In order to prevent malicious or dangerous nodes from sending wrong requests, the consortium chain uses this to determine the functions that the device can have after executing the smart contract. The process of consortium chain verification equipment is shown in Fig. 4.

## 4 Experiment and Evaluation

Our simulation experiments are performed on four computers with one CPU as Intel i7, NVIDIA MX50 as GPU, and 8 G RAM; 2 CPUs as Intel i5, GPU as Intel Graphics620

and NVIDIA MX150, and RAM as both 4 GB computer; 1 CPU is Intel 8th generation i5, GPU is Intel UHD Graphics620, RAM is 4 GB computer, and the above 4 devices are correspondingly simulated into 4 organizations, on which Linux system is installed and Java is configured PairingBased Cryptography related environment. We use MATLAB R2020a to simulate this scenario. In order to simulate the accuracy of the experiment, we test the reliability of the method from the time required to generate wallets, block generation time and overall identity authentication from different devices.

### 4.1 Time for 4 Devices to Generate Blocks

Because the IoT node identity authentication scheme of the alliance chain will be accompanied by the generation of new blocks, the specific information operation process can also be stored in the block. Based on the analysis of the experimental results of block generation, this paper also studies the time required to create a block and compares the time required to create a new field faster on 4 different physical platforms. From Fig. 5, we can see that the generation time of the four IoT device nodes is within an appropriate and reasonable range, and as more and more IoT device nodes are generated, the growth rate will slow down. This is because the Byzantine algorithm is slower when the number of blocks is relatively large because network communication is important and nodes must coordinate to generate new blocks.

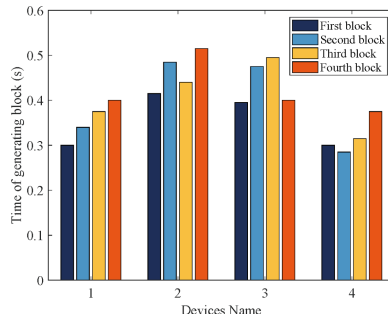


Fig. 5. Time of block generation by different devices

### 4.2 The Speed at Which 4 Devices Generate Wallets

The process of creating a wallet for a consortium chain device is very similar to that of an IoT device. Both use cryptography-related technologies to create wallets for users and at the same time store the private key in the node. Broadcast. Figure 6 illustrates the speed at which different institutions or organizations generate wallets. The number of generated wallets is not affected by the device factor, but as the number of IoT nodes increases, the speed of wallet generation will increase slightly. When a node needs to generate a new wallet, there will be a new block to store specific transaction information. But the generation speed of the wallet will slowly increase with the increase of nodes, which is because of the growth of the block that stores the transaction information.

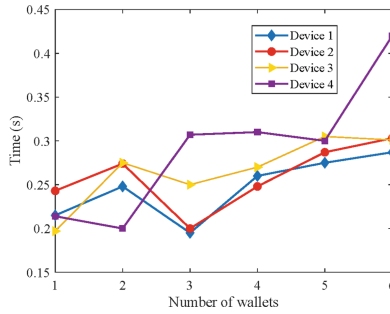


Fig. 6. Number of wallets generated by different devices

### 4.3 Analysis of Identity Authentication Results

This paper focuses on the authentication of IoT devices. After the IoT node completes the creation of the wallet, these devices need to join the consortium chain network, and in this process, it is necessary to authenticate these devices to ensure the security of the IoT. This section mainly tests whether these IoT devices can successfully join the alliance chain. The analysis of the identity authentication results is shown in Fig. 7. According to the experimental results, it can be seen that the required time for successful and unsuccessful authentication is different. The difference in time lies in the post-authentication operation. If the authentication is successful, the organization will perform access control on each device, and if the authentication fails, it will not.

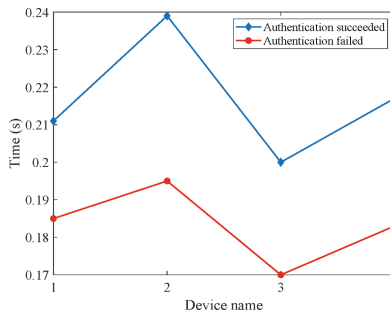


Fig. 7. Time spent on equipment certification

## 5 Conclusions

This paper builds a consortium chain-based IoT identity authentication scheme on the basis of blockchain to realize a decentralized identity authentication system and peer-to-peer communication between blocks. In addition to the information of the block header, each block also contains transaction input and transaction output related to IoT identity authentication. Once recorded in a block, these data are irreversible. After the blockchain



system is completed, through the analysis and testing of the platform, the platform is more in line with the specific scenarios of the Internet of Things, thereby promoting the combined development of the Internet of Things and blockchain technology. However, from the perspective of the entire IoT security identity authentication system, the frequent operation of devices uploading and broadcasting transaction information in the alliance chain puts forward higher requirements on the storage capacity and computing power of the device, which is difficult to meet the needs of large-scale IoT scenarios. At the same time, this paper mainly focuses on the security of the blockchain data layer, and there is no complete authentication system for the Internet of Things from the network layer to the application layer.

**Acknowledgements.** This work is supported by Science and Technology Project of State Grid Zhejiang Electric Power Co., Ltd (B311HZ21000G).

## References

1. Tange, K., Donno, D.M., Fatoutis, X., et al.: A Systematic survey of industrial internet of things security: requirements and fog computing opportunities. *IEEE Commun. Surv. Tutorials* **22**(4), 2489–2520 (2020)
2. Wu, J.Y., Li, W.J., Cao, J., et al.: AIoT: a taxonomy, review and future directions. *Telecommun. Sci.* **8**, 1–17 (2021)
3. Zhou, W., Jia, Y., Peng, A.N., et al.: The effect of IoT new features on security and privacy: new threats, existing solutions, and challenges yet to be solved. *IEEE Internet Things J.* **6**(2), 1606–1616 (2019)
4. Li, X., Niu, J., Bhuiyan, M.Z.A., et al.: A robust ECC-based provable secure authentication protocol with privacy preserving for industrial internet of things. *IEEE Trans. Industr. Inf.* **14**(8), 3599–3609 (2018)
5. Li, X., Peng, J., Niu, J., et al.: A robust and energy efficient authentication protocol for industrial internet of things. *IEEE Internet Things J.* **5**(3), 1606–1615 (2018)
6. Deebak, B.D., Al-Turjman, F., Aloqaily, M., et al.: An authentic-based privacy preservation protocol for smart E-healthcare systems in IoT. *IEEE Access* **7**, 135632–135649 (2019)
7. Beltran, M., Calvo, M., Gonzalez, S.: Federated system-to-service authentication and authorization combining PUFs and tokens. In: *International Symposium on Reconfigurable Communication-Centric Systems-on-Chip, Spain* (2017)
8. Bartolomeu, P.C., Vieira, E., Hosseini, S.M., et al.: Self-sovereign identity: use-cases, technologies, and challenges for industrial IoT. In: *IEEE International Conference on Emerging Technologies and Factory Automation, Zaragoza* (2019)
9. Yu, Y., Zhao, Y., Li, Y., et al.: Blockchain-based anonymous authentication with selective revocation for smart industrial applications. *IEEE Trans. Industr. Inf.* **16**(5), 3290–3300 (2020)
10. Shen, M., Liu, H., Zhu, L., et al.: Blockchain-assisted secure device authentication for cross-domain industrial IoT. *IEEE J. Sel. Areas Commun.* **38**(5), 942–954 (2020)
11. Zhang, L.P., Li, F.Q., Wang, P.C., et al.: A blockchain-assisted massive IoT data collection intelligent framework. *IEEE Internet Things J.* **9**, 14708–14722 (2022)
12. Wang, M., Qian, C., Li, X., et al.: Collaborative validation of public-key certificates for IoT by distributed caching. *IEEE/ACM Trans. Networking* **29**(1), 92–105 (2020)
13. Tang, J., Gao, D., Shao, S., et al.: Construction of alliance system model based on blockchain technology-taking the catering industry as an example. *Dahe Fortune China Forum And Chinese High-educational Management Annual Academic Conference, Zhengzhou* (2020)

14. Shafique, K., Khawaja, B.A., Sabir, F., et al.: Internet of Things (IoT) for next-generation smart systems: a review of current challenges, future trends and prospects for emerging 5G-IoT scenarios. *IEEE Access* **8**, 23022–23040 (2020)
15. Chegini, H., Naha, R.K., Mahanti, A., et al.: Process automation in an IoT-fog-cloud ecosystem: a survey and taxonomy. *The Internet of Things* **2**, 92–118 (2021)
16. Hewa, T., Ylianttila, M., Liyanage, M.: Survey on blockchain based smart contracts: applications, opportunities and challenges. *J. Netw. Comput. Appl.* **177**, 102857 (2021)



# Financial Accounting Risk and Prevention System Construction Analysis Based on Fuzzy Influence Diagram Evaluation Algorithm

Sitong Chen<sup>(✉)</sup>

Beijing Institute of Technology, Beijing 100081, China  
chen05090333@163.com

**Abstract.** At present, the research on financial risk focuses on market risk, liquidity risk, credit risk, operational risk and legal risk of financial instruments(FI), and there is no research on accounting risk(AR) of FI. Based on the fuzzy influence diagram(FID) evaluation algorithm, this paper studies and analyzes the financial accounting risk(FAR) and the construction of the prevention system, briefly analyzes the concept of FAR, the causes of risk formation, and puts forward preventive measures; The related definition and transformation principle of FID, the construction of FID and the evaluation algorithm of FID are discussed; Finally, through the analysis of FARs, the paper constructs a FAR prevention system, which is of great significance for the reduction of FARs in the future.

**Keywords:** Fuzzy Influence Diagram · Evaluation Algorithm · Financial Accounting · Prevention System Construction

## 1 Introduction

Nowadays, all kinds of products obtained and rights and interests formed by enterprises should be accounted according to the actual cost or actual amount incurred at the time of acquisition or formation. In the case of drastic changes in prices, the historical cost principle will not apply, and the relevance of the accounting information provided thereby can be imagined. Accounting standards are the norm for accountants to engage in accounting work, but there are a lot of uncertain words in them, which leads to too much prompting accountants to use professional judgment, which will inevitably lead to the subjective arbitrariness of accountants in dealing with accounting business. FA has this big risk, and the urgent task at present is to build a FAR prevention system. For this reason, this paper proposes a FID evaluation algorithm, which is studied and analyzed.

In the late 1990s, China began to take “AR” as an academic concept and began to study it. The researchers focus on defining the concept and classification of AR and how to prevent AR. Yiwei Dou gave a comprehensive explanation of AR. This document defines the concept of AR. The article points out that AR is mainly the risk of monitoring the loss caused by the authorities due to false accounting information, and believes that external FA reports are the main factors that generate AR, and that AR

is mainly related to FA reports [1]. Shengli Dai believes that macroeconomic factors, socio political and legal factors, and socio-cultural factors affect AR in China. The macroeconomic factors mainly include the diversification of accounting objectives and the economic consequences of accounting; Social, political and legal factors mainly refer to the conflict between accounting objectives and the status of accountants, and the lag of laws. In a word, the common point of current research on AR is that the concept of AR is basically analyzed, and quantitative measurement and empirical research on AR are relatively insufficient [2].

This paper sums up the definition of AR by drawing on and synthesizing the different understandings of various scholars on the definition of AR. This paper defines AR as: due to the wrong choice of accounting policies, the abuse of accounting elements, the misuse of accounting techniques, the lack of competence and quality of accountants, and other factors, the financial statements of enterprises or other information carriers that reflect the financial status of enterprises are ultimately wrong and incomplete, and accounting information users, enterprises, regulatory agencies and other relevant personnel or organizations Risk of loss caused by the institution. On this basis, a FAR and prevention system based on FID evaluation algorithm is proposed [3, 4].

## 2 Research on FAR

AR refers to the risk caused by false financial report. AR can be subdivided as shown in Fig. 1.

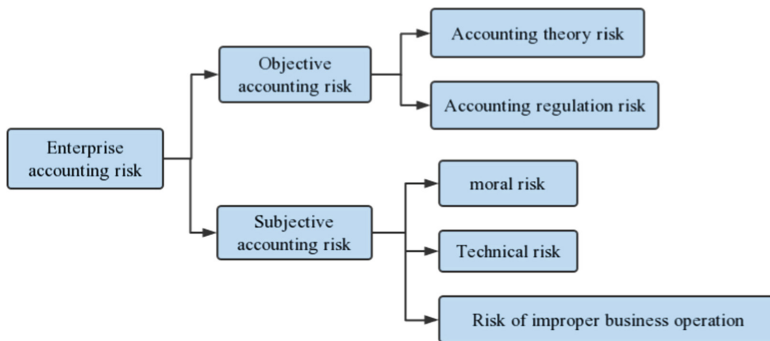


Fig. 1. Enterprise AR composition

### 2.1 Characteristics of AR

#### (1) Uncertainty of AR

The concept of risk itself refers to the possibility of loss, which is only a possibility rather than a certainty. AR is hidden in the daily work of accounting. When there are no factors inducing AR, AR will not be converted into real losses. Only when conditions are ripe can this potential uncertainty be converted into real losses [5].

## (2) Severity of AR

Accounting information has economic consequences. The false accounting information not only damages the economic interests of the enterprise, but also brings the loss of the image of the enterprise, which leads to the credit crisis of the enterprise, and also affects the healthy development of the market economy. At the same time, in view of the wide use of accounting information, involving suppliers, investors, banks and other relevant interest institutions, the inaccuracy of accounting information will also mislead the economic decision-making of these enterprises [6, 7]. From a macro perspective, accounting information can guide resource allocation, while inaccurate accounting information leads to inefficient resource allocation, inferior goods market and adverse selection, which will result in waste of resources and is not conducive to the healthy development of the economy.

## 2.2 Classification of ARs

The classification methods of ARs mainly include the following:

According to whether AR can be controlled, it can be divided into controllable risk and uncontrollable risk. As the name implies, controllable risk is the AR that can be controlled, prevented and predicted [8]. Uncontrollable risks are ARs that cannot be controlled, prevented and predicted.

According to whether AR can be avoided, it can be divided into systematic risk and non systematic risk. System risk refers to the risk of influencing the authenticity of accounting information due to the limitations of accounting itself, such as accounting assumptions, accounting systems, accounting standards and other reasons when each link in the accounting system is playing its own role. Non systematic risk is the risk caused by factors other than the accounting system. Non systematic risks mainly include accounting technology risks, accounting behavior factors, accounting management factors and accounting legal factors [9].

### (1) Impact of accounting environment

Here, the accounting environment refers to the external macro environment and the internal business management environment of the enterprise. Although the emergence of AR is directly related to the problems of the accounting system itself, the accounting system is also part of the external macro environment, so AR will inevitably be affected by the external macro environment. The external macro environment includes market economic environment, law, culture and other aspects, which have more or less influence on AR [10]. For example, when the economic system is transformed from a planned economy to a market economy, due to changes in the economic environment, accounting treatment has to become more and more complex. In addition, there are many ambiguities in the formulation of accounting related laws in China, which will reduce the function of laws to restrict the self-discipline of relevant personnel, thus increasing the AR. A healthy and orderly operating environment of an enterprise can ensure the authenticity and reliability of accounting information [11].

(2) Impact of accounting theory itself

Accounting theory is the action guide of accounting practice. A complete and scientific accounting theory is undoubtedly conducive to the development of accounting practice. If the accounting theory has its own insurmountable limitations, it will lead to the distortion of accounting information, which will lead to the risk of loss. Accounting is a very practical work, but its theoretical structure is composed of a series of abstract concepts. In particular, the accounting assumptions, accounting standards and general principles of accounting make the basis of accounting not so reliable and credible. It can be found from the above that the factors that affect the size of AR include people, materials, time and other aspects, which all have an impact on AR [12].

### 3 Evaluation Algorithm of FID

#### 3.1 Relevant Concepts of FID

The influence diagram theory, formed in the early 1980s, is a relatively new decision-making analysis method and tool. It can be represented by  $G = (N, A)$ . Where,  $G$  represents the influence diagram,  $A$  represents the directed arc set (information arc, influence arc, correlation arc, and forgetting arc),  $N$  represents the composition of node set and decision node set  $D$ , that is,  $N = (C, D, V)$  represents various interrelationships between variables. The specific meanings and representations of the three types of nodes are shown in Fig. 2.

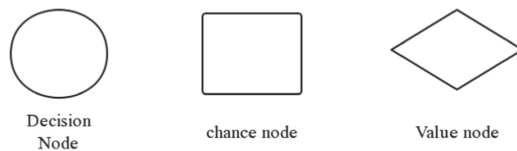


Fig. 2. Schematic Diagram of Node Types

Decision node  $d$  ( $d \in D$ ): the optimal decision  $d^*$  is a mapping corresponding to it, indicating that the arrow node is the optimal choice when the node information of the directed arrow tail is known;

Opportunity node  $x$  ( $x \in C$ ): also known as uncertain node. In mathematical sense, its corresponding mapping is a conditional fuzzy distribution function  $\Pi x, \Pi x = P(x|C(x))$ ;

Value node  $v$  ( $v \in V$ ): in mathematical sense, its corresponding mapping is the utility function  $U, \Omega c(v) \rightarrow \Omega U$ , indicating that the expected value of utility can be expressed by the function of the value node's predecessor node.

#### 3.2 Relevant Definitions and Transformation Principles of FID

Definition 2: If there are two or more directed paths from node  $x_i$  to node  $x_j$  at the same time, then node  $x_i$  can be said to be the multi way preorder node of node  $x_j$ .

Definition 2: If the influence diagram meets the following three conditions at the same time, it is called the specification influence diagram: there is no loop in the diagram; There are value nodes without subsequent nodes in the graph; There is a directed path containing all decision nodes in the graph. When the third condition becomes that there is only one value node in the graph, the influence diagram is called the directional specification FID.

### 3.3 Structure of FID

FID is a commonly used method and modeling tool, which has the characteristics of visualization and intuition, and is suitable for people to accept. After simplification and computer programming, the biggest difficulty is how to construct the influence diagram. The FID describes the object of study from three levels: relationship level, numerical level and function level. The three levels represent different research depths. The construction idea of fuzzy influence map is as follows:

Determine the function layer. Quantitative analysis methods such as mathematical statistics can be used. In the actual indemnificatory housing BT financing project, the function level is generally determined by referring to similar projects and historical statistical data through expert survey and scoring method.

Determine the numerical layer. In the FID, the determination of the numerical level refers to giving the variable state of the problem studied in the diagram and its corresponding fuzzy distribution value, which generally includes edge blur and condition blur.

In the process of using the FID to analyze the risk of the actual indemnificatory housing BT financing project, the determination of the numerical level is generally synchronous with the determination of the functional level, which determines the fuzzy relationship between the risk factors, and the numerical level determines the specific fuzzy value of the risk factors. In the actual safety risk analysis of cable-stayed bridge foundation construction project, the numerical layer is generally determined by referring to similar projects and historical statistical data through expert investigation and scoring method and function layer.

The fuzzy influence map has two basic construction methods: target oriented construction and two-way construction:

The fuzzy influence map generally adopts the target oriented construction method. First, the boundary nodes in the influence map are divided by disciplines, giving full play to the professional advantages of experts in different fields, and reducing the difficulty in the construction and estimation of the influence map. Then, all local influence maps are integrated and expanded. After reaching a satisfactory scale, a detailed inspection is carried out. The parts that can be merged are merged. Some insignificant factors are discarded, and the influence map is simplified. Finally, an influence diagram model with reasonable structure and scale and beautiful structure is obtained.

### 3.4 Evaluation of Fuzzy Influence Map

Independent node evaluation algorithm. An independent node means that the node has no antecedent node and no other node points to it on the graph. In the FID, it is shown

in Fig. 3. Let X represent a node without preorder. Assume that the possible state vector of node X is:

$$P_X = \{P_{X_1}, P_{X_2}, P_{X_3}, \dots, P_{X_n}\} \quad (1)$$

Among them,  $P_{X_1}, P_{X_2}, P_{X_3}, \dots, P_{X_n}$  are fuzzy sets defined by language vocabulary. The frequency vector of independent node X is:

$$f_x = \{f_{x_1}, f_{x_2}, \dots, f_{x_n}\}^T \quad (2)$$

where,  $f_{x_1}, f_{x_2}, \dots, f_{x_n}$  are the frequency ambiguity sets corresponding to each possible state in the frequency vector of node X.

Then, the frequency matrix of independent node X is:

$$F_X = (f_{x_1} \times P_{X_1}) \cup (f_{x_2} \times P_{X_2}) \cup \dots \cup (f_{x_n} \times P_{X_n}) \quad (3)$$

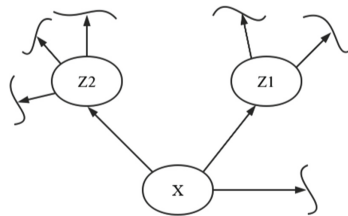


Fig. 3. Schematic Diagram of Independent Nodes

## 4 Analysis of FAR and Prevention System Construction Based on FID Evaluation Algorithm

### 4.1 Prevention and Management of FARs

#### (1) Improve the fair value measurement method of FA instruments

It is a general trend for derivative FI to adopt the measurement attribute of fair value to be included in the form of financial statements. At present, we should focus on the valuation technology of the fair value of derivative FI. This paper argues that China can use the "B-S option pricing model" and matrix method to value the fair value of derivative FI on the basis of the present value method, which is widely used at this stage.

#### (2) Valuation techniques for improving fair value measurement

The application of B-S option pricing model in China has great applicability. The main variables involved include price and time, and the variables involved are few. Moreover, any derivative financial instrument is determined by the price and time of the basic financial instrument. B-S option pricing model can also be used for other derivative FI other than options as long as other variables are added, which

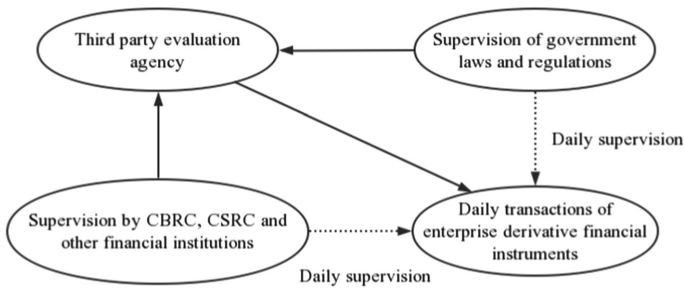


can expand its applicability. However, as mentioned above, derivative FI are based on a series of assumptions, and the current economic conditions may not meet their assumptions. Many experts in academia expect to expand the application scope of B-S option pricing model and relax its assumptions, so these scholars are committed to studying the expansion of the model. And with the extensive use of computer technology, B-S option pricing model should be widely used.

**4.2 Construction of FAR Prevention System**

- (1) Strengthen the supervision of fair value measurement and the construction of authoritative valuation institutions

The volatility of the value of derivative FI and the characteristics of financial leverage make derivative FI have greater volatility on the data in the financial statements of enterprises, especially when the external economic environment is more volatile. Therefore, the determination of the fair value of derivative FI should not only proceed from the technology, but also solve the fairness problem of fair value beyond the technology, and prevent the behavior of profit manipulation using fair value. Therefore, this paper constructs a supervision mechanism for the fair value valuation of derivative FI as shown in Fig. 4.



**Fig. 4.** Supervision mechanism for fair value valuation of FI

First of all, the government should establish relevant laws and regulations to increase the illegal costs of enterprises and relevant personnel as a daily supervision of enterprises. At the same time, we should establish a third-party authority responsible for the value evaluation of the fair value, and change the current situation that the fair value evaluation is completed by the enterprise itself. This can not only give play to the professionalism of the third-party organization, but also reduce the subjectivity and free operation of the enterprise. Of course, the rights of third-party appraisal agencies should also be supervised by CBRC, CSRC and other regulatory agencies, and their evaluation results should be supervised. Establish a series of mechanisms based on this to ensure the fairness of the value of derivative FI.

- (2) Improve accounting information disclosure of derivative FI  
 First of all, we should add a “detailed list of derivative FI”. Second, a simple detailed list of derivative FI only describes the value and changes of derivative FI in figures.

From this table, we can not see the risks involved in the derivative FI that enterprises are trading. It should be explained from the quantitative and qualitative perspectives of credit risk, liquidity risk and market risk. For other types of risks that cannot be quantified at present, a detailed explanation should be made in the notes so that the statement users can know the risk level of derivative FI.

## 5 Conclusions

For a long time, foreign and domestic scholars have been actively engaged in the research on various risks of derivative FI. This paper puts forward the FID evaluation algorithm, studies the FAR and constructs the prevention system. Due to the limitation of my knowledge and the lack of research time, the research on the problems related to FAR in this paper is only a general study, for example, the analysis of the problems in FA treatment is relatively shallow, and no in-depth research has been done; As the research content belongs to a very frontier issue in the field of FA, the research process is faced with many difficulties, and there is still room for further improvement in the feasibility of AR control measures for FI.

## References

1. Dou, Y.: The debt-contracting value of accounting numbers and financial covenant renegotiation. *Manag. Sci.* **66**(3), 1124–1148 (2020)
2. Dai, S.: ARS interactive teaching mode for financial accounting course based on smart classroom. *Int. J. Emerg. Technol. Learn.* **14**(3), 38–50 (2019)
3. Bos, T., Frasinca, F.: Automatically building financial sentiment lexicons while accounting for negation. *Cogn. Comput.* **14**(1), 442–460 (2021). <https://doi.org/10.1007/s12559-021-09833-w>
4. Hamal, S., Senvar, Ö.: A novel integrated AHP and MULTIMOORA method with interval-valued spherical fuzzy sets and single-valued spherical fuzzy sets to prioritize financial ratios for financial accounting fraud detection. *J. Intell. Fuzzy Syst.* **42**(1), 337–364 (2022)
5. Kalantonis, P., Delegkos, A.E., Sotirchou, E., Papagrigoriou, A.: Modern business development and financial reporting: exploring the effect of corporate governance on the value relevance of accounting information – evidence from the Greek listed firms. *Oper. Res.* **22**(3), 2879–2897 (2022)
6. Dhar, S., Chowdhury, M.A.F.: Impact of environmental accounting reporting practices on financial performance: evidence from banking sector of Bangladesh. *Int. J. Asian Bus. Inf. Manag.* **12**(1), 24–42 (2021)
7. Hamal, S., Senvar, Ö.: Comparing performances and effectiveness of machine learning classifiers in detecting financial accounting fraud for Turkish SMEs. *Int. J. Comput. Intell. Syst.* **14**(1), 769–782 (2021)
8. Shen, J., Han, L.: Design process optimization and profit calculation module development simulation analysis of financial accounting information system based on particle swarm optimization (PSO). *ISEB* **18**(4), 809–822 (2019). <https://doi.org/10.1007/s10257-018-00398-0>
9. Dameri, R.P., Garelli, R., Resta, M.: Neural networks in accounting: clustering firm performance using financial reporting data. *J. Inf. Syst.* **34**(2), 149–166 (2020)

10. Janvrin, D.J., Mascha, M.F., Lamboy-Ruiz, M.A.: SOX 404(b) audits: evidence from auditing the financial close process of the accounting system. *J. Inf. Syst.* **34**(3), 77–103 (2020)
11. Durgadevi, A., Shanmugavadivoo, N.: Availability capacity evaluation and reliability assessment of integrated systems using metaheuristic algorithm. *Comput. Syst. Sci. Eng.* **44**(3), 1951–1971 (2023)
12. Markiewicz, M., Koperwas, J.: Evaluation platform for ddm algorithms with the usage of non-uniform data distribution strategies. *Int. J. Inf. Technol. Syst. Approach* **15**(1), 1–23 (2022)



# Development of Rural Labor Resource Optimization Platform Based on Spring Boot Technology

Jingyu Ye<sup>1</sup>, Wei Huang<sup>2</sup>(✉), Jian Sun<sup>2</sup>, Bin Huang<sup>1</sup>, and Longfei Li<sup>1</sup>

<sup>1</sup> School of Electrical and Information Engineering, Jilin Agricultural Science and Technology University, Jilin, Jilin, China

<sup>2</sup> Economic and Management School, Jilin Agricultural Science and Technology University, Jilin, Jilin, China

junerye1314@163.com

**Abstract.** With the development of big data and cloud computing technology, the use of digital means to manage information resources has gradually become the mainstream of information processing. The outflow of rural labor force has greatly affected residents' labor operations, and temporary work is seriously lacking in normative and professional management and docking. This paper combined with the idea of information system project exploration, analyzed and classified the rural labor resource information through the current situation analysis method, experience summary method, design and exercise method, and initially built the rural labor resource platform with the help of SpringBoot technology. The combination of Internet and modern rural labor management can improve the shortcomings of traditional rural labor management and provide information technology support for rural revitalization.

**Keywords:** Rural labor Force · Resource Optimization · Temporary Management

## 1 Introduction

The proposal of the CPC Central Committee on Formulating the 14th Five-Year Plan for National Economic and Social Development and the Long-range Goals for 2035 proposes to develop the county economy, promote the integrated development of the primary, secondary and tertiary industries in rural areas, enrich rural economic forms, and expand the space for farmers to increase their income.[1] The newly issued “14th Five-Year Plan for Promoting Agricultural and Rural Modernization” also points out that the traditional driving forces supporting the increase of farmers' income are gradually weakening, and the new driving forces need to be cultivated. The construction of digital countryside should be further deepened to provide impetus for the promotion of agricultural and rural modernization. In recent years, with the tightening of resource and environmental constraints, the weak competitiveness of local agriculture, the promotion

of urban-rural integration and other factors, more and more young rural labor force outflow, resulting in a sharp increase in the number of left-behind elderly and children.[2] At the same time, there is also a part of the labor force because of various factors can not travel to work. This has resulted in a serious imbalance in the distribution of surplus Labour. In order to better regulate the two-way flow of urban and rural factors, develop the value of surplus labor and promote the integration of rural industries, it is of great significance to analyze and design the optimization platform of rural labor resources for the construction of new countryside.[3].

### **1.1 Domestic Research Status**

At present, there are still no perfect rural labor platforms in our country, and more rely on “three edge” relationship to disseminate the relevant information. The existing rural labor platform mainly collects information of rural labor force to recommend employment opportunities and promote rural plans, which has drawbacks such as long waiting time for employment, unsuitable work content and incomplete follow-up basic security. A platform that can provide short-term income is similar to the daily settlement in Shenzhen. In an area with intensive personnel flow, the rural labor platform with short-term, efficient, fast and short distance needs to be supplemented and improved in order to adapt to the changes of posts.[4].

### **1.2 Foreign Research Status**

Up to now, the research and practice of rural development in various countries, such as rural planning, agriculture-related policies, such as urban-rural integration, farmers moving to the city, capital to the countryside and so on. Whether it is large-scale farming in the United States, specialized farming in Japan or medium-sized farming in France. They have complete basic service facilities and professional equipment, technology to the countryside, capital to the countryside is the mainstream rhythm of their rural labor development, and these are suitable for their own conditions of development of scale management path. Foreign economists’ analysis of economic development and the full use of rural labor information system make rural labor management more inclined to scale management and the transfer of surplus labor. In addition, for western countries, the development of computer network technology is more mature, and the way of computer information management is also the earliest applied in their work and life, so the scope of this management is relatively more extensive in the West [5, 6].

## **2 Significance and Objectives of the Development of Rural Labor Resource Optimization Platform**

### **2.1 Development Significance**

The rural labor force platform is mainly a way to support the development of modern rural labor force industry. The traditional rural labor force management mode is still in the offline management stage, and the management efficiency of this management mode

is extremely low. With the transfer and loss of rural labor force, this offline management mode has been difficult to meet the needs of existing users. But with the arrival of the information age, the transformation of labor digitalization, can greatly facilitate the management of labor resources, solve the problem of imbalance between supply and demand of rural labor operations [7, 8].

## 2.2 Development Objectives

### (1) Realize grid management of rural labor force

Based on the village, TUEN MUN, township, according to the principle of regional management, fine management and exploration of surplus labor resources, improve the efficiency of rural labor management and information immediacy, maximize the advantages of network management information.

### (2) Simplify the platform registration process and audit process

Users can register an account with their real names through the network platform, and input, update and delete personal information and job skills information. A hierarchical and itemized administrator is set up to conduct professional and timely audit of personal certification, skill certification, recruitment authenticity and other contents, so as to better ensure the accuracy and security of platform information.

### (3) Realize efficient allocation of rural labor resources

The system preferentially recommends personnel or positions that meet professional skills. Users can also choose more suitable temporary jobs according to the spatial distance to meet the needs of both sides.

### (4) Realize the function of employment evaluation

Through the evaluation of both employers to improve the personal reputation and service level, to provide more and more real evaluation cases for both employers.

## 3 System Development Technology

### 3.1 System Architecture

The technology used in the development of the platform is mainly the currently very popular B/S structure development mode, Java object-oriented programming language and Mysql database, and finally through the overall module design, database design, functional module design, system page production to make a detailed plan [9, 10].

### 3.2 Database Selection

This project is designed to apply the data information in the distributed storage of Mysql database. Mysql is a large database system launched by Microsoft, with strong scalability and a very large scale of data storage. For the Mysql database, it is mainly to carry out the code and view the data, and has been applied to the database in many design schemes. In the course of use, you can carry out the search and composition of the basic data, so in the application of Mysql database, as long as write a small section of data can complete the relative function.

### 3.3 SpringBoot Framework

This project is designed to apply the data information in the distributed storage of Mysql database. Mysql is a large database system launched by Microsoft, with strong scalability and a very large scale of data storage. For the Mysql database, it is mainly to carry out the code and view the data, and has been applied to the database in many design schemes. In the course of use, you can carry out the search and composition of the basic data, so in the application of Mysql database, as long as write a small section of data can complete the relative function.

### 3.4 B/S Architecture

The B/S structure works by making a request from the browser and getting a response from the server. When the server processes the browser's request, it returns the results and information to the browser. And B/S structure has the advantages of low cost, convenient maintenance, strong distribution and simple development. [11] It is suitable for the development of this platform.

## 4 Design of Optimization Platform for Rural Labor Resources

### 4.1 Overall System Function Design

Through the system demand analysis, the development and implementation of the rural labor platform mainly includes three modules: the administrator module provides individual management, employer management, employee management, recruitment management, order management, service management, education and training management, system management functions; The employer module provides personal management, recruitment management, order management, service management, learning and training management; The module provides functions of personal management, order management, service management, learning and training management, my collection and so on. [8, 12] The specific functional structure is shown in Fig. 1.

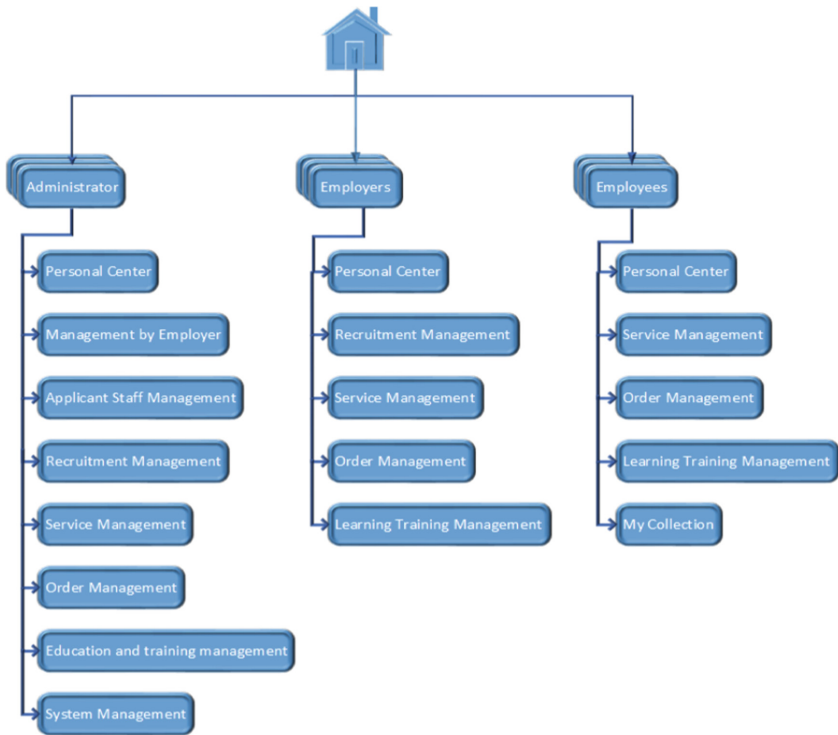


Fig. 1. System function diagram

## 4.2 Database Relation Tables

- (1) Administrator information table: mainly used to store all system information, including personnel numbers, passwords and other entities.
- (2) Employer database table: it is mainly used to store all employer information in the system, including: personnel number, employer name, gender, telephone number, recruitment information and other attributes.
- (3) Candidate information database table: it is mainly used to store all staff information in the system, including: personnel number, personnel name, email, mobile phone number, professional position and other attributes.
- (4) Employment evaluation management information database table: it is mainly used to store all information of the system, including employer evaluation, employee evaluation and other attributes.
- (5) News and information management information database table: mainly used to store all the information of the system, including: professional positions, policies and regulations, education and training, technical certification and other attributes.

The relationship between the above five data tables is shown in Fig. 2.



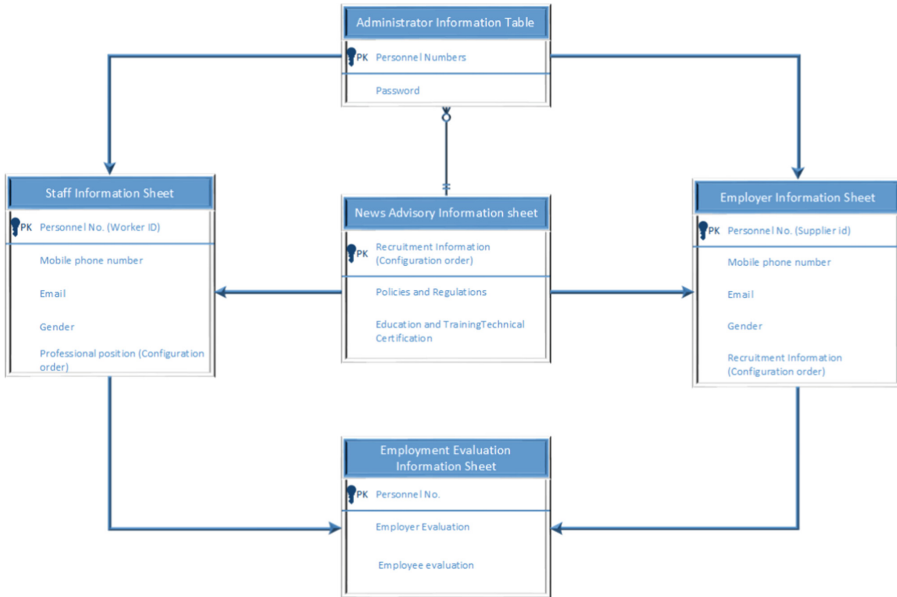


Fig. 2. Data relationship diagram

### 4.3 Setting System Functions

#### (1) System setting module

(1) Login management. When logging in to the system, enter the account and password, select the role for logging in, and click Log in. The system will compare the input information with the database information.

(2) Administrator function management. The hierarchical administrator can review, input, screen, modify, delete and view the information of employers and employees in the region, and release the news consultation, education and training, etc.

(3) Employer function. After logging in the account, employers can publish and modify their recruitment information, review orders, evaluate employees, terminate services and other operations.

(4) Staff function. After logging in the account, employees can receive orders, bookmark recruitment information, participate in skill training, and make complaints and terminate the service when there is a problem with the order.

#### (2) Information management module

(1) Recruitment information management. A tier administrator can review, filter, view, and delete recruitment information.

(2) Order receiving information management. This section describes how to configure a tier administrator who can audit, query, and delete orders.

(3) Data export. Can generate reports and print according to different screening conditions.

## 5 Conclusions

At present, China's rural labor resources management information construction is still in the advancing stage. More policy support and social support are needed to design a comprehensive management system which meets the regional characteristics and can meet the regional demands. Increasing the education, training and skill upgrading of labor resources at all levels is the basic means to improve the physical demand. Accelerating the improvement of agricultural science and technology infrastructure and the extension of entrepreneurship are the necessary conditions for increasing production and income. Consolidating the human resources management system and social service guarantee is a powerful guarantee for the overall development of rural economy. [9, 13] Making full use of Internet technology to design and implement rural labor resources management system can better serve rural revitalization, optimize the local labor resources management system. Build a platform for government management, resident employment, social services, promote rural modernization, and promote the construction of a new countryside.

**Acknowledgements.** This work was supported by Social Science Foundation of Jilin Province "Research on the Development Path and Innovation Mechanism of Rural Industrial Integration in Jilin" Province (No. 2022B49).

## References

1. Xiang, W., Zhu, K., Teo, B.M., Talib, Z.: Influence of the development of human capital and social identification of the rural laborers' mobility in rural strategy revitalization in China. *J. Eastern Eur. Cent. Asian Res. (JEECAR)* **9**, 628–637 (2022). <https://doi.org/10.15549/jeecar.v9i4.1036>
2. Ma, L., Dou, H., Shanshan, W., Shi, Z., Li, Z.: Rural development pressure and "three-stay" response: a case of Jinchang City in the Hexi Corridor, China. *J. Rural. Stud.* **91**, 34–46 (2022)
3. Shui, Y., Xu, D., Liu, Y., Liu, S.: The influence of human capital and social capital on the gendered division of labor in peasant family in Sichuan, China. *Soc. Indic. Res.* **155**(2), 505–522 (2021). <https://doi.org/10.1007/s11205-020-02598-z>
4. Weber, J.A.: (2021) Partnering with resellers in business markets. *Ind. Mark. Manage.* **30**(2), 87–99 (2001)
5. Cheng, F.: Talent recruitment management system for small and micro enterprises based on SpringBoot framework. *Adv. Educ. Technol. Psychol.* **5**(2), 99–105 (2021)
6. Teng, Z., Fan, S.: Research on the strategy of computer network security in the big data age. *Wirel. Internet Technol.* (2019)
7. Digilina, O.B., Teslenko, I.B.: Transformation of the labor market in the context of digitalization. *RSUH/RGGU Bull. Ser. Econ. Manage. Law* (2020)
8. Hongju, L., Yinghua, H.: Perfect service and guidance to speed up the rural surplus labor transfer. *J. Jilin Agric.* **17**, 49 (2021). <https://doi.org/10.14025/j.carolcarrollnkijlly.2021.17.003>
9. Nechaev, V.I., Arzhantsev, S.A., Mikhailushkin, P.V., Khoruzhy, L.I., Bondarenko, T.G.: Eurasian agricultural technological platform for technical and technological modernization of agriculture in the EAEU member states: thematic research (2021)

10. Jiangang, H.: Shenmu, the plight of the rural labor resources development and countermeasures. *Rural Sci. Technol.* **12**(08), 26–27 (2021). <https://doi.org/10.19345/j.carolcarrollnki.1674-7909,2021.08.019>
11. Qi, L.: Discussion on the existing problems and countermeasures of rural labor transfer. *Read. World* **09**, 243 (2020)
12. Li, J.: Research on employment survey and employment strategy of rural labor force. *Inner Mongolia Sci. Technol. Econ.* **20**, 10+12 (2021). (in Chinese)
13. Wu, Y., Wang, Z.: Secondary vocational education: an important platform for rural labor transfer. *Southland Today (Theoret. Innov. Ed.)* **11**, 77–78+89 (2020)



# Research on Performance Enhancement of Parallel Computing in Intelligent Proofreading System

Zhishuai Guo<sup>(✉)</sup> and Fucheng Zhang

Neusoft Group Co., Ltd, Shenyang, Liaoning, China  
guozs@126.com

**Abstract.** This paper describes the status quo of intelligent checking technology and the efficiency problems existing in the current checking system, and analyzes the reasons why the processing complexity of the checking system determines the time consuming of serial execution. Through the introduction of parallel computing technology, the paper finds out the tasks suitable for multi-process and multi-thread in the checking system. Through the reference of openMPI and Pthread, through multi-scene and multi-parameter experiments, the direction and method of parallel computing performance improvement and parameter value measures are obtained, and the parallel computing optimization strategy of the existing intelligent checking and checking system is proposed.

**Keywords:** Intelligent Proofreading · Parallel computing · FPMMax

## 1 Introduction

With the continuous development of artificial intelligence technology, especially after the emergence of deep learning technology, the application fields of artificial intelligence continue to expand, especially in the media industry is more and more widely used. The proofreading of news articles is an important link in the process of news production and release. In the context of artificial intelligence, the traditional manual proofreading has been unable to meet the requirements of information transmission speed in the new media era, and the application of artificial intelligence-related technology to the proofreading of news manuscripts has become the mainstream proofreading method at present. Therefore, in order to enhance content production, improve production verification efficiency, and enrich news content, it is of practical significance to build an intelligent proofreading system for media and we-media industry.

## 2 Development of Intelligent Proofreading System

The content intelligent proofreading of news articles can be divided into three levels: vocabulary level, whole sentence level and opinion level. Intelligent proofreading system is also developed according to these three levels.

The first generation proofreading system is mainly based on the lexical level proofreading system. Based on the long-term accumulation of a large number of wrong word database, the system matches the text content of the manuscript word by word and word by word. In the process of proofreading, if it is found to match the content of the wrong word database, the system will identify it as a word error and give a hint. The system can assist the checker to identify some common vocabulary errors and improve the efficiency and accuracy of the check. However, due to the infinite number of collocation combinations between words, the first-generation proofreading system could not detect the whole sentence level errors for collocation errors in context [1].

The second generation proofreading system is mainly based on the whole sentence level checking system. By learning a lot of correct corpus, the system enables the computer system to analyze and summarize the idioms and patterns of the language, so that the system can form a certain understanding and judgment ability of the sentences, so as to realize the analysis and judgment of the words and words in the dimension of a sentence, identify the abnormal and unreasonable content, and achieve the purpose of checking and proofreading. In view of the shortcomings and weaknesses of the first generation system, the second generation of checking system adopts artificial intelligence technology to realize the whole sentence level text check, and can realize the whole sentence level verification [2].

The third generation proofreading system is mainly based on viewpoint level checking system. The system is a humanoid system, which can imitate the recognition way of human, and carry out comprehensive analysis and understanding of manuscript content. Based on the point of view and keynote of the full text, it can judge whether each sentence and each word is reasonable, and whether there is any contradiction or illogical place, so as to realize the point of view level verification. At the same time, the third-generation system will also be able to check manuscripts based on pre-set ideas and themes to identify content that runs counter to mainstream values or common sense. It can be said that the third generation of intelligent proofreading system has been able to replace manual proofreading work to a large extent [3].

### **3 Analysis of the Current Intelligent Checking and Rectifying System**

#### **3.1 Current Status of Intelligent Checking and Rectifying System**

The current intelligent proofreading system completes the training and construction of the checking model by combining manual and machine learning, which can realize the checking and checking of specific words and data docking, the data checking and checking of Chinese fixed terms, the checking and checking of punctuation marks in Chinese language, and the checking and checking of common text errors. Through CA/LDA, SIFT/HO, GMM, HMM, SVM, decision tree, neural network, deep learning, reinforcement learning multi-modal feature extraction methods to identify news events, judge the truth and falsity of news. The intelligent checking and checking system also provides many services [4].

(1) Interface service

By connecting the API interface service, the interface can return the content and location of suspected problems found in the manuscript, which is convenient for the editing system to display in the corresponding position of the manuscript.

(2) Statistical Service

Intelligent proofreading products can count the number of proofreading or The Times of proofreading by checking the text, which is conducive to the business identification ability of products and lays the foundation for the charging mode of cloud services.

(3) Account management capabilities of cloud services

Intelligent verification products can be SAAS through cloud services. In this case, independent multi-account service capabilities and statistical capabilities are required.

(4) Technical standards and user experience

The UI and UE experience design of software should conform to the form of current mainstream Internet software. Technically, the distributed micro-service architecture of Spring Cloud is adopted, and python and related NPL plug-ins are used to develop and iterate the artificial intelligence kernel [5].

### 3.2 Problems Existing in the Current Intelligent Checking and Checking System

(1) Model-based single-threaded execution affects performance

The inspection model trained on a network of neurons analysed from a billion words of digital data, 13,000 Chinese idioms, more than 7000 sensitive word, careful use of Words, forbidden words, More than 15,000 punctuation checks, more than 20,000 citation words, homophone words, sound words, shape words, homophone words synonyms is huge. Each model adopts a single-thread processing mode, so that in the intelligent check of a large article, as long as there is a thread is not finished, it needs to wait for its return result, resulting in the overall performance is weak [6].

(2) When multiple users use cloud resources, cloud resources are improperly allocated

In SAAS deployment, the intelligent checking and verification system cannot allocate cloud resources according to the usage of different users. The cost of cloud resources using GPU and common ECS servers is uncontrollable and cannot be adjusted flexibly.

(3) The interface has downward compatibility problems. Procedure

The interface of intelligent checking needs to be compatible with the version upgrade. Due to the high coupling degree in the original program architecture, the reconstruction cost increases.

(4) The program design of GPU is not applicable

In the current checking system, GPU is used in parallel computation of MIMD (multi-instruction stream and multi-data stream), but in fact, GPU programming tends to directly parallel computation of a single computing model. The unsuitable programming occupies the computing resources of the GPU, which makes the whole situation unreasonable.

## 4 Optimization of Intelligent Checking and Checking System by Parallel Computing

The parallel computing is used to optimize the intelligent checking and checking system, which usually adopts multi-process and multi-thread processing. Combined with the multi-process processing experiment, OpenMPI multi-process is used to carry out the experiment in the supercomputer center. After the semantic model, political language, fallible word, sensitive word, disable corpora that needs to be checked is initialized with the data square, the column and row data of the check task is formed, and the column and row data that needs to be processed is evenly distributed in each process [7]. MPI was broadcast for the column and row data, so that each process received the broadcast task pushed from the data array, and the experimental data was shown in Table 1.

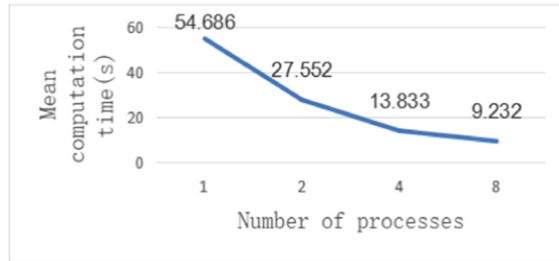
**Table 1.** Number of processes and runtime relationship table

Number of experiment	Times (s)			
	Number of processes			
	1	2	4	8
1	54.627	27.588	13.816	9.211
2	54.647	27.449	13.805	9.214
3	54.821	27.551	13.83	9.194
4	54.731	27.621	13.82	9.242
5	54.693	27.584	13.847	9.249
6	54.694	27.575	13.802	9.247
7	54.577	27.525	13.896	9.252
8	54.594	27.541	13.803	9.233
9	54.766	27.585	13.861	9.272
10	54.709	27.499	13.848	9.203
Mean value	54.686	27.552	13.833	9.232

It can be seen from the data in Table 1 that the NLP, NLP model, political language, fallible word, sensitive word, disable corpora that inspection effect of 8000 words manuscript with 1, 2, 4, and 8 processes is shown in Fig. 1. As the number of NLP model, processes increases, the checking time decreases from 54.686 s to 9.232 s, indicating that parallel computing improves the checking speed.

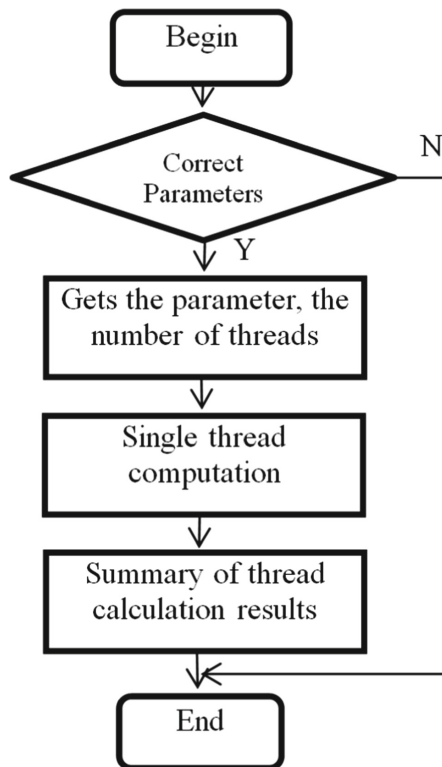
With the increase of the number of processes, the running time is gradually shortened and the acceleration ratio is gradually increased. But the running time slope is decreasing. One possibility is that there are too many processes, and a lot of up time is spent switching between processes and communicating with each other, which slows down the speedup rate.

In addition to the parallel computation of multiple processes, Pthread multithreading can be used in each process to deal with the relative same type of check tasks [8]. For



**Fig. 1.** Calculate the graph of the change of time with the number of processes

example, the examination objects of political language are divided into task packages, and then the contents of the task packages are checked with the political language corpus. The experiment is repeatedly run with different numbers of threads, and the average time of multi-threading processing of political language is obtained. The flow chart of specific implementation is shown in Fig. 2.



**Fig. 2.** Flow chart of multithreaded parallel computing



According to the above process, multithreading parallel computing experiment was carried out, and the experimental data results were shown in Table 2.

**Table 2.** Relationship between number of threads and time

Time (S)	1.2	0.7	0.45	0.4	0.3	0.25	0.31	0.33	0.4
Number of threads	1	2	3	4	5	10	15	20	30

As you can see from the table above, adding more threads increases the computation speed when the number of threads is small. However, when the number of threads is too large, the scheduling of threads also needs to consume time, thus affecting the computing speed.

From the conclusion of the above process and thread parallel experiment, the application of intelligent checking and checking system should be optimized from the following aspects:

(1) Serial threads are parallelized directly

The parallel mechanism of serial algorithm is used to directly transform serial algorithm into parallel algorithm. The method of direct parallelization by serial algorithm is one of the most common methods of parallel algorithm design, but not all serial algorithms can be directly parallelized. In intelligent checking and checking system, for idioms, common words, polyphonics, fallible words, political words and other dictionary models can be directly parallelized. This can greatly enhance the performance of model verification.

(2) Parallelization of sorted groups of elements checked by the same class

The semantic analysis elements of Chinese language are massive, which can be grouped into different processors for multi-process and multi-threaded parallel computation. In the checking and checking system similar to the semantic model of Chinese language, some algorithms need to be used to group, sort and parallelize the model elements. The basic algorithms include multi-candidate generation algorithm, pattern growth algorithm and vertical format algorithm, and parallelization scheme using association rule mining algorithm. Using the FPMMax algorithm for recursive processing, and then using the width first processing method to model mining and comparison of FP-TREE data to be checked, this parallel nonlinear processing mechanism is an effective way to improve the speed of the natural language semantic analysis check.

(3) Optimization of GPU programming

With the continuous enhancement of the programmability of the GPU, the cooperation between the CPU and GPU becomes a strong computing body [9]. The CPU is the core of serial processing, while the GPU is the core of strong parallel computing. The serial part of the program runs on the CPU, while the parallel computing part runs on the GPU. Make the checker run much faster than a system using a multi-core CPU [10].

## 5 Conclusions

In summary, the application of parallel computing can improve the performance of intelligent checking and checking system in many aspects and extend the life cycle of software, which indirectly drives the reconstruction and business innovation of intelligent checking and checking products. For the processing of intelligent and parallel computing, a lot of excellent algorithms and program architecture need to be reformed.

## References

1. Min, J.: Research on the application of computer intelligent proofreading system in college English teaching. *J. Phys. Conf. Ser.* **1915**(3) (2021)
2. Ren, X.: Research on the design of the computer intelligent proofreading system for English translation. *Basic Clin. Pharmacol. Toxicol.* **126** (2020)
3. Zheng, H.: Research on computer intelligent proofreading system of improved English phrase translation model. *J. Phys. Conf. Ser.* **1871**(1) (2021)
4. Zou, Y.: Application of computer intelligent proofreading system in English phrase translation. *J. Phys. Conf. Ser.* **1881**(4) (2021)
5. Zhao, J.: Development of computer intelligent proofreading system from the perspective of English translation application, 996–1001 (2022)
6. Kong, M., Yu, F.: Design and application of computer conference+Thai corpus-oriented word segmentation consistency checking system. *J. Phys. Conf. Ser.* **1915**(3) (2021)
7. Song, R., Song, X., Zhang, Y., Ma, Y.: Experiment in parallel computing for the Julia programming language. In: *Conference Proceeding of 2020 3rd International Conference on Algorithms, Computing and Artificial Intelligence (ACAI 2020)*, pp. 215–221 (2020). <https://doi.org/10.26914/c.cnkihy.2020.052176>
8. Gokalp Yavuz, F., Schloerke, B.: Parallel computing in linear mixed models. *Comput. Stat.* **35**(3), 1273–1289 (2020). <https://doi.org/10.1007/s00180-019-00950-7>
9. Skorych, V., Dosta, M.: Parallel CPU–GPU computing technique for discrete element method. *Concurrency Comput. Pract. Exp.* **34**(11) (2022)
10. Auradkar, P., Gagan, G.R., Deva, S., et al.: Optimized closest pair computation with CPU-GPU *Comb. Model.*, 743–755 (2022)



# Ethical Reflections on Artificial Intelligence: Practical Dilemmas and Mitigation Strategies

Hao Cheng<sup>(✉)</sup>

School of Marxism, Sichuan University, Chengdu, Sichuan, China  
617187051@qq.com

**Abstract.** Artificial intelligence is not a pure labor tool, but an innovative force that drives the technological revolution and promotes industrial innovation. A series of ethical dilemmas such as the responsibility attribution dilemma, fairness and justice dilemma, privacy and security dilemma, and threat to humanity dilemma have arisen accordingly. The reason for all this is due to the development of artificial intelligence technology itself and the transformation of the production mode in which human beings become data producers and the whole society is reshaped. In the face of dilemma relief strategies, we should not only adhere to the basic law between productive forces and production relations and the main principle of people-oriented to benefit mankind, but also ensure that the development principle of human-machine integration which guarantees that humans and machines evolve together, while following ethical principles such as the trend of sharing and openness contained in artificial intelligence itself.

**Keywords:** Artificial Intelligence · Ethical Reflections · Practical Dilemmas

## 1 Introduction

The fundamental goal of artificial intelligence is to understand the essence of human intelligence and simulate it, or to create similar intelligent machines [1]. The rapid development of artificial intelligence has brought profound changes to human production and life, and has also brought many ethical dilemmas to human beings. How to avoid the risks brought by artificial intelligence has become a necessary research [2]. It is an objective requirement of society and reality to systematically investigate the ethical dilemmas of artificial intelligence, analyze the specific root causes of the dilemma and make corresponding strategies to alleviate them.

Avoiding the risk caused by limitations of AI technologies has become a grand challenge.

## 2 The Ethical Dilemma of Artificial Intelligence

The rapid development and popularization of artificial intelligence technology has inevitably brought many ethical dilemmas to society. On the one hand, these dilemmas warn us to pay attention to the relationship between human beings and artificial

intelligence. The attitudes and values that human beings hold towards artificial intelligence will directly affect the development of human society itself. On the other hand, the dilemmas also require people to carry out scientific and reasonable use and control of artificial intelligence technology and always adhere to the basic ethical principles of science for the benefit of mankind.

## **2.1 The Ethical Dilemma of Attribution of Responsibility**

The large-scale application of artificial intelligence has brought a large number of issues related to the subject of responsibility, represented by self-driving cars. Artificial intelligence is an important part of autonomous vehicle [3]. In the event of a traffic accident, who will hold the responsibility? Is it the user of the car, or the car manufacturer, or even the car itself? The self-driving car itself is only controlled by intelligent programs. It can have certain human-like intelligence, but it cannot become an independent subject with corresponding responsibilities. Its logical behavior is determined by the algorithm of its own program, and limiting the discussion of accident liability to the users of the vehicle and the manufacturers is a more reasonable consideration. For one thing, it is irrational to blame the car users for the accident, because the automatism of automatic driving lies in the fact that the cars are driven without the involvement of the users that is, the driving behaviors of the cars don't depend on the users' will, so the users are not responsible for the accident; For another, if the cars are delivered to the users without obvious defects in their designs, then the subsequent accident liability should have nothing to do with the designers, then the design manufacturers are not responsible for the accidents of the self-driving cars.

## **2.2 Ethical Dilemmas of Fairness and Justice**

The "intelligence" of artificial intelligence mainly relies on algorithms in the program, and the fairness and justice problems brought about by algorithm discrimination are increasing. The development process of artificial intelligence is also the development process of algorithms from simple to complex, and the code generated inside artificial intelligence is increasingly difficult to be explained by the decisions generated by machine learning, and what leads to people's inability to accurately predict the results of the algorithm is the "black box" problem inside the algorithm, which is an unfair behavior produced by the algorithm in the specific application process.

## **2.3 Ethical Dilemmas to Privacy Security**

The abuse of data collection and accurate prediction technology of artificial intelligence technology has led to more and more privacy security issues. There is a remarkable increase in the number of Web 2.0 tools like online social media and e-commerce websites where users freely express their ideas and thoughts [4]. Through the collection and calculation of personal information, big data technology infers our preferences, habits and needs. On the one hand, face recognition, camera surveillance, biometric collection and other technologies have been integrated into our lives, and almost everyone

is accustomed to this. However, whether the data collected through these methods is reasonable and whether these data will not be used to calculate users' privacy information and predict users' privacy behavior are difficult to guarantee. On the other hand, if these technologies are abused by criminals to steal state secrets, carry out human flesh search and disseminate privacy letter maliciously, they will also cause great damage to individuals, collectives and national interests.

### **3 The Root Cause of the Ethical Dilemma of Artificial Intelligence**

The rapid progress of artificial intelligence technology has brought great, rapid and all-round changes to society and human beings. On the one hand, it provides convenient technical basic support for the entire human society, resulting in the comprehensive intelligence of social life. On the other hand, it also brings corresponding ethical dilemmas, making the relationship between humans and artificial intelligence increasingly distorted and ambiguous.

#### **3.1 Objective Root Causes**

AI has become more capable in planning and decision-making [5]. The rapid development of artificial intelligence technology itself for human beings has created corresponding ethical dilemmas. Artificial intelligence needs to be gradually developed through the basic field to the application field. For example, Artificial intelligence-based technologies are gradually being applied to psychiatric research and practice [6]. Virtual health assistance is a successful example of the application of artificial intelligence in the field of medical and health care, such as Babylon Health in the United Kingdom and Kangfuzi in China. AI even gave birth to a computational medicine [7]. Artificial intelligence has penetrated into all aspects of life, extracting people's private data and directly "controlling" human beings through algorithmic decision-making. We enjoy the convenience and superiority brought by scientific and technological progress, but the price we pay is surrounded by various artificial intelligence devices, such as smart phones, smart bracelets, various monitoring devices, etc.

#### **3.2 Subjective Root Causes**

In the age of artificial intelligence, everyone has become a data producer. Different from the traditional labor object, whose scope is always narrow, either the materials produced by nature, or the products processed by some people, labor object in the era of artificial intelligence, is the data. All people who access the network will produce data, so the issue of data privacy is getting more and more attention. More and more personal privacy data is obtained in a lower, more private and simpler way and cost, and the channels for obtaining private data have even become a complete gray industry chain, which greatly increases the cost of maintaining data security for enterprises and reduces the public's trust in artificial intelligence and its industry. Data information is permanently retained online, even if it is deleted artificially, it will definitely leave clues on the network. Therefore, some people call the era the naked era is not unreasonable.

### **3.3 Social Background**

Production mode of the whole society is reshaped. The continuous informatization, intelligence, unmanned and automation of social economy, cultural, political and other activities has brought about a great increase in labor productivity, resulting in a significant increase in productivity and a corresponding expansion of labor scope. AI can even be used to predict the distribution of energy [8]. Data information has become one of the new and most important resources, so the corresponding data enterprises and intelligent industries have also emerged. The repetitive, monotonous and simple physical labor in the traditional factory has been replaced by artificial intelligence robots, and more and more mental work, even including computing and medical work, is constantly being replaced by intelligent machines.

## **4 Strategies for Solving Artificial Intelligence Dilemmas**

With the development of artificial intelligence and the change of human interaction, more and more people are beginning to express concerns about the technological development potential of artificial intelligence. Artificial intelligence has adapted to the social and natural environment [9]. Elon Musk asserts that “artificial intelligence may be the demon we are summoning”. Artificial intelligence is only a technical means, and the changes in productivity and production relations brought about by artificial intelligence are unified in people themselves. The essential attribute of artificial intelligence is the materialization and objectification of human labor and the materialization and objectification of the intelligent activity viewing man as a subject.

### **4.1 Follow the Trend of Sharing and Openness Contained in Artificial Intelligence Itself**

The deep learning of artificial intelligence requires artificial intelligence to be continuously used by various people in order to become smarter and smarter quickly, so artificial intelligence itself requires continuous use once generated. The development and evolution of artificial intelligence can only be realized by sharing. The foundation of artificial intelligence, big data itself is open. Data itself is difficult to be occupied alone due to its own form, so openness itself is the requirement of big data. And big data itself can be more accurate and intelligent only if it is open. Developing countries should keep up with the development trend of artificial intelligence, and developed countries should shoulder international responsibilities. The whole world should unite together to strengthen the global quality of artificial intelligence.

### **4.2 The Principle of People-Oriented Should Be Adhered to for the Benefit of Mankind**

Service should be a basic feature of artificial intelligence, and the purpose of service must be people-oriented. Artificial intelligence is to help people become healthier and happier, for example, to help people treat cancer [10]. Artificial intelligence technology,

like traditional technology, should always implement the basic principle of putting people first. Artificial intelligence cannot be developed at the cost of human beings' eternal basic right to privacy, but should provide better, safer and more convenient protection for the basic right to human privacy. No matter how amazing the artificial intelligence world will be in the future, without people, even the most magical artificial intelligence will become meaningless. Marx believed that people not only carries out the production of material materials, but also the production of social relations. "The continuous enrichment of social relations in the process of practice pushes subjective relations into public relations." Artificial intelligence machines can never replace human subjectivity.

### **4.3 The Development Principle of Human-Machine Integration Ensures that Humans and Machines Evolve Together**

The relationship between humans and AI machines is symbiotic rather than binary opposite. The combination between human and machine means that artificial intelligence and human beings play their respective strengths and organically combine to jointly realize the existence value of human and artificial intelligence. The relationship between computers and people is getting closer and closer, and computers can replace the human brain for more and more jobs. The digital surgery assisted by artificial intelligence has been able to realize digital decision support [11]. Intelligent computers are no longer simple machine tools, but play the role of the mediator human relationships. We don't have to think of artificial intelligence as a beast, the relationship between human intelligence (HI) and artificial intelligence (AI) is symbiotic rather than antagonistic. The birth of brain networking in the future means that the development of human-machine integration has entered a new mode of human-machine integration.

## **5 Summary**

To sum up, artificial intelligence technology has not only brought a lot of convenience, but also a lot of ethical dilemmas. The essence of artificial intelligence is a human creation, which is only the further development of technology from the extension of the limbs to the extension of the brain. Therefore, artificial intelligence can not replace humans. Under the guidance of the principle of people-oriented, from the perspective of technology and people, restricting and stipulating artificial intelligence, are not only the inherent development requirements of artificial intelligence technology itself, but also conform to the inevitable law of the development of artificial society.

## **References**

1. Poo, M.: Towards brain-inspired artificial intelligence. *Natl. Sci. Rev.* **5**(06), 785 (2018)
2. Xue, J., Hu, B., Li, L., Zhang, J.: Human-machine augmented intelligence: research and applications. *Front. Inf. Technol. Electron. Eng.* **23**(08), 1139–1143 (2022)
3. Ma, Y., Wang, Z., Yang, H., Yang, L.: Artificial intelligence applications in the development of autonomous vehicles: a survey. *IEEE/CAA J. Automatica Sinica* **7**(02), 315–329 (2020)

4. Habimana, O., Li, Y., Li, R., Gu, X., Yu, G.: Sentiment analysis using deep learning approaches: an overview. *Sci. China (Inf. Sci.)* **63**(01), 21–56 (2020)
5. Hepworth, A.J., Baxter, D.P., Hussein, A., Yaxley, K.J., Debie, E., Abbass, H.A.: Human-swarm-teaming transparency and trust architecture. *IEEE/CAA J. Automatica Sinica* **8**(07), 1281–1295 (2021)
6. Cao, X., Liu, X.: Artificial intelligence-assisted psychosis risk screening in adolescents: practices and challenges. *World J. Psychiatry* **12**(10), 1287–1297 (2022)
7. Lyu, L., Cui, H., Shao, M., Fu, Y., Zhao, R., Chen, Q.: Computational medicine: past, present and future. *Chin. J. Integr. Med.* **28**(05), 453–462 (2022)
8. Mollaiy-Berneti, S.: Developing energy forecasting model using hybrid artificial intelligence method. *J. Central South Univ.* **22**(08), 3026–3032 (2015)
9. Qiu, J.: Research and development of artificial intelligence in China. *Natl. Sci. Rev.* **3**(04), 538–541 (2016)
10. Tonini, V., Zanni, M.: Early diagnosis of pancreatic cancer: what strategies to avoid a foretold catastrophe. *World J. Gastroenterol.* **28**(31), 4235–4248 (2022)
11. Hardy, N.P., Cahill, R.A.: Digital surgery for gastroenterological diseases. *World J. Gastroenterol.* **27**(42), 7240–7246 (2021)





# Analysis of the Direction of Library Development in the Future Based on Knowledge Graph Analysis

Jie Dong<sup>1</sup> and Lingxiang Luo<sup>2</sup>(✉)

<sup>1</sup> School of Electrical and Control Engineering, Shenyang Jianzhu University, Shenyang, Liaoning, China

<sup>2</sup> School of Computer Science and Engineering, Shenyang Jianzhu University, Shenyang, Liaoning, China

2413048310@qq.com

**Abstract.** Visual analysis of library reader service research and linear regression analysis are conducted to point out the closeness and changing rules of different influencing factors and the overall service quality of the library, providing suggestions for the optimization direction of library service problems in the future. In the CNKI database, the data in the field of “library reader service” is used as the initial data, and the knowledge map is used to analyze and study the research hotspots in this field. The questionnaire is designed based on the results, and the regression analysis is conducted on the questionnaire data. The research hotspots in the field of library reader service mainly focus on four aspects: reader service, reading service, environmental service and activity organization. By combining these four aspects to provide multi-level services, libraries can adapt to the rapid development of libraries in the new century.

**Keywords:** Reader Service · Public Libraries · Knowledge Graph · Regression Analysis

## 1 Introduction

With the development of the times and the progress of the society, the library is also constantly developing and changing [1]. On the other hand, with the rapid development of social and material conditions, the spiritual and cultural demands of the people have become more and more prominent, while the most groups, due to their special social experience, cultural background, physical conditions and reading habits, more often choose public libraries as the main venue and way to carry out their own spiritual and cultural life [2]. In this context, It is of great significance to study how to diversify and optimize future library services [3].

From the existing literature, scholars mostly understand and put forward suggestions for improvement from the perspective of their own professional fields in the study of library reader service, but the macro development and coordination of the problems are insufficient.

On the basis of previous work, we use data visualization analysis methods to reveal research hotspots and explore improvement methods through combing research literature in related fields, so as to provide reference for improving the service quality of library readers [4].

## 2 Data Sources and Research Methods

### 2.1 Data Sources

A fuzzy search was conducted on China Knowledge Network with the theme of “Library reader services” for the past twenty-five years (1998–2022). A total of 313 valid documents were obtained. The selected documents are representative and can be used in this study [5].

### 2.2 Research Methodology

Visualisation techniques were used to analyse the research hotspots, regression analysis is conducted on the questionnaire data to understand the current situation of reader service in Hefei Library, and then countermeasures for service improvement are discussed [6].

## 3 Knowledge Graph Analysis

Visual analysis can highlight the relevance between data, and conduct visual analysis on relevant documents of “research on public library reader service” (see Fig. 1).



**Fig. 1.** Keyword co-occurrence mapping

The co-occurrence network diagram of high frequency keywords shows that the keywords “public libraries”, “services” and “elderly readers” are distributed in the centre of the diagram as a whole. In the structure of the co-occurrence network diagram, the keywords “public libraries”, “services” and “elderly readers” are surrounded by the keywords “readers services”, “reading services”, “barrier-free design”, “ageing-friendly”,

“cultural ageing” “This indicates that the above keywords are closely related to the keywords “public libraries”, “elderly readers” and “services”. This is a hot issue under the research topic of “library services for elderly readers”.

The chart shows that the research hotspots can be divided into four main hotspots: library patron services for older readers, library reading services, library environmental services and library activity organisation services.

## 4 Questionnaire Survey on Senior Reader Services in Public Libraries

From the findings of the visual analysis, the current research on public library services for elderly readers mainly focuses on four aspects: reader services, reading services, environmental construction and activity organisation, while the Hefei City Library, as a municipal public library, is also an important window for providing services for the spiritual and cultural life of elderly readers in Hefei. Therefore, this paper designs a survey questionnaire based on its results to investigate the Hefei City Library.

### 4.1 Confidence Analysis

Questionnaire reliability is an indicator of the degree of truthfulness of the respondents. The questionnaire was analysed for reliability (see Table 1) and the Cronbach coefficient is currently the most commonly used coefficient that can represent the level of reliability. Among them, the reliability coefficient of the total scale is 0.889, which meets the survey requirements.

**Table 1.** Results of the reliability test

Reliability analysis		
Clone Bach Alpha	Clone Bach Alpha terms based on normalized	terms
.889	.894	21

### 4.2 Validity Analysis

Validity is a method used to analyse the degree of validity of a research object. It shows the extent to which the findings match what was set as expected and the higher the validity, the more the findings reflect the nature of the research object. The questionnaire was analysed for validity (see Table 2), where the KMO value of 0.813 was met.

**Table 2.** Validity test results

KMO and Bartlett's test		
KMO Sampling suitability number		.813
Bartlett's test of sphericity	Approximate chi-square	204
	Degree of freedom	120
	Significance	.000

### 4.3 Regression Analysis

Regression analysis is used to examine the pattern of quantitative changes between variables in order to analyse the degree of influence between them. In this paper, regression analysis is used to examine the linear relationship between reader service, reading service, environmental service and organisational service factors and overall service quality (see Table 3). Where the dependent variable was chosen to be overall service quality and the independent variables were those in the table above, the regression equation was set to be

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \mu \quad (1)$$

As seen from the coefficient values in Table 3, the regression coefficients of reader service, reading service, environmental service and activity organization service passed the significance test ( $\text{Sig} < 0.05$ ), with the regression coefficients of 0.173 for reader service, 0.187 for reading service, 0.166 for environmental service and 0.227 for activity organization service, indicating that These factors all have a significant positive influence on the overall service quality, resulting in the regression equations of the influencing factors of reader service, reading service, environmental service and activity organisation service with the overall service quality as

$$Y = 0.637 + 0.173X_1 + 0.187X_2 + 0.166X_3 + 0.227X_4 \quad (2)$$

From the above regression equation, all else being equal, for every 1 unit increase in the quality rating of reader services, the overall service quality rating improves by 0.173 units; for every 1 unit increase in the quality rating of reading services, the overall service quality rating improves by 0.187 units; for every 1 unit increase in the quality rating of environmental services, the overall service quality rating improves by 0.166 units; and for every 1 unit increase in the quality rating of organizational For every 1 unit increase in service quality rating, the overall service quality rating increases by 0.227 units.

As shown by the coefficients of the regression equation, for the elderly reader group, organisational services have the greatest impact and absolute dominance on the overall service quality of the library, followed by reading services, reader services and environmental services. Therefore, with limited resources, libraries should focus on improving their own organisational capacity for elderly readers' activities in order to maximise the overall service quality of the library [7].

**Table 3.** Table of regression equation coefficients

Model	Coefficient <sup>a</sup>				Covariance statistics		
	Unstandardised coefficients	Standardised coefficients	Beta	t	Significance	Tolerance	VIF
(constant)	.637	.256		1.899	.033		
Reader services	.173	.103	.132	2.645	.047	.538	1.859
Reading services	.187	.078	.139	2.444	.003	.429	2.233
Environmental services	.166	.093	.113	3.933	.042	.499	2.003
Organizational services	.227	.064	.299	1.547	.036	.456	2.191

a. Dependent variable: Overall service quality

## 5 Conclusions

From the findings of the visualisation analysis of public library services for older readers, the four dimensions of patron services, reading services, environmental services and organisational services were selected to effectively reflect the study of ‘public library services for older readers’. The results of the regression analysis further indicate that the improvement of the variables in this dimension has a more significant impact on the quality of services for elderly patrons in public libraries, can effectively promote the diversified development of library services in the future [8].

- (1) Organizational services have the greatest impact on the quality of elderly patron services in libraries

The coefficients of the regression analysis show that organisational services have the greatest impact on the elderly patron services in libraries. Improving organizational services can enrich the spiritual and cultural life of older readers, Promote the harmony of the future society and the improvement of the spiritual civilization environment [9].

- (2) Patron services have a significant positive impact on the quality of elderly patron services in libraries

The coefficients of the regression analysis show that patron services have a significant positive impact on the quality of library services for elderly patrons. Strengthening and improving patron services can effectively improve the quality of elderly patron services in public libraries [10].

- (3) Reading services can effectively improve the quality of services for elderly readers

The regression analysis coefficients show that reading services have a positive impact on the quality of library services for elderly readers. Targeted book resource recommendation can enrich the aging book content, resource form and other reading services in a diversified way [11].

- (4) There is a clear demand for environmental services for elderly readers

From the coefficients of the regression analysis, the absolute value of the coefficient for environmental services is also higher. The library can make the elderly readers’ use of public library resources more convenient by increasing investment in the aging adaptation of environmental facilities, So as to improve the service quality of future elderly readers at multiple levels.

## 6 Conclusions

This paper analyzes the hot research topics in the field of library reader service through visualization method, and quantifies the influencing factors of library reader service quality. It is hoped that the research of this paper can provide reference for the diversified and multi-level development of libraries in the future, and add a touch of color to the spiritual life of library readers.

## References

1. Twum, K.K., Yalley, A.A., Agyapong, K.Q., et al.: The influence of Public University library service quality and library brand image on user loyalty. *Int. Rev. Publ. Nonprofit Mark.* **18**, 207–227 (2021)
2. Choshaly, S.H., Mirabolghasemi, M.: Using SEM-PLS to assess users satisfaction of library service quality: evidence from Malaysia. *Libr. Manag.* **40**(3/4), 240–250 (2019)
3. Afthanorhan, A., Foziah, H., Rusli, R., et al.: Modeling reflective constructs in generalized structure component analysis: an application to service quality and customer satisfaction in UniSZA library. *Int. J. Innov. Creativity Change* **7**(10), 33–41 (2019)
4. Keisling, B.L., Sproles, C.: Reviewing and reforming library service points: lessons in review and planning services, building layout, and organisational culture. *Libr. Manag.* **38**(8/9), 426–436 (2017)
5. Kumar, A., Mahajan, P.: Evaluating library service quality of University of Kashmir: a LibQUAL+ survey. *Perform. Meas. Metrics* **20**(1), 60–71 (2019)
6. Abumandour, E.S.T.: Public libraries' role in supporting e-learning and spreading lifelong education: a case study. *J. Res. Innov. Teach. Learn.* **14**(2), 178–217 (2020)
7. Khasseh, A.A., Yamchi, S.R., Azimi, H., et al.: Library services to the disabled in the public libraries of Iran (a case of East Azerbaijan Province). *Libr. Philos. Pract. (e-journal)* **3551** (2020)
8. Omid, M., Taheri, S.M., Sarrafzadeh, M.: Investigating library services for the blind and visually impaired in Iranian public libraries and suggesting some solutions to improve them. *Libr. Inf. Sci. Res.* **11**(1), 84–108 (2021)
9. Dahan, S.M., Taib, M.Y., Zainudin, N.M., et al.: Surveying users' perception of academic library services quality: a case study in Universiti Malaysia Pahang (UMP) library. *J. Acad. Librarianship* **42**(1), 38–43 (2016)
10. Halim, S., Felecia, I., et al.: Digital natives: its characteristics and challenge to the library service quality. In: *Proceedings of Second International Conference on Electrical Systems, Technology and Information 2015 (ICESTI 2015)*, pp. 487–494 (2015)
11. Serholt, S., Eriksson, E., Dalsgaard, P., et al.: Opportunities and challenges for technology development and adoption in public libraries. In: *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*, pp. 311–322 (2018)



# Integrated Energy Service Business Applicability Based on MRW Model

Hongcai Dai<sup>(✉)</sup> and Rui Chen

State Grid Energy Research Institute Co., Ltd., Beijing 102200, China  
ZXiaoTong99@163.com

**Abstract.** Since the start of the new round of electricity reform, the electricity-side market has been liberalized, and the market players have become more diversified. Domestic enterprises have set off a boom in the transformation to comprehensive energy services. However, the business forms of comprehensive energy services are diverse, and the application of traditional business models in such projects is obviously limited. In this process, it is necessary to study and explore the business model to strengthen business model innovation. Based on the analysis of the core elements of business model, this paper studies China's integrated energy service business model and explores the applicability of MRW model to explain the current economic growth. The results indicated that the MRW model showed differences in the test results of different samples. Through the development of Jiangsu Province, the applicability of the typical business model of integrated energy is analyzed.

**Keywords:** Power Grid Enterprises · Integrated Energy Services · Business Model · Key Elements · Energy Storage

## 1 Introduction

The concept of business model is often confused with the concepts of strategy and tactics. Strategy determines the overall development direction and top-level design of enterprises. For the construction of comprehensive energy services, the top-level design of enterprise strategy determines the overall method and strategy adopted by enterprises in the construction of comprehensive energy services, and tactical implementation determines the operation and implementation of enterprises at different business levels.

The integrated energy service system architecture is generally considered to include four levels: perception layer, network layer, platform layer and application layer. Among them, the application layer carries all kinds of business directions and contents, which is the main level of the business model. The specific implementation of different businesses will be reflected in specific technical projects and application projects. Therefore, the landing application of the project is set and constructed by specific business directions, and determines the success or failure of the construction and operation of enterprises comprehensive energy services. It can be found that the business model of integrated energy service presents three levels of logical relationship: company level, business

level and project level. Against the background of economic globalization, the economic development of various countries is constantly colliding. The global economy is accelerating its shift from the Atlantic to the Pacific, undergoing periodic fluctuations and emerging new trends and features. The study on the growing trend of the world economy is the focus that scholars around the world continue to pay attention to. Many studies try to provide new ideas and guidance for countries to formulate economic policies and guide economic growth through a more in-depth and clear interpretation of the development trend of the world economy. On the other hand, the study of the internal factors that affect the development of the world economy, such as technological progress, human capital and physical capital, in an attempt to explain the differences in the development of the world economy, and the exploration of effective ways to accelerate the economic growth of various countries have also been the front-burner issues of economists over the years. In the 1950s, Solow (1956) et al. [1], on the basis of in-depth research on the Harrod-Domar model, further proposed a neoclassical growth theoretical model to explain the differences in economic growth between countries and their sources of growth, which provided a basic framework for the subsequent economic growth analysis. Subsequently, Mankiw et al. [2] extended Solow model in a paper of pioneering contribution (hereinafter referred to as MRW model). The model regarded the saving and population growth as endogenous variables and introduced human capital way to enrich the Solow model, which improves the explanatory power of the model for the differences of economic growth among countries.

On this basis, a large number of scholars have examined the MRW model. Most of the researches focus on the MRW model and expand the model by further introducing variables, thus enriching the economic growth theory. Current researches focus on the level of empirical analysis of MRW model to explain the difference of economic growth in specific regions [4, 5, 11–13]. Rubina Vohra [14] applied MRW model to American states for empirical analysis and verified the adaptability of MRW model to the analysis of economic growth in American states with more homogeneous economic data and more unified institutional factors. Regression results give greater weight to human capital index, which has certain explanatory ability for economic growth, but it cannot replicate the empirical analysis results of Mankiw et al. [2]. Chengliang Yan [15] extended the MRW model by introducing R&D investment and health investment variables, and investigated the relationship between capital investment and economic growth with the data of 31 provinces in China. Dihai Wang et al. [16] used transnational data to investigate the impact of health, education and material capital investment on economic growth under the framework of MRW analysis. Changlin Yu [17] also used panel data to explore the impact of MRW model on explaining the human capital stock and its investment structure in the path of economic growth. Above all, in the process of such empirical studies, scholars mostly focus their research perspective on human capital index and conduct in-depth studies by expanding or redefining it [3, 6–10]. A few scholars explored the application environment and structural parameters of MRW model from the perspective of the research on its self-explanatory ability. For example, Bernanke [18] et al. re-tested Solow model and other growth models, finding that compared with Solow model, MRW model is more suitable for the analysis of balanced growth path. On this basis, the analysis framework of MRW model is extended, and its research path



provides reference for the evaluation of later growth model. Elena Pelinescu [19] studied the selection of human capital variables on the basis of MRW model, analyzed the value difference of different human capital indicators, and verified the important role of human capital in economic growth. Felipe [20] et al. tested the MRW model by using OECD data, and put forward some skeptical views on the MRW model combined with the empirical analysis results, showing that different procedures and methods should be selected for prediction test if the framework is to be used for economic growth analysis.

## 2 MRW Model

MRW model is extended by introducing human capital variable on the basis of Solow model, which is used as an econometric model to strictly test and derive economic convergence. On the basis of inheriting the traditional neoclassical economic growth theory, MRW model establishes a good starting point for empirical analysis of economic growth. In literature [2], the data samples were divided into three types of countries, namely 98 countries with oil production, 76 countries without statistical data problems and small countries, and 22 OECD countries with a population of over 1 million. Meanwhile, the Cobb-Douglas production function was used as the tool for empirical analysis. The model assumes that the production function is:

MRW model is extended by introducing human capital variable on the basis of Solow model, which is used as an econometric model to strictly test and derive economic convergence. On the basis of inheriting the traditional neoclassical economic growth theory, MRW model establishes a good starting point for empirical analysis of economic growth. In literature [21], the data samples were divided into three types of countries, namely 98 countries with oil production, 76 countries without statistical data problems and small countries, and 22 OECD countries with a population of over 1 million. Meanwhile, the Cobb-Douglas production function was used as the tool for empirical analysis. The model assumes that the production function is: Domestic comprehensive energy service business model research.

$$Z(t) = O(t)^a Y(t)^b [B(t)W(t)]^{1-a-b} \quad (1)$$

where,  $Z(t)$  represents output,  $O(t)$  represents physical capital,  $Y(t)$  represents human capital stock,  $B(t)$  represents the level of labor force and  $W(t)$  represents the level of technology. Meanwhile, it is assumed that the growth of  $n$  and  $g$  given outside the level of labor force and technology is:

$$L(t) = L(0)e^{nt} \quad (2)$$

$$A(t) = A(0)e^{gt} \quad (3)$$

$L(t)$  represents the effective labor, and its growth rate is  $A(t)$ .

### 3 Model Validation and Testing

#### 3.1 Stationary Test

In order to prevent pseudo-regression, after logarithmic processing of each variable, the stationarity test of economic variable data of each sample is carried out in this section. Four methods of unit root test were used to test the stationarity of each variable, and the test results of three samples were shown in Table 1 to Table 2. According to the analysis of test results, all the economic variables are stationary.

**Table 1.** Results of unit root test of Ln

Method\Statistic	Sample 1	Sample 2	Sample 3
LLC Analysis	-3.89056*	-3.52341*	-4.40000*
IPS Analysis	-3.34391*	-3.05572*	-1.62368***
ADF Analysis	322.401*	270.408*	106.184*
PP Analysis	238.113*	188.875**	80.1908**

Note: \* means significant at 1% level; \*\* means significant at 5%; \*\*\* means significant at the 10% level.

**Table 2.** Results of Ln ( $s_k$ ) unit root test

Method\Statistic	Sample 1	Sample 2	Sample 3
LLC Analysis	-6.12626*	-5.03730*	-1.02478
IPS Analysis	-5.02972*	-4.13118*	-2.08756**
ADF Analysis	294.274*	228.422*	79.6604 **
PP Analysis	287.399*	226.552*	73.5032***

Note: \* means significant at 1% level; \*\* means significant at 5%; \*\*\* means significant at the 10% level.

#### 3.2 Model Specification

Through the above unit root test, it can be known that all sample data are stationary sequences, indicating that there is a long-term stable relationship between economic variables and a regression model can be established.

On the basis of the above work, the configuration form of the panel data is verified. Combined with the economic theory and the characteristics of panel data, it can be judged that the sample is not suitable for the mixed panel data model.

## 4 System Architecture of Integrated Energy Service Business Model

### 1) Typical scenarios of integrated energy services

Definition of business model With the advancement and development of China's integrated energy services, the improvement of the business field of integrated energy services and the innovation of mechanism and system, China has carried out integrated energy services business in different kinds of systems, and at present, the typical scenes of China's integrated energy services can be mainly categorized into 8 kinds of commercial complexes, intelligent communities, oasis solitary networks, transportation hubs, new rural areas, heavy industrial parks, entertainment and tourism areas, photovoltaic poverty alleviation areas shown in Table 3.

**Table 3.** Typical scenerios of energy services

Category	Load characteristics	Resource Endowment	Self-governance capacity	Policy environment
Commercial complexes	Large and stable	Poor resource endowment, little room for renewable energy development	Poor autonomy and high dependence on large power grids	Low policy support
Intelligent community	Stable	Landscape primary energy can be developed according to the size of the community	Poor autonomy and high dependence on large power grids	Higher level of subsidies for residential distributed clean energy
Oasis Orphan Network	Stable	Abundant wind and light resources	Local independent power grid, low dependence on large power grid	Low policy support
Transportation Hub Area	Large	Less space available for primary energy development	Poor autonomy and high dependence on large power grids	Medium level of policy support
New Rural	Stable	There is more space available for scenic primary energy in rural areas	Poor autonomy and high dependence on large power grids	High policy support, with government subsidies and price advantage

(continued)

**Table 3.** (continued)

Category	Load characteristics	Resource Endowment	Self-governance capacity	Policy environment
Heavy Industrial Park	Large and stable	Industrial production waste heat and pressure resources are abundant	Some autonomy, large High grid dependency	Low policy support
Entertainment and tourism area	Large	Less space available to develop primary energy within the region	Poor autonomy and high dependence on large power grids	Low policy support
Photovoltaic Poverty Alleviation Zone	Small	Solar energy resources are abundant and can be used to develop a large space for primary energy	Self-generation and feed-in of surplus within the region, with low dependence on the grid	High policy support, with government subsidies and price advantage

## 2) Integrated Energy Services Profit Model

There are mainly the following four profit models.

### A. Integrated distribution and sales model

The difference between integrated power distribution and sales companies and other power sales companies is that they can earn revenue from the power sales business while also earning distribution revenue from the distribution grid business. Within the operation scope of the distribution network, if a customer signs a contract with a distribution company directly, the company can only pay transmission costs to transmission network operators, and the rest of the revenue belongs to the company. In either case, the integrated distribution and sales company can guarantee a profit source, which is the guarantee for the company's sustainable operation and development. As a distribution and sales company, it is easier to take the lead in the electricity sales market because it has distribution resources, becoming a guaranteed electricity sales company and gaining more customers for the company, while also using distribution network resources to carry out value-added electricity sales services, such as combined demand-side response and energy management, and also using customer resources to participate in the electricity auxiliary market.

### B. Supply and marketing cooperative model

The electricity sales company in the supply and marketing cooperative model combines power generation with electricity sales. The members of the cooperation have the power generation resources to sell electricity directly to other members in the form of electricity supply and sale cooperation, and part of the proceeds from the sale of electricity will continue to be invested in the construction of power plants to achieve a win-win situation for both parties. The biggest advantage of the

collaborative supply and marketing model is that it can obtain high-quality power generation resources, especially for distributed renewable energy power stations. Power plant operators are more willing to join the cooperative model of power sales, and the cost of power purchase for the company is relatively reduced.

#### C. Integrated Energy Service Model

The integrated energy service model refers to the fact that some foreign electricity sales companies carry out electricity sales along with other energy services to the area, also known as integrated city energy companies. These enterprises generally provide electricity and gas services, and customers can sign a separate electricity or gas contract with the company, which will also provide an integrated energy package. This is an important means of attracting and retaining customers, as it allows them to enjoy multi-faceted energy services. This bundled sales approach attracts more customers, increases customer loyalty, and diversifies profit sources. Aggravates the financial burden of the company, thus leading some foreign regional integrated energy service companies into financial difficulties and even on the verge of bankruptcy.

#### D. Electricity sales discount model

To better attract customers, discounters offer not only lower basic rates, but also attractive discounts for new customers. Many new commercial and industrial customers can significantly reduce their initial electricity costs through these packages, while residential customers can potentially reduce their electricity bills by 20% in the first year through discounts and discounts. For some customers, even lower discount prepaid electricity can be provided. Power retailers are the link between mass production and small-scale sales of electricity, and must participate in both wholesale and retail electricity markets.

## 5 Conclusions

The results of empirical analysis show that the MRW model has a high goodness of fit for the three sample countries, and the test results are in line with the original expectation for the samples from petroleum countries and intermediate samples. However, for OECD countries, the model shows a certain degree of inadaptability, which is mainly reflected in the action path of human capital. Firstly, for the first two types of samples, each economic variable obtained the expected symbol and was very significant, with little difference from the original influence path. Secondly, in the process of testing the economic data of OECD countries, human capital index shows variability. Through further analysis, it can be seen that the inapplicability of the model mainly stems from the fact that human capital is introduced into the model as a simple factor of production, which is the most likely reason for the difference. Meanwhile, the choice of human capital index measurement method will also have a certain impact on the result. Finally, the significant difference between the alpha and beta values calculated from the regression results indicates that it is difficult to replicate the original results of the MRW model. At present, most studies try to maximize the explanation of the difference in economic growth by improving the structure of human capital, but the mode of its entry into the model has not been perfected, which to some extent affects the stability of the empirical analysis results, which can be used as the main direction of future research.

At present, China's comprehensive energy service level and comprehensive energy efficiency level compared with the world's major economies, there is still a big gap. Make full use of information technology and physical energy system for close integration, in the supply side to large-scale use of clean energy to replace traditional fossil energy, in the demand side to large-scale use of electricity to replace. It should be noted that in the process of promoting digital empowerment, we should strive to innovate in basic theory and operational technology, and use reasonable business models. If we only call for or discuss concepts, it is difficult to make substantive breakthroughs.

**Acknowledgment.** This work is supported by Science and Technology Project of State Grid Corporation of China (Research on Integrated Energy Service Based on the Energy Internet Strategy (SGSDJY00GPJS1900057), No.520940210007).

## References

1. Solow, R.M.: A contribution to the theory of economic growth. *Q. J. Econ.* **70**, 65–94 (2020)
2. Mankiw, N.G., Romer, D., Weil, D.N.: A contribution to the empirics of economic growth\*. *Q. J. Econ.* **107**, 407–437 (2020)
3. Benhabib, J., Spiegel, M.M.: The role of human capital in economic development evidence from aggregate cross-country data. *J. Monet. Econ.* **34**, 143–173 (2021)
4. Lucas, R.E.: On the mechanics of economic development. *J. Monet. Econ.* **22**, 3–42 (2020)
5. Klenow, P.J., Rodríguez-Clare, A.: The neoclassical revival in growth economics: has it gone too far? *NBER Macroecon. Annu.* **12**, 73–103 (2021)
6. Barro, R.J., Lee, J.-W.: International comparisons of educational attainment. *J. Monet. Econ.* **32**, 363–394 (2020)
7. Engelbrecht, H.-J.: Human capital and economic growth: cross-section evidence for OECD countries. *Econ. Rec.* **79**, S40–S51 (2019)
8. Neycheva, M.: How might the negative impact of higher education on growth be explained? The role of vertical qualification (Mis) match in an MRW-type model. *Econ. Transition Inst. Change* **27**, 943–969 (2020)
9. Neycheva, M., Joensen, K.: Higher educational attainment for growth: the MRW model for Iceland. *Scand. J. Educ. Res.* **63**, 301–316 (2020)
10. Krueger, A.B., Lindahl, M.: Education for growth: why and for whom? *J. Econ. Lit.* **39**, 1101–1136 (2021)
11. Islam, N.: Growth empirics: a panel data approach\*. *Q. J. Econ.* **110**, 1127–1170 (2020)
12. Eberhardt, M., Teal, F.: Econometrics for grumblers: a new look at the literature on cross-country growth empirics. *J. Econ. Surv.* **25**, 109–155 (2020)
13. Vohra, R.: How fast do we grow? *Growth Chang.* **27**, 47–54 (2019)
14. Yan, C.: The impact of capital investment on China's economic growth – an analysis based on the expanded MRW framework. *J. Quant. Tech. Econ.* **28**, 3–20 (2021)
15. Wang, D., Li, X., Huang, L.: How does health investment affect economic growth: a study from transnational panel data. *Econ. Sci.* 5–17 (2022)
16. Yu, C.: Human capital investment structure and its economic growth effect – an endogenous growth theory and empirical study based on the extended MRW model. *J. Quant. Tech. Econ.* 117–125 (2021)
17. Bernanke, Gokoenak, R.S., Zhong, L., et al.: Exophytic discussion of economic growth – analysis based on MRW framework. *Econ. Soc. Syst. Comparison*, 31–39 (2021)

18. Pelinescu, E.: The impact of human capital on economic growth. *Procedia Econ. Finance* **22**, 184–190 (2021)
19. Felipe, J., McCombie, J.S.L.: Why are some countries richer than others? A skeptical view of Mankiw–Romer–Weil’s test of the neoclassical growth model. *Metroeconomica* **56**, 360–392 (2019)
20. Chang, S., Shao, Y.: Explanation of labor market fluctuations and new trends in OECD countries. *Theory Mon.* 135–140 (2020)
21. Nerlove, M.: Comments on: panel data analysis - advantages and challenges. *TEST* **16**(1), 42–46 (2019)



# Application of Internet of Things and Virtual Reality Technology in the Intelligent Reconstruction of Old Residential Quarters

Ming Lv, Miao Yu<sup>(✉)</sup>, and Wenting Ding

Shenyang Jianzhu University, Shenyang, Liaoning, China  
ym13604674881@126.com

**Abstract.** In order to speed up the reconstruction of old communities in our country and promote the transformation into a modern smart city. Taking Shenyang as a reference, analyzing the reconstruction situation of old communities, summarizing a series of problems that existed in the initial project design, and having a deeply research on the intelligent reconstruction of the old community in Shenyang. Under the influence of the development trend of artificial intelligence, the feasible method of using Virtual Reality Technology and Internet of Things to assist the initial design is proposed. Can be used as a reference for old communities' reconstruction in Shenyang, and further providing a new idea for the construction of a modern smart city in the whole country.

**Keywords:** Internet of Things · Virtual Reality Technology · Intelligent Transformation

## 1 Introduction

Along with the steadily social development and the gradually increasing of people's needs for material life, the old communities built in 1980s and 1990s can no longer meet the living requirements to current residents. After 20 to 30 years worn with long time service, various problems have shown out with the old communities that was built in early time, like house main body damage, circuit aging, supply shortage of convenient facilities, all these issues urging the old community's reconstruction projects imminent. In order to improve the residents' living standards and the overall style of modern cities, and improve the happiness index of urban residents as well, the reconstruction projects of old communities in our country have begun to be promoted on a large scale under the guidance of the relevant national policies.

The transformation of old residential areas in foreign countries started early, and was divided into two stages in the 1960s and 1970s. Countries have issued policies or government led reforms. The transformation scheme is mainly divided into specific technical research and management evaluation optimization. The transformation of old residential areas in China started relatively late, which can be divided into three stages: before the reform and opening up, from the beginning of the reform and opening up to



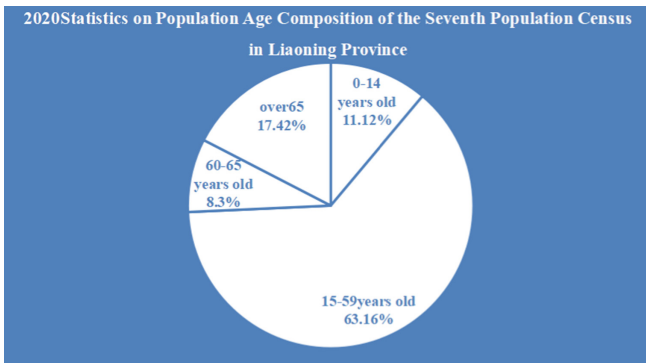
the end of the 20th century, and from the 21st century to the present. According to the different research done by different students, the reconstruction of old residential areas in China has achieved rich results, which provides reference for the subsequent practical projects.

According to the survey, there are many problems mentioned above in some old residential areas in Shenyang, which have become a major obstacle in the process of urbanization. The Internet of Things and virtual reality technology have good operability. Based on the combination of Internet of Things and virtual reality technology, this paper proposes the application of three aspects of old residential area terrain route planning, scene, facility setting and mechanism linkage optimization, which provides new inspiration for the future intelligent community.

## 2 Old Community Transformation Status and Existing Problem in Shenyang City

### 2.1 Overview of Old Community Transformation

The old communities were built in early years, with low construction standards, and lack of public service facilities as well. In addition, many of these old communities have not been maintained or repaired for many years, the performance of buildings is gradually degraded, plus the mixing traffic, also have the safety management issues, all these issues and concerns have seriously affected the cities' modernization process in our country. The old communities already became the shortcoming to city quality improvement. Based on the seventh census result, the population of 65-years-old people in Liaoning Province accounted for 17.42%, which is ranking our country the first place. This means Liaoning has entered a stage of deep aging and that will bring resistance to the development of modern cities. Therefore, improving the living environment has become an important way to maintain urban vitality and satisfy social equity [1], as shown in Fig. 1.



**Fig. 1.** Statistics of population age composition in the 7th population census of Liaoning Province in 2020

In July 2020, according to the “Guidelines on Comprehensively Promoting the Reconstruction of Old Communities in Urban Areas” issue by Office of the State Council, which is required to combine the actual situation of various places to fully promote the reconstruction project of the old community, the guidelines bring forward that we need to benefit people’s livelihood, expand domestic demand, to further transform the mode of urban renewal and also for development and construction. On March 24, 2021, the Information Office of Liaoning Provincial Government held a press conference, proposing that Liaoning Province will build a habitable, green, tenacious, smart and humane city, innovate the urban renewal model, and take the road of urban renewal in the old industrial base.

Under the guidance of various documents and policies, the old community reconstruction project can better find a new route that is suitable for its own development. Under the guidance of intelligent technology of IoT, the reconstruction project will show more comprehensive and more thorough characteristics.

## 2.2 Transformation Methods and Content

In the early days, our country was mostly employed the shed reform mode. Because of the long-time rebuild for the shed reform, the original living circle has been destroyed, plus the traditional neighborhood relations and residential cultural atmosphere cannot be established in a short period of time. With the construction of urban residential areas and the housing prices increasing across the country, the demolition of sheds came to the end. So far, the residential communities that built in the 1980s and 1990s have become key targets for reconstruction.

The old community reconstruction is closely connected with people livelihood, and also closely related to city modernization. According to the document of “Reconstruction Standards for the Reconstruction of Old Residential Districts in Shenyang City in 2021”, there will be three kinds of reconstruction standards of Shenyang old communities: Basic, Improve and Upgrade. The Basic category means the project must be reconstructed so that can meet the residents’ safety requirements and basic living requirements. The Improve category means the projects that can be modified, the purpose is to meet the needs of residents’ convenience and improve livelihood. The Upgrade category means the projects that should be reconstructed, main purpose is to enrich the supply of community service, including reconstruction and construct the comprehensive service facilities in the community or the surroundings, to achieve the needs of modern communities.

According to the “Three-year Working Plan for the Reconstruction of Old Communities in Shenyang (in the park)”, in order to meet the needs of residents as much as possible, the project is planned to be carried out in the form like one designer-one community and one drawing-one community. The old community reconstruction is not only for a single community, but also for the entire district. To bring up the entire district as whole, expanding the courtyard projects to the outside streets. Not only in the community inside, the residents will feel the new look in the entire district as well.

### 2.3 Problems in the Transformation

After studying the issues from various aspects, starting from the initial design view, refining the issues by focusing on the improve and upgrade categories that be pointed out in reconstruction standard, and bring up the improvement direction [2]:

- ① The community parking space is insufficient. Need to design and develop the ground and underground spaces. Induction monitoring system is adopted to investigate in advance the residents' routes of daily activities, and to plan the distribution of mobile, non-motor vehicles and pedestrian lines in the community.
- ② The community basic convenient facilities is insufficient. It needs to be improved, and the facilities style need to be integrated with surrounding environment.
- ③ No clear distinguish between greening and culture activities area, or even no plan for this area, it is necessary to collect data and replan the region.
- ④ Security services and property management have weakness. The formation of intelligent community needs systematic improvement through IoT.
- ⑤ Construction period is too long that will affect the normal life of residents. Needs to shorten the early time design lead time.

## 3 Realization of IoT and VR Technology

The design must be emphasized people-oriented, keep user experience as the core. The old community construction project must fully cooperate with the dynamic development of city, so that can avoid irreversible errors. Thus, the early design requires more efficient technical to support and to save cost [3].

The Internet of Things (IoT) monitors objects and collects data in real-time through various information sensors, positioning systems, identification technologies, laser scanners and other devices and technologies, so as to realize the extensive connection between things-things and human beings-things. Virtual Reality (VR) technology can make fully utilize computer-assisted equipment to create a virtual three-dimensional space, so that can create a simulated artificial world. The traditional human-computer interaction can be broken under the virtual environment, so that can realize the multi-sensory interaction, like vision, hearing touch, smell, etc. Compared with traditional design drawing shows, virtual reality technology has outstanding advantages: Firstly, it can make people feel like immersive. Secondly, the experienter's behavior is able to have a certain effect on the virtual world. Thirdly, human-computer interaction experience has been enhanced. This powerful interaction function can give people an impressed immersion and the designer's intentions [4]. The combination of IoT and VR is the future development trend of artificial intelligence. At present, the Internet of Things is operated and monitored by 2D data interface, not intuitive enough to operate. If virtual reality technology is used, the scene can be visualized through 3D. Through data interconnection, the actual situation of the scene is reflected into the 3D scene, creating an artificial environment, and solving problems intuitively and simply. Enable users to make more real feedback on the design scheme [5].

Use high-tech electronic equipment to close the user's sense into a virtual environment. With immersive technical support, designers and users can enter a limited space

at any stage of design, and observers can experience the design results from the first perspective. The combination of the two technologies can not only make the preliminary design work achieve strong immersion, interaction and conception, but also provide core technical support for the later construction of intelligent residential area [6].

Specific Application of IoT and VR in Transformation Projects.

The advantage of combination of the two technology is good maneuverability. According to the aging degree, and different needs of the residents, the requirements of infrastructure and environment layout are also different, the new technology can use its own unique advantages to achieve design autonomize, to improve the efficiency of communication when modify the plan, the early design period can be shortened. At the same time. In the overall planning of intelligent system, enhance the security needs of residents. Designers should put themselves in service for the residents, should abandon the fantasy, to create the condition to make residents take part in innovation. Only the residents' wishes are respected, the service projects such as reconstruction can deeply rooted in peoples' heart.

### **3.1 Application of IoT and VR in Terrain Route Planning of Old Community**

During the visitation and investigation, it was found that the old communities exist outstanding issues, like unclear division of parking areas, the confusion of entertainment areas and activity areas, and the interspace underutilized. When do the planning, the designers should fully consider the residents' activity line, designers cannot plan the layout based on their own feelings. The optimization of road routes for pedestrians and vehicles needs to coordinate with the habits of community residents.

Using VR, the road structure, facility layout, environment construction of the community can be shown out. In the road planning, first through sensor, positioning and identification technology collect and record the data of the road topography inside the community and surrounding cities, and then construct a virtual environment similar to the objective world. Through observe the resident's life style, action track, use the IoT processing system to integrate the data of his moving line and optimize the road: firstly, to simulate the people gathering point and flow point, to broaden the road with high traffic timely. Secondly, setting the fixed facility and entertainment at low mobility area, to make sure no affect each other. Thirdly, inside the feasible community, expand the underground space for parking, storage, emergency exit and other functions.

The road scene will be displayed through holographic projection technology after design finished. The observer can pre-view the actual effect under the virtual environment, and can record the observer's feedback at the same time. It helps residents to accept the convenience brought by the reconstruction intuitively, and can also save other departments' review time and costs. The virtual reality scene model can be considered from the public's perspective, which reflects the design conception based on people-oriented, the early designers can collect residents' suggestions timely, so that the reconstruction projects can better solve the reality problems.

### 3.2 Application of IoT and VR in Old Community Scenes and Facilities Layout

The community facilities must keep their own characteristics, and cooperate with the overall community scene and even the surrounding cities' modern development. Designers should optimize the various facilities and scene inside the community, should divide and allocate the space function of community, exploit the fixed area on the ground as public activity area, adding aging-forward and barrier-free facilities. From the vehicle charging equipment, fitness equipment, to the public health room facilities, kindergarten education facilities, and intelligent protection facilities, the designer should consider all these facilities to make a rational layout [7].

The intelligent technology of IoT is used to collect the regional data in real time and transmit it to the processor. Multiple feasible schemes are provided through calculation. VR can control the design of the flow line layout freely, do the interaction of human-computer during simulation, so can let users experience the roaming scene at different times and locations. Through the virtual experience of scenarios and facilities layout, users will evaluate the degree of design coordination, and have an intuitive experience of environmental design plan. IoT and VR allow people to react in a virtual environment, improving the overall user experience [8].

### 3.3 Application of IoT and VR in Mechanism Linkage Optimization

The old communities generally have a lot of problems, such as insufficient functional matching settings and lack of property management mechanisms, and there are more elder people, when emergency happens, kinds of dangerous will occurred [9]. IoT can realize environmental monitoring and data collection through various information sensors, positioning systems, identification technologies, infrared sensors, laser scanners and other devices and technologies. IoT sensors that collect mechanical data enable predictive maintenance. At the same time, by using VR, it becomes possible to simulate community scene and operating process, unexpected problems in reality can be solved by using the virtual.

Take the intelligent security system as an instance, collect the scene data for the community border protection, the building entrance, community gate and the public area. Real-time safety monitoring is carried out by complex sensors and the generated data is processed. In the early system design, VR can be used to simulate the entry process of community residents and visitors, do the virtual exercise for special situation and emergency handling. According to the simulation result, upgrade the old community systems and processes timely, to centralize transaction monitoring and control for unexpected problems, ensure the harmony and coordination between community internal safety mechanism and the residential environment. By this way, the residents will be in an intelligent smart community [10].

In additional, the design should be adapted to local conditions, to build a management method that suitable for the characteristics of this community, and also proceed the initial design analysis for fire-warning system, intelligent patrol system, to make the seamless connection of information change between people-things, things-things as a reality. The efficient and coordinate operation between sub-system and systems of the community be beneficial to improve community management level and social development efficiency.

## 4 Conclusions

Our life is always under the dynamic development. Therefore, the huge cost of reconstruction project makes it become an irreversible procedure. The designer should have an ability of spatial thinking, and focus on user's feeling. Using the IoT and VR can utilize technological information to assist the design process, realize multi-party reality experience such as audio-visual in the virtual scene, making the reconstruction design more creative and flexible. Artificial intelligence needs the data generated by the IoT, while VR combines all this with immersion and visualization, to create more powerful simulations that enhance the user experience and understanding of the scene. The application of intelligent community has covered kinds of aspects of human life. The combination of IoT and VR is the trend of intelligent development. They will integrate the unique advantages of multiple aspects, to open a new breakthrough for the development of industries. In the future, it will be inevitably that virtual reality technology will be applied into intelligent city construction widely, and bring a new inspiration to the modern lifestyle.

## References

1. Zwierzchowska, I., Haase, D., Dushkova, D.: Discovering the environmental potential of multi-family residential areas for nature-based solutions. In: A Central European cities perspective", *Landscape and Urban Planning*, vol. 206, pp. 103975–103987 (2021)
2. Nooringsih, K., Susanti, R.: Implementation of smart city concept for sustainable development in Semarang old town area. *IOP Conf. Ser. Earth Environ. Sci.* **1082**(1) (2022)
3. Kilani, R., Zouinkhi, A., Bajic, E., Abdelkrim, M.N.: Socialization of smart communicative objects in industrial internet of things. *IFAC Papers OnLine* **55**(10), 1924–1929 (2022)
4. He, P., Almasifar, N., Mehbodniya, A., Javaheri, D., Webber Julian, L.: Towards green smart cities using Internet of Things and optimization algorithms: a systematic and bibliometric review. *Sustain. Comput. Inform. Syst.* **36** (2022)
5. Allal-Chérif, O.: Intelligent cathedrals: using augmented reality, virtual reality, and artificial intelligence to provide an intense cultural, historical, and religious visitor experience. *Technol. Forecast. Soc. Chang.* **2022**, 178 (2022)
6. Bayramova, R., Valori, I., McKenna Plumley Phoebe, E., Callegher Claudio, Z., Farroni, T.: The role of vision and proprioception in self-motion encoding: an immersive virtual reality study. *Attention Percept. Psychophys.* **83**(7), 1–14 (2021)
7. Kent, G., Simon, D.: Upscaling community transformation. *Action Res.* **18**(4), 407–413 (2020)
8. Sánchez Mario, R., Palos-Sánchez Pedro, R., Velicia-Martin, F.: Eco-friendly performance as a determining factor of the adoption of virtual reality applications in national parks. *Sci. Tot. Environ.* **798** (2021)
9. Anna, G.: Karl Linn and the foundations of community design: from progressive models to the war on poverty. *J. Urban Hist.* **46**(4), 794–815 (2020)
10. Nathan, L., Michael, H., Myounghoon, J.: Ergonomics in AI: designing and interacting with machine learning and AI. *Ergon. Des. Q. Hum. Fact. Appl.* **28**(3) (2020)



# Construction of English Teaching Multimedia Resource Library Based on Cloud Computing Technology

Huan Jiang<sup>(✉)</sup>

Department of Basic, Liaoning Vocational University of Technology, Jinzhou, Liaoning, China  
550138468@qq.com

**Abstract.** Cloud computing, as a new type of network service model, shares computer resources such as computing, data and storage in the all-round way, provides dynamic and scalable virtualization resources, and provides the good basic support platform for college education reform. Multimedia as an auxiliary means, the use of multimedia technology can make teaching methods rich and diverse. The construction of the multimedia resource library based on cloud computing technology is in line with the development direction of English teaching. Taking the development of mobile information technology and teaching reform as the opportunity, we have studied the core technology of cloud computing and cloud computing technology architecture, designed the structure of the multimedia resource library for English teaching, and proposed the practical application strategy of the multimedia resource library for English teaching, and actively promoted English teaching and research is undergoing the comprehensive transformation and upgrading towards modernization.

**Keywords:** Cloud Computing Technology · Core Technology · Cloud Computing Architecture · Multimedia Resource Library · Construction Program

## 1 Introduction

As the emerging technology, cloud computing can use the network to integrate different computers to form the super-large platform, and share the computing, data, and storage resources of the computer in the all-round way [1]. The rapid development of information technology, cloud computing as the new network service model, cloud users can break through the boundaries of time and space, use terminal facilities, through the network platform to obtain the information resources they need [2]. Cloud computing not only has the function of cloud storage and shared pool, but also provides dynamic and scalable virtualized resources, providing the good basic support platform for college education reform. With the help of cloud computing to provide functions, build the multimedia resource library, and improve the practicality of education and teaching data information. At the same time, it supervises the effectiveness of the implementation of various tasks, focuses on the resource supply of colleges and universities, and through in-depth exploration of data information, helps colleges and universities optimize and integrate

resources such as resource retrieval, resource browsing, resource recommendation and resource collection, and improve the level of education informatization.

The construction of the multimedia resource database is the important content for advancing teaching reform. Through multimedia resource sharing, English educators can find suitable materials on the Internet for reprocessing, or upload wonderful courseware to the resource database. English teachers can also use the resource to improve teaching ability and professional level while sharing to broaden horizons. By using the multimedia resource library, students can develop the wide range of online learning, select resources according to their own interests and hobbies, and improve their independent learning ability. Teachers are not only the transmitters of knowledge, but also the guides and creators of learning. Reflecting that the charm of education is the charm of creation, it is the charm of creating the development of life. Through the research of this subject, we will help English educators overcome the difficulties in online teaching, make online teaching more people choose and accept, and help the development of English education online teaching.

## 2 Cloud Computing Core Technology

Cloud computing is the integration of several technologies. The core technology is composed of “Distributed computing, Internet technology, and Resource management”, as shown in Fig. 1.

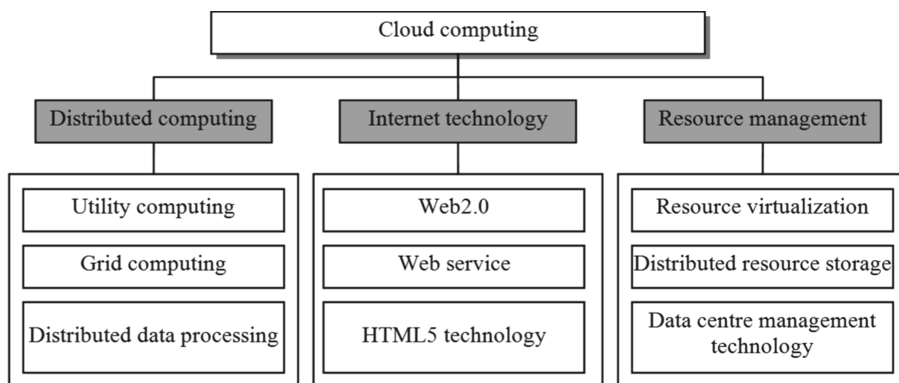


Fig. 1. Cloud computing core technology

### 2.1 Distributed Computing

Distributed computing is the result of improved personal computer performance. It is more convenient to use multiple computers to process tasks at the same time and share resources [3]. The problem that requires huge computing power to be solved is broken down into many small parts, and these small parts are allocated to many computers for processing, and then the calculation results are combined to obtain the final result. The soul of distributed computing is to balance load and share resources.



### 2.2 Internet Technology

Internet technology is the sign of entering the information society, accelerating the speed of information transmission and broadening the channels for obtaining information. HTML5 is currently the most important technology. It strengthens the performance of Web pages, separates content and display, introduces the more reasonable brand new logo, replaces Flash’s position in mobile devices, and becomes the most excellent performance design language for text-based display [4].

### 2.3 Resource Management

Cloud computing uses the sharing of resources to improve resource utilization efficiency, thereby achieving the purpose of reducing the cost of computing. Virtualization breaks the boundaries of time and space and is the most significant feature of cloud computing. Resource regulation plays the key role in overall system performance. Distributed resource storage stores data in various corners and uses location servers to locate and store information.

## 3 Cloud Computing Technology Architecture

Cloud computing technology has developed over time and has been applied in many industries. Using cloud computing technology to minimize the hardware differences in different integrated network system management equipment, improve the management

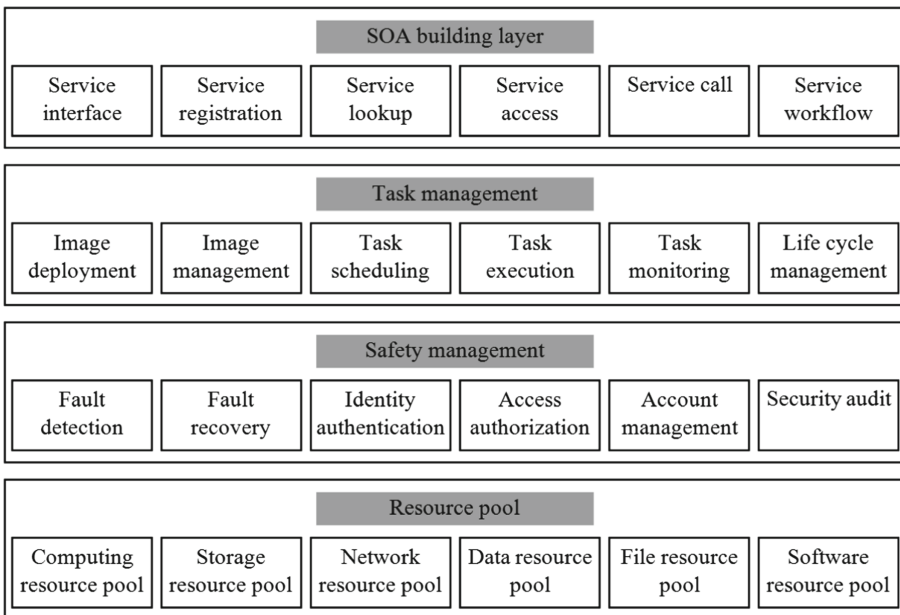


Fig. 2. Cloud computing technology architecture

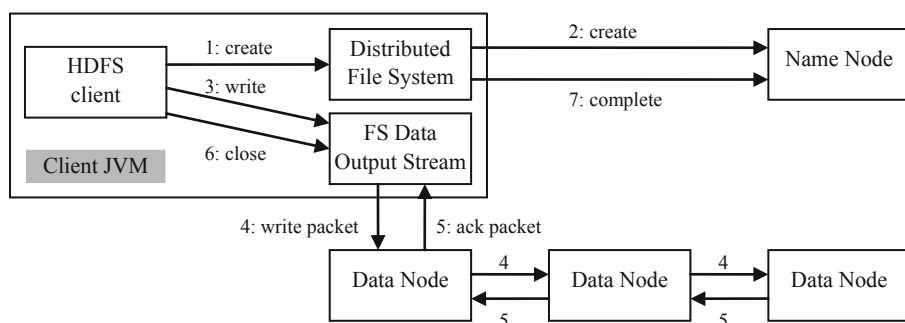
level and scope of data integration, can effectively reduce management costs, and keep it within the relatively reasonable range. The cloud computing architecture is shown in Fig. 2 [5–7].

The cloud computing architecture shown in Fig. 2 is composed of SOA building layers, task management, security management, and resource pools. Security management is the quite important content. Data security directly affects the application and development of cloud computing. The low-cost, virtualized, multi-tenant, and ubiquitous features of cloud computing cater to the needs of information services. However, the unique distributed features of cloud computing are easily targeted by intruders. Data security and privacy protection are also security issues that users are most concerned about. Measures such as data privacy protection, network transmission security, and cloud service security management and monitoring are required.

#### 4 Upload File Process of HDFS

The Hadoop distributed file system (HDFS) has become the preferred underlying storage system for computing platforms of Internet enterprises due to its high reliability, high throughput, and high scalability. The HDFS uses Datanodes as data storage nodes, and NameNode maintains file directory information. Files in the HDFS consist of data blocks of the same size and are evenly distributed among Datanodes in the cluster based on the data block distribution algorithm. When a user requests to query the MetaData of a file, the NameNode feeds back the living storage nodes in the file to the user based on the existence of the current node. Then the user accesses the living DataNodes for data access [8–10].

Transferring the finished teaching resources to the network storage space is an important work in the construction of teaching resource database. This system uses the Writer object in the API provided by Hadoop to upload the data of teaching resources to the distributed storage system. The uploading process is shown in Fig. 3.



**Fig. 3.** Upload file process of HDFS

For the Upload file process of HDFS shown in Fig. 3, the execution process is briefly described as follows: The client calls the create () method to create a distributed file;

The create() method requests the data node number from the NameNode; The client calls the write() method to write data to the output stream, divides the data according to the size of the data block in the configuration file, and backs up the data. Data is transmitted to Datanodes in the form of packets. Regardless of whether the data node is successfully written to the data, it will return a message indicating whether the data is successfully written. After the data write operation is completed, the client calls the close() method to close the file transfer stream. FSDataOutputStream returns the result and notifies NameNode that the data upload is complete. When Hadoop's Writer object writes data, it needs to configure the main parameters, as shown in Table 1.

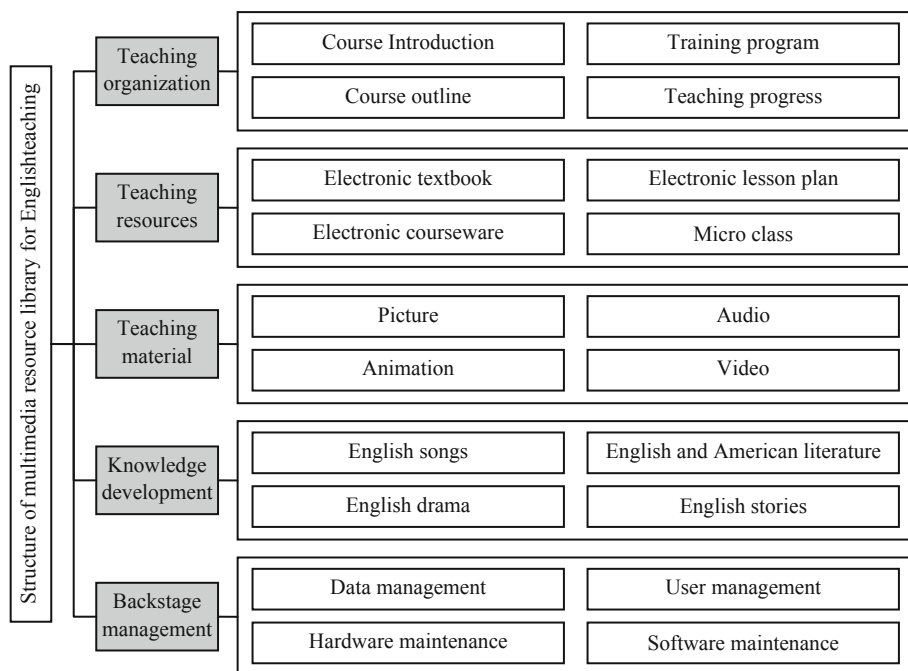
**Table 1.** Main parameters of Hadoop's writer object

No	Parameter name	Parameter description
1	Default FS	Name Node address of the file system
2	File Type	Type of file to write
3	path	Path to the file system
4	File Name	File name when writing
5	Write Mode	Patterns for data cleansing and processing before writing
6	Field Delimiter	Field separator on write
7	compress	File compression type
8	encoding	Encoding configuration of file
9	Parquet Schema	Describe the structure of the target file
10	Hadoop Config	Relevant advanced parameters can be configured
11	Datax Parquet Mode	Synchronization mode of Parquet file
12	Have Kerberos	Whether there is Kerberos authentication

Datanodes are the working nodes of the file system. They are like the peripheral nerves that provide data for each perception. The Name Node is like the human brain, synthesizing neural feedback. Data Node stores and retrieves databases as needed, and periodically sends a list of storage blocks to Name Node, which records how to reconstruct files. Based on the data storage mechanism of the distributed storage system, developers need to select an appropriate read/write consistency algorithm to meet the performance requirements of storage applications. Since storage mechanisms of replicas and erasure codes differ greatly, distributed storage systems need to deploy corresponding read and write consistency algorithms to adapt to data read and write performance under various storage mechanisms. Because the networks between storage nodes are heterogeneous, the read/write consistency algorithm needs to adjust the network routing policies between storage nodes based on the network status of the distributed storage system to ensure the data read/write performance of the distributed storage system.

## 5 Structure of Multimedia Resource Library for English Teaching

The multimedia resource library under cloud computing refers to the use of more efficient network technology to integrate a series of applied electronic resources. Using cloud computing technology to strengthen the application and management of the multimedia resource library, it facilitates the user's acquisition and utilization of data information, reduces the user's time cost, ensures the accuracy and safety of the acquired data and information, and improves the computer compatibility of the operating system. The structure of the multimedia resource library for English teaching is shown in Fig. 4.



**Fig. 4.** Structure of multimedia resource library for English teaching

As the application of modern educational technology, multimedia resources can provide vivid audiovisual images to stimulate students' imagination and play the positive role in promoting teaching. However, if the opportunity is not accurate, it may have counterproductive effects. When explaining new words, you can use pictures; when teaching grammar, you can use pictures to increase the vividness of grammar teaching; when teaching culture; you can use pictures, music, and videos. Teachers should grasp the opportunity of the use of multimedia resources according to the goals of English teaching, reasonably arrange the frequency of use of multimedia resources, and ensure that their use achieves the best result. Multimedia has the characteristics of diversification and integration of information carriers, and the integration of multimedia and foreign language teaching is the inevitable development trend of future teaching models. The

use of multimedia resources to properly extend the learning content will further enhance the teaching effect.

## 6 Conclusions

Multimedia as the auxiliary means, the use of multimedia technology in teaching can make teaching methods rich and diverse. Teachers can master the basic skills and resource production methods of multimedia technology, combine the needs of offline teaching and online learning, and choose content reasonably, through video, audio and animation methods are widely used in multimedia technology, combining multimedia and teacher teaching content, breaking through traditional teaching methods, improving the quality of English teaching, and enhancing students' familiarity with knowledge.

## References

1. Fabrizio, M., Loris, B.: Cloud computing for big data analysis. *Appl. Sci.* **12**(20), 424–431 (2022)
2. Prachi, D., Mukund, K., Sanjay, N., Anil, N.: Taxonomy of load balancing practices in the cloud computing paradigm. *Int. J. Inf. Retrieval Res. (IJIRR)* **12**(3), 1–15 (2022)
3. Nath, R., Nagaraju, A.: Genetic algorithm based on-arrival task scheduling on distributed computing platform. *Int. J. Comput. Appl.* **44**(9), 887–896 (2022)
4. Kang, S., Lee, J.: Improving rendering speed of 3D geospatial data based on HTML5/WebGL using improved arithmetic operation speed. *Int. J. Urban Sci.* **23**(3), 303–317 (2019)
5. Guzman, A.: Phase model of adopting cloud computing in the implementation of e-government. *J. Res. Sci. Eng.* **4**(7), 78–84 (2022)
6. Rajeshwari, S., Singh, R., Mohan, M., Varun, B.: A multi-objective optimization scheduling method based on the genetic algorithm in cloud computing. *Int. J. Cloud Appl. Comput. (IJCAC)* **12**(1), 1–21 (2022)
7. Muriqi, G., Mutuku, M.: Security issues and challenges in cloud computing among public institutions in Africa. *J. Bus. Manag. Sci.* **10**(3), 131–137 (2022)
8. Gaykar, R.S., Khanaa, V., Joshi, S.D.: Faulty node detection in HDFS using machine learning techniques. *Revue d'Intelligence Artificielle* **36**(4), 553–560 (2022)
9. Santhana, M.V., Natarajan, V.: Efficient big data security analysis on HDFS based on combination of clustering and data perturbation algorithm using health care database. *J. Intell. Fuzzy Syst.* **43**(3), 3355–3372 (2022)
10. Senthil, K.P., Nurul, B., Aboo, A.B.: Data integrity verification using HDFS framework in data flow material environment using cloud computing. *Mater. Today Proc.* **60**(3), 1329–1333 (2022)



# Design and Application of Japanese MOOC Teaching Resources System Based on User Collaborative Filtering Recommendation Algorithm

Xuemei Zhang<sup>(✉)</sup>

College of Foreign Languages, Bohai University, Jinzhou, Liaoning, China  
2965738909@qq.com

**Abstract.** The core of personalized recommendation is an algorithm, which searches for a group of users with similar interests to the target user as “nearest neighbors”, and realizes recommendation by finding “nearest neighbors”. According to the actual needs of MOOC teaching in Japanese major, a teaching resource system is built, and based on the user collaborative filtering recommendation algorithm, the teaching resources of interest are recommended to users to meet the needs of MOOC teaching in the mobile information age. The core work completed in this paper mainly includes two aspects: First, a mathematical model is constructed according to the user-based collaborative filtering recommendation algorithm process. Second, based on the use case diagram method, the functional requirements analysis is carried out for the student users, the teacher users and the system administrators, and the key problems of the system development are solved.

**Keywords:** User Collaborative Filtering · Recommendation Algorithm · MOOC Teaching Resources System · Design and Application · Mathematical Model · Functional Requirement Analysis

## 1 Introduction

With the advent of the information age, teaching resources have grown rapidly, and digital teaching resources have grown exponentially, resulting in a serious “information overload”. It is a difficult problem for students to find the teaching resources they need accurately and quickly [1]. The construction of teaching resources is the basic work for carrying out MOOC teaching. Although many MOOC platforms provide teaching resources, they still cannot meet the needs of teaching. Teaching resources include textbooks, audios, videos, question banks and courseware, etc. Japanese teachers should make full use of their own advantages, use Internet technology to integrate diversified network teaching resources, develop novel and applicable Japanese teaching resources based on modern educational technology, and upload them to the MOOC platform for sharing. It is also possible to develop a small Japanese teaching resource library according

to its own teaching needs, with the help of the network, we can innovate and optimize our own teaching mode, enrich teaching resources and effectively improve teaching quality, and expand students' knowledge space [2–4]. Personalized recommendation is an intelligent technology based on massive data mining, from the mass of information mining users interested in content. Make full use of the content mined from massive data, in view of the actual needs of MOOC teaching in Japanese, this paper builds a teaching resource system to meet the needs of MOOC teaching in the mobile information age.

## 2 Mathematical Model Based on User Collaborative Filtering Recommendation Algorithm

This algorithm is divided into five steps, and the algorithm flow is shown in Fig. 1.

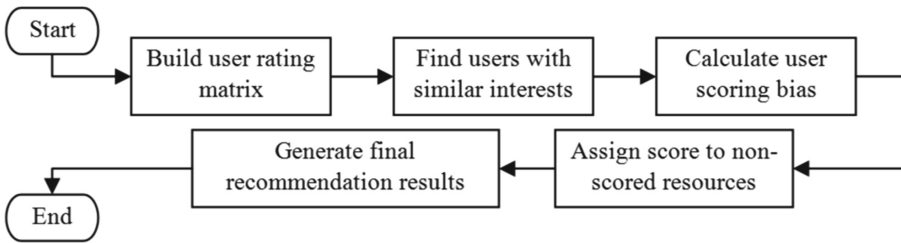


Fig. 1. Algorithm flow

In this paper, a mathematical model is constructed for the five steps shown in Fig. 1.

### 2.1 Build User Rating Matrix

The user set is  $U = \{U_1, U_2, \dots, U_i, \dots, U_m\}$ , and  $m$  represents the number of users; the resource set is  $S = \{S_1, S_2, \dots, S_j, \dots, S_n\}$ , and  $n$  represents the number of resources. The user rating matrix composed of  $m$  users and  $n$  resources is:

$$R < m \times n > = \begin{bmatrix} R_{1,1} & R_{1,2} & \dots & R_{1,j} & \dots & R_{1,n} \\ R_{2,1} & R_{2,2} & \dots & R_{2,j} & \dots & R_{2,n} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ R_{i,1} & R_{i,2} & \dots & R_{i,j} & \dots & R_{i,n} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ R_{m,1} & R_{m,2} & \dots & R_{m,j} & \dots & R_{m,n} \end{bmatrix} \tag{1}$$

### 2.2 Find Users with Similar Interests

The similarity calculation uses the correlation coefficient calculation method, and the commonly used methods are as follows [5–7]:

- (1) Cosine similarity coefficient, calculate the angle between two vectors to reflect whether the two vectors are similar. The dimensions of the vectors are not limited, any dimension can be compared. In the algorithm, cosine similarity is to treat the rating matrix as multiple rating vectors. Assuming that the vectors  $u$  and  $v$  are the rating vectors of two users, the similarity calculation formula is:

$$sim(u, v) = \cos(\bar{u}, \bar{v}) = \frac{\bar{u} \cdot \bar{v}}{|\bar{u}||\bar{v}|} = \frac{\sum_{i \in I_{uv}} R_{ui}R_{vi}}{\sqrt{\sum_{i \in I_{uv}} R_{ui}^2} \sqrt{\sum_{i \in I_{uv}} R_{vi}^2}} \quad (2)$$

- (2) The modified cosine similarity coefficient, the difference in the scoring habits of the two users is not considered in the cosine similarity. If the similarity used to calculate the score is directly used, the recommendation result may be inaccurate. Subtract the mean of all his ratings from the user's rating and this is the modified cosine similarity. The similarity calculation formula is:

$$sim(u, v) = \frac{\sum_{i \in I_{uv}} (R_{ui} - \bar{R}_u)(R_{vi} - \bar{R}_v)}{\sqrt{\sum_{i \in I_u} (R_{ui} - \bar{R}_u)^2} \sqrt{\sum_{i \in I_v} (R_{vi} - \bar{R}_v)^2}} \quad (3)$$

- (3) The Pearson correlation coefficient is used to describe the trend of the change and movement of the two sets of linear data. In the algorithm, when the Pearson coefficient of the two user ratings is close to 1, it is considered to be a user with similar interests. The similarity calculation formula is:

$$sim(u, v) = \frac{\sum_{i \in I_{uv}} (R_{ui} - \bar{R}_u)(R_{vi} - \bar{R}_v)}{\sqrt{\sum_{i \in I_{uv}} (R_{ui} - \bar{R}_u)^2} \sqrt{\sum_{i \in I_{uv}} (R_{vi} - \bar{R}_v)^2}} \quad (4)$$

In the above formulas, the scored set of user  $u$  is denoted as  $I_u$ , and the scored set of user  $v$  is denoted as  $I_v$ .  $|I_u|$  represents the number of elements in user  $u$ 's rated set,  $|I_v|$  represents the number of elements in user  $v$ 's rated set, and  $I_{uv} = I_u \cap I_v$  represents the intersection of the two ratings.

### 2.3 Calculate User Scoring Bias

Data bias occurs at the user-to-data stage. The observed ratings are not a representative sample of all ratings, and ignoring the internal bias in the data can lead to many problems, as users prefer to rate particularly good or particularly poor resources [8]. The calculation formula is:

$$DEV_{uv} = \frac{\sum_{i \in I_{uv}} (R_{ui} - R_{vi})}{|I_{uv}|} \quad (5)$$



**2.4 Assign Score to Non Scored Resources**

There are many system resources, and users’ practices are inconsistent. There must be many users who use resources but do not score. The easiest way to do this is to uniformly set the user’s unrated resources to a fixed value rating. However, this approach cannot fundamentally solve the problem of reduced recommendation efficiency caused by data sparseness. The system predicts the user’s rating of the resource through the ratings in the similar resource set, thereby increasing the user’s rating rate for the resource, and promoting a higher common rating set among the users. The formula is:

$$R_{ui} = \frac{\sum_{i \in N(u)} |N(i) \cap N(j)| \times (R_{ij} - Dev_{ij})}{\sum_{i \in N(u)} |N(i) \cap N(j)|} \tag{6}$$

In the above formula,  $N(u)$  represents the resource rated by user  $u$ .

**2.5 Generate Final Recommendation Results**

According to the rating information of similar neighbors, the recommendation result is generated for the target user. The simplest prediction method is to average the ratings of the neighbors. The calculation formula is [9, 10]:

$$\tilde{R}_{ui} = \frac{1}{k} \sum_{v \in N} R_{vi} \tag{7}$$

In order to obtain a better recommendation effect, a weighted average method can be used.

The weighted average of the scores of the neighbors is directly performed, and the calculation formula is:

$$\tilde{R}_{ui} = \frac{\sum_{v \in N} sim(u, v) R_{vi}}{\sum_{v \in N} |sim(u, v)|} \tag{8}$$

Calculate the increment of user ratings first, and then calculate the weighted average. The calculation formula is:

$$\tilde{R}_{ui} = \bar{R}_u + \frac{\sum_{v \in N} sim(u, v) (R_{vi} - \bar{R}_v)}{\sum_{v \in N} |sim(u, v)|} \tag{9}$$

Considering the real-time nature of the recommendation and the length of the recommendation list, the Top-N recommendation list is finally generated according to the predicted score.

### 3 Functional Requirement Analysis of MOOC Teaching Resources System for Japanese Major

Using MOOC for related teaching can shorten the distance between related subjects, enhance the interaction between related subjects, and provide a more convenient platform for effective communication between teachers and students, which is conducive to the further consolidation and deepening of teaching content. Through the analysis of functional requirements, we can understand the actual needs of users for the teaching resource system, and lay the foundation for the development of a fully functional MOOC teaching resource system. The system consists of three types of users, each type of user is analyzed separately based on the use case diagram.

#### 3.1 Students User Functional Demand Analysis

After student user registration and approval by the system administrator, they can use the MOOC teaching resources. An example of student user functional demand analysis is shown in Fig. 2.

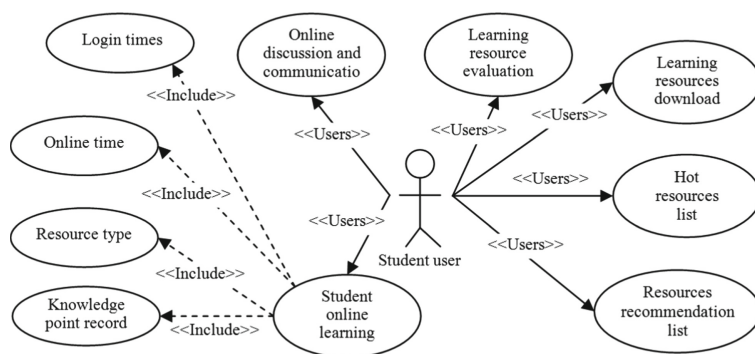


Fig. 2. Use case diagram of student user

For the student user use case diagram shown in Fig. 2, each function is briefly described as follows: First, the Resources recommendation list, a list of all resources recommended to the student user according to their interests. The second is the Hot resources list, the current hot resources recommended to student users. Third, Learning resources download, student users can download resources to the local computer for offline use. Fourth, online discussion and communication, in the process of using resources to learn, students can discuss and communicate to improve together. Fifth, Learning resource evaluation, student users evaluate resources in the process of using resources. Sixth, Student online learning, the core function of student users, the system records information such as “Login times, Online time, Resource type, Knowledge point record” of student users.

### 3.2 Teacher User Functional Demand Analysis

After the teacher user is registered and approved by the system administrator, he can guide the students. An example of teacher user functional demand analysis is shown in Fig. 3.

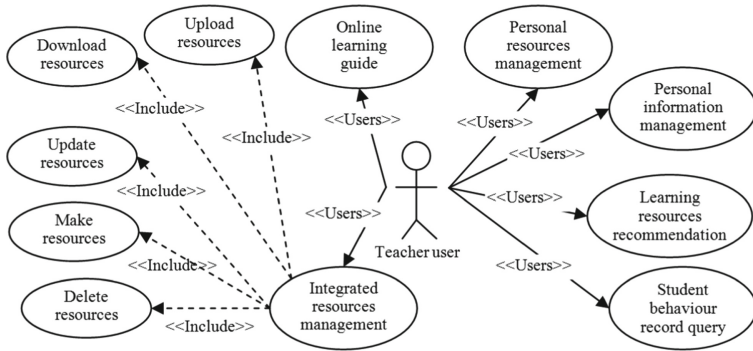


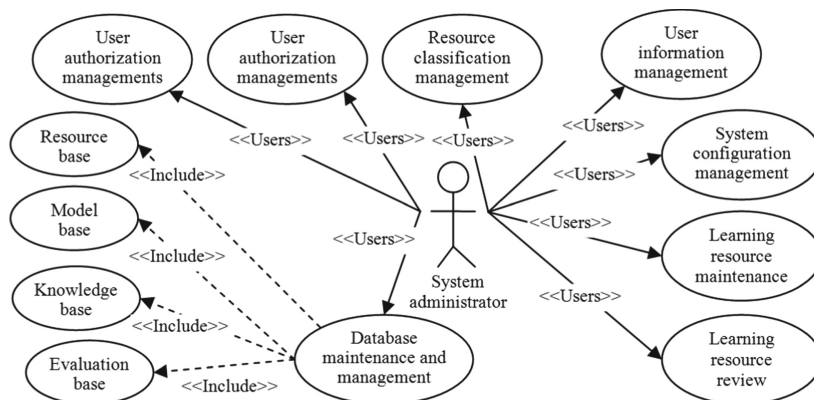
Fig. 3. Use case diagram of teacher user

For the teacher user use case diagram shown in Fig. 3, each function is briefly described as follows: First, Student behavior record query, the teacher understands the students' learning situation using MOOC resources, and improves the pertinence of the guidance. Second, with Learning resources recommendation, teacher users can recommend resources to student users based on their own knowledge and experience. The third is Personal information management, the maintenance of personal information by teacher users, including functions such as entry and update. Fourth, Personal resources management, teacher users can manage personal teaching resources, and can upload excellent personal resources to the system. Fifth, online learning guide, teachers guide students online to solve problems in the use of MOOC teaching resources. Sixth, Integrated resources management, which is the core function of teacher users, including "Upload resources, Download resources, Update resources, Make resources and Delete resources" and other functions.

### 3.3 System Administrator Functional Demand Analysis

The system administrator user has the highest authority and can directly enter the system without registration and auditing. The system administrator user functional requirement analysis use example is shown in Fig. 4.

For the system administrator use case diagram shown in Fig. 4, each function is briefly described below. First, Learning resource review, in order to improve system security, MOOC teaching resources uploaded by teacher users can only be used by student users after being reviewed by the system administrator. Second, Learning resource maintenance, update and modification of MOOC teaching resources, etc. The third is System configuration management. In order to improve the compatibility and portability of the system, configuration management is required. The fourth is User information



**Fig. 4.** Use case diagram of system administrator

management, which manages the registration information of teacher users and student users. The fifth is Resource classification management, which manages the categories of MOOC teaching resources. Sixth, User authorization management, different user roles have different permissions to authorize teacher users and student users. Seventh, QR code generation is used to generate QR codes for teaching resources, as well as QR codes for teacher users and student users to log in, which is convenient for users to use. Eighth, Database maintenance and management, which is the core function of system management, specifically includes functions such as “Resource base, Model base, Knowledge base and Evaluation base”.

## 4 Conclusions

The research results of this paper are the basic work of developing a teaching resource recommendation system, which is in line with the development direction of education in the new era, and can promote the application of the MOOC-based blended teaching model for Japanese major, and achieve the goal of talent training for Japanese major. In future research and practice, it is necessary to further use data mining technology and cluster analysis technology to effectively alleviate the impact of the various problem and improve the accuracy and intelligence of personalized teaching resource recommendation.

## References

1. Guan, F., Gao, Y.H., Lin, Q.: Empirical research on intelligent recommendation method of printed resources in university libraries based on collaborative filtering. *Inf. Res.* **34**(4), 109–115 (2021)
2. Mei, Y., Wang, J.R., Yu, J., et al.: Online and offline teaching practices based on MOOC. *Comput. Educ.* **19**(3), 1–5 (2021)
3. Lin, S.N.: Research on hybrid Japanese teaching from the perspective of “Internet +.” *Educ. Teach. Forum* **13**(37), 125–128 (2022)

4. Wei, H.Y.: Research on the reform measures of online and offline Japanese teaching mode under the background of “Internet +”. *Off. Informatization* **27**(13), 17–19+9 (2022)
5. Wu, J.F., Zeng, S.P., Zheng, L.: Research on user-based collaborative filtering recommendation Algorithm. *Mod. Comput.* **34**(19), 27–29+67 (2020)
6. Qiu, D.P.: Application of collaborative filtering recommendation algorithm based on user behavior in video recommendation. *Changjiang Inf. Commun.* **35**(2), 118–120 (2022)
7. Liu, J., Yang, J., Song, S.S.: Collaborative filtering recommendation algorithm based on purchasing intention of users. *J. Jilin Univ. (Sci. Ed.)* **59**(6), 1432–1438 (2021)
8. Liu, J.J.: Improvement and implementation on personalized book recommendation algorithm based on user collaborative filtering. *J. Libr. Inf. Sci.* **6**(1), 38–42 (2021)
9. Lei, K.W., Wang, X.H.: Design of personalized recommendation system for scenic spots based on WeChat public platform. *Inf. Technol.* **46**(1), 56–61 (2022)
10. Wei, H., Liu, X.Y., Zhang, W.: A collaborative filtering recommendation algorithm of the improved user similarity. *Electron. Des. Eng.* **29**(17), 30–34 (2021)



# A Study of Teaching Resource Retrieval System Based on Semantic Web

Weiwei Zeng<sup>1</sup>(✉) and Shirui Peng<sup>2</sup>

<sup>1</sup> College of Foreign Languages and Literature, Wuhan Donghu University, Wuhan, Hubei, China

zengweiwei@wdu.edu.cn

<sup>2</sup> College of Media and Art Design, Wuhan Donghu University, Wuhan, Hubei, China

**Abstract.** This paper starts from the architecture of the Semantic Web, and studies the core technologies of the Semantic Web, such as RDF/RDFS, OWL and SPARQL, from the perspective of implementation. Secondly, the theoretical basis and related technologies of the Semantic Web are studied and analyzed from the aspects of the Semantic Web's functions and hierarchical composition, grid technology, ontology technology, etc., especially its knowledge organization, knowledge management, knowledge Advantages in discovery and knowledge retrieval. Finally, the research analysis teaching resource management related technologies under the Semantic Web framework, it provides theoretical and technical guarantees for the next research work.

**Keywords:** Semantic Web · Semantic Web Technology · Teaching Resources Construction

## 1 Introduction

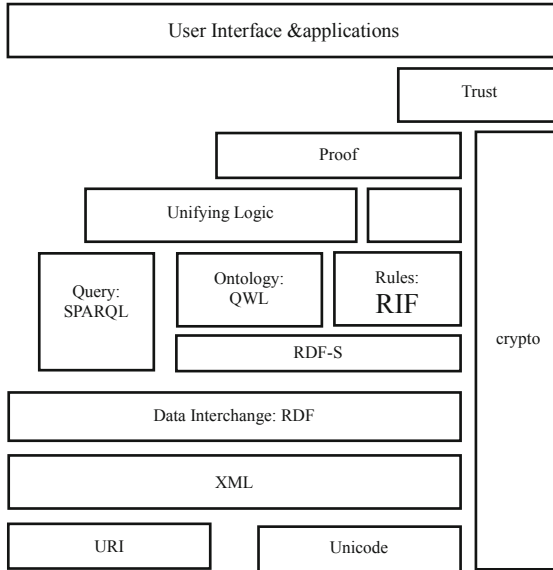
Network teaching resources are the premise and basis for carrying out network education. With the development of Web technology and network information technology, network teaching resources are becoming more and more abundant. There are various types of educational resources and rich forms of expression, which provide learners with a wide range of choices and broad development space.

## 2 Concept of Semantic Web

BERNERS-LEET believes that the Semantic Web defines the meaning of information on the Web on the basis of the current Web, and enables humans, machines, and machines to work better together on the basis of machines understanding their meanings [1]. After the concept of the Semantic Web came out, the “W3C Semantic Web Activity” organization was also established. The organization has long been committed to the promotion of the Semantic Web in the World Wide Web Consortium (W3C) and formulated related standards for it. At present, the research related to it is developing rapidly, and it has become an important research field in the field of computer science.

## 2.1 Architecture of the Semantic Web

The architecture of the Semantic Web was proposed by BERNERS-LEET, as shown in Fig. 1.



**Fig. 1.** Architecture of the Semantic Web

In this architecture is mainly used to uniformly encode information on the Web, while URI is used to identify resources and objects. The second layer is the grammar layer. Extensible Markup Language (XML) can not only be used as the metalanguage of markup language, but also as the description grammar of Semantic Web [2]. The resource description framework (Resource Description Framework, RDF) of the third layer is a data model that describes Web resources and their relationships, expressed in XML syntax. The fourth layer is the ontology (Ontology) layer, including RDF Schema, Ontology Web Language (Ontology Web Language, OWL), query language and data acquisition protocol (Protocol and RDF Query Language, SPARQL) and rule exchange format (Rule Interchange Format, RIF) and so on. RDF Schema provides a vocabulary for describing RDF resources, OWL is a network ontology description language based on description logic, SPARQL is a language for querying RDF data, and RIF is used to describe the rule exchange format [3]. Layers 5 through 7, in turn, also cover logic, validation, credit, and user interface and applications. Each layer on the ontology layer mainly performs operations such as logical reasoning and rule query based on the functions provided by the lower layer. At present, W3C is still improving the architecture of the Semantic Web, and relevant technical specifications and standards are also being formulated.

## 2.2 Core Technology of Semantic Web

In the semantic Web architecture, the underlying technologies are all hypertext web technologies, including HTTP, Unicode, XML, XML Name space, etc. The middle layer mainly uses the semantic Web core technology promulgated by W3C, including RDF, RDFS, OWL and SPARQL [4]. In addition, there are some Semantic Web -related technologies, such as SKOS, RIF, and GRDDL. At present, the core technologies (RDF, RDFS, OWL and SPARQL, etc.) to realize the Semantic Web have been standardized by W3C. Although these technologies are not enough to build the Semantic Web, they can be used to build some semantic applications.

### (1) XML/XMLS technology

XML does not have a fixed tag set, and users can customize the tag set through XML schemas. Document Type Definition (Document Type Definition, DTD) and XMLS (XML Schema) are two commonly used XML schemas, but DTD is relatively old and has many restrictions. XMLS is a substitute for XML-based DTD, supports data types and name spaces, and can be expanded for future needs, so XMLS is widely used.

### (2) RDF/RDFS technology

RDF describes resource information on the World Wide Web, such as describing information about web pages, including content, author, and date of creation and modification. RDF includes RDF Data Model, RDF Schema and RDF Syntax.

## 2.3 Ontology and Its Description Language

RDF is based on the XML standard syntax when describing resource objects, and can express certain semantics, so the machine can understand the description information. However, due to the limitations of the expressive ability of RDF(S), the relationship between words cannot be expressed. In order to further improve the description ability of Web information resources, the concept of Ontology originally belonging to the field of philosophy is applied to the field of computer science, especially the field of Semantic Web, which is very suitable for describing concepts and the relationship between concepts.

Ontology-based knowledge visualization.

“The systematic description of objective existence in the world, that is, ontology” is the original philosophical concept of Ontology. On the Semantic Web, it is the purpose of the ontology to enable information and applications of different development tools and application platforms to communicate, share and reuse without redesigning, and it overcomes the problem that the new knowledge system cannot effectively use the existing knowledge system [5]. Defects, thus saving a lot of manpower, material and financial resources. The “knowledge acquisition” and “knowledge representation” methods need to be readjusted in the network information age. It is precisely this need that people use ontology to organize information resources and knowledge. “Resources” and “knowledge” in the information age are based on ontology Theories can be dynamically understood, so as to obtain an ability to adapt to the network information age. Ontology is a very complex and abstract concept. It is an important part of the Semantic



Web, and it is also the key core technology and cornerstone for realizing the Semantic Web.

On the basis of scientific computing visualization, data visualization, and information visualization, a new research field of knowledge visualization (Knowledge Visualization) has been developed. Literally, through knowledge visualization, students can acquire knowledge or understand knowledge in a self-learning manner [6]. It is a way to show the knowledge structure of experts to students. The knowledge visualization researched in this paper refers to the use of graphic visualization to organize the conceptual knowledge in the network professional course resources, establish the mutual connection between knowledge and knowledge, and let students understand the knowledge system of the course relatively intuitively. Thereby improving the learning effect. Figure 2 is an ontology-based knowledge visualization model.

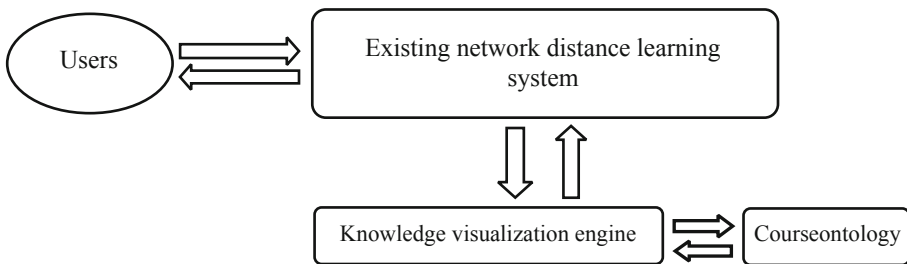


Fig. 2. Ontology-based knowledge visualization model

The fundamental purpose of Ontology is to enable communication, sharing and reuse of information and applications of different development tools and application platforms. Acquiring knowledge or understanding knowledge through self-directed learning.

### 3 Functions and Contents of the Teaching Resource Library

#### 3.1 Functions of the Network Teaching Resource Library

Generally speaking, nine types of teaching resources, including media materials, test questions, test papers, cases, courseware, FAQs, resource catalogue indexes and online courses, are the main content of teaching resource construction. The construction of educational resources not only needs to form a resource warehouse with a large amount of data, but also has a strong support and service function for education and teaching. These functions have a major impact on the reform of the teaching structure itself, whether it is teachers' lesson preparation, teaching, or There will be corresponding fundamental changes in student learning, including the management of academic staff.

##### (1) Resource sharing

Resource sharing is the most obvious service function of educational resource construction. Whether it is at the national or regional level, the organization and construction of educational and teaching resources must achieve resource sharing within the region.

(2) Electronic lesson preparation

An important service function of educational resource construction is that it can provide a large amount of reference information for teachers' lesson preparation and realize a constructive teaching.

(3) Resource-based learning model

The resource library provides learners with extremely rich learning resources, including digital libraries, electronic reading rooms, online newspapers and various databases, multimedia e-books, etc.

(4) Knowledge Accumulation

In a broad sense, an educational resource bank is a knowledge base, and the development and construction of resources is a process of knowledge accumulation.

(5) Data Management

Educational resources include not only learning resources, but also management and decision-making resources. Therefore, the development and construction of educational resources not only provide objects that can be learned and recognized, but also provide information for managing and coordinating various educational activities. Data management refers to the collection, transmission, storage, processing, maintenance and analysis of educational management and educational statistical information [7]. After digitizing various information in educational units, it is organized and classified and applied to all aspects of teaching and management.

### 3.2 Main Contents of Network Education Resource Database Construction

The construction of educational resource database mainly includes the following four levels:

- (1) Educational and teaching resources of media materials, such as: text materials, literature, test question bank, test papers, courseware, teaching cases and educational resource indexes.
- (2) Network courses formed in accordance with the comprehensive integration of disciplines, professional knowledge systems and network teaching requirements;
- (3) Form excellent educational resources through evaluation;
- (4) Development of Web - based education and teaching resource management system.

Among the four levels of network education resources construction, material teaching resources and network course education resources are the focus and core that need to be standardized in the construction, and they are the basis for the construction of network education resources. In order to evaluate and screen high-quality educational resources, the evaluation criteria must be standardized. In addition, due to the ever-changing content and various forms of online educational resources, the effective management and retrieval of online educational resources is the basic guarantee for making full use of online teaching resources [8].

## 4 Research on Related Technologies of Teaching Resource System Under Semantic Web Architecture

### 4.1 Educational Resource Library Architecture

The online education and teaching resource library provides materials, teaching software, supplementary materials and so on for students and teachers in the process of learning and teaching. Its most important feature is its strong pertinence. It is a structured resource library specially designed for learning and teaching. The education resource library system includes media materials, test questions, courseware and network courseware, cases, literature, frequently asked questions, resource directory index and network courses, as well as the research and development of teaching support systems suitable for various teaching modes and modern distance education management systems. The first problem to be solved in constructing the educational information resource library is the content system of the library. Collecting, organizing, processing, and organizing educational information to construct an orderly structure is the starting point for establishing the content system of the educational information resource library.

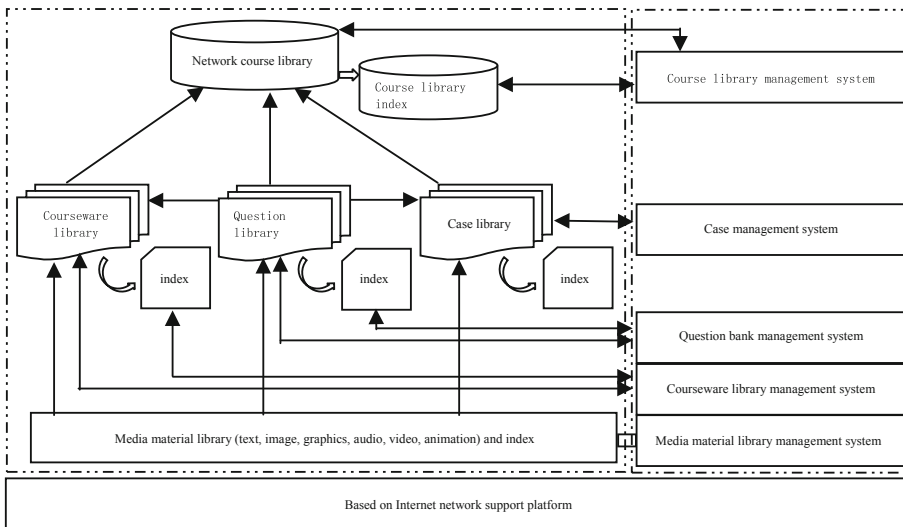


Fig. 3. Architecture of Educational Resource Database

As shown in Fig. 3. Taking the Internet network support environment as the platform is the environment for the construction, implementation and operation of modern distance education resources. In the whole resource, the media material library is the most basic, and the resources or data in other resource libraries may come from the data in the media material library. The network course library is composed of cases, course ware and corresponding. The course ware and test questions of the teaching link are comprehensively formed.

Within the educational resource system, the collection, exchange, and exchange of educational information resources are usually complicated and disorderly, and there are

serious regional differences in the storage of information resources, which greatly weakens the sharing benefits of information resources. To build an educational information resource bank is to make the collection, circulation, and exchange of information orderly, so as to achieve the purpose of rational allocation of resources, efficient development, and benefit sharing.

## 4.2 Organization of Teaching Resources and Meta Data Representation

At present, some LMSs and teaching websites make the teaching content fixed in the course ware. If the teaching content is not properly packaged and separated from the teaching platform, it will be quite time-consuming and laborious to expand new teaching content. Even if the course has been objectified, if the design of appropriate meta-data for the teaching content is not considered, it is impossible to further use information retrieval (Information Retrieval, Agents and other technologies to improve the ease of course objects to be searched, The value of being reused. That is to say, these learning content or software (LMS) based on E-learning do not have a unified mechanism in terms of acquisition, query, delivery, sharing, and reuse [9]. Therefore, there must be an effective description, The method of organizing and expanding teaching resources, that is, establishing a general resource framework and hierarchical structure description method for teaching resources. At present, the typical international standard systems include ADL-SCORM, IMS, AICC-AGR, etc. Among them, the most influential ones are ADL-SCORM.

## 4.3 Analysis of Knowledge Base Teaching Resources Construction

Knowledge Base is a structured, easy-to-operate, easy-to-use, comprehensive and organized knowledge cluster in knowledge engineering. The knowledge base system is a development tool and environment that provides support for users to develop knowledge-based systems. In summary, the existing knowledge base system should have the following characteristics. Of course, due to the inherent characteristics of each system, there may also be different processing technologies or a different function:

- (1) It should have the function of accessing and managing knowledge.
- (2) It should have the function of efficient access to large quantities of data, including functions such as data sharing, concurrency control, data security, fault recovery, etc., and support all functions of the database.
- (3) It has a reasoning mechanism, and the reasoning mechanism completes the processing of knowledge.
- (4) It has proper storage and management functions of rules and facts.

In the subject professional knowledge resource base established on the basis of the domain knowledge base in ontology, what kind of architecture and what kind of system should be adopted for the integration, integration, organization and management of large-scale, multi-type and multi-level knowledge resource content? What about strategy? This paper adopts the design pattern of "three-layered knowledge network" to construct the frame structure of knowledge resource base.

**Basic information base:** The basic information base consists of source information and edited data for retrieval, such as various multimedia databases, course ware databases, and numerical databases composed of scientific data established according to the knowledge classification system and media classification system [10].

**Knowledge base:** Knowledge base is a database composed of knowledge elements. The so-called knowledge element refers to a knowledge unit or knowledge element with independent meaning, which is formed by excavating and refining various resources of the subject professional knowledge base, including three types of basic knowledge elements including theory and method, fact, and value. Among them, the theoretical and method-based and fact-based knowledge bases are classified according to subject knowledge systems, and the numerical knowledge bases are classified according to data types.

**Knowledge warehouse:** knowledge warehouse is the basis for realizing knowledge management and knowledge service. The construction of the knowledge warehouse should take the knowledge element as the basic element and the knowledge element link as the hub; its construction generally goes through the knowledge element extraction and indexing of document information, the construction of the knowledge element library, and finally the construction of the knowledge warehouse.

The knowledge association methods of different spans are used between the above layers of databases to build a multi-dimensional knowledge network, but at the same time they can be used independently. In addition, the international advanced universal data structure standard is used to build the knowledge base, which enables the simultaneous use of three-tier databases on the unified retrieval platform. In this way, the three-tier database finally forms a super-functional and super-large-scale knowledge resource system through knowledge association methods and a unified retrieval platform.

## 5 Epilogue

Fully realizing the Semantic Web requires large-scale adoption of ontology-based approaches to sharing information and resources. Information resources can be marked with ontology representation language, semantic features can be recognized by the knowledge organization system, and the semantic information obtained by it is stored in the knowledge base. The teaching resource system model oriented to the Semantic Web has teaching content covering multiple disciplines and networks. Knowledge and other characteristics, fully demonstrate the learning characteristics of the information age, provide learners with newer knowledge and broader knowledge, and make the teaching process more personalized and improve the quality of teaching.

**Acknowledgments.** This paper is supported by the Hubei Province Educational Science Planning Project under grant 2019GB126. Thanks for the support of the Education Department of Hubei Province.

## References

1. Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web in Scientific American (2001)
2. Pascal Hitzler, I.: A review of the semantic web field. *Commun. ACM* **64**(2), 76–83 (2021)
3. Vidyamv, P.: A survey on ontology tools. *Int. J. Sci. Eng. Res.* **3**(10), 987–994 (2012)
4. Wade, B.: Was humpty dumpty right? An exploratory study of semantic priming and lexical access. *Educ. Psychol.* **4**(8), 98–99 (1995)
5. Berners-Lee, T., Hendler, J., Lassila, O.: the semantic web. *Sci. Am.* (5), 443 (2001)
6. Roussey, C., Calabretto, S., Pinon, J.M.: A Multilingual Information System Based on Knowledge Representation, 98–111 (2001)
7. Koper, R.: Use of the semantic web to solve some basic problems in education: increase flexible, distributed lifelong learning, decrease teacher's workload. *J. Interact. Media Educ.* **6**, 1365–1371 (2004)
8. Paulsson, F., Naeve, A.: Virtual work space environment: a taxonomy and service oriented architecture (SOA) framework for modularized virtual learning environments (VLE)-applying the learning object concept to the VLE. *Int. J. E-Learn.* **5**(1), 45–57 (2006)
9. Naeve, A.: The human semantic web-shifting from knowledge push to knowledge pull. *Int. J. Semant. Web Inf. Syst.* **1**(3), 1–30 (2005)
10. Guarino, N., Masolo, C., Vetere, G.: Onto seek: content-based access to the web. *IEEE Intell. Syst.* **14**(3), 70–80 (1999)



# Construction of Indoor Space Atmosphere by UVLED Light Source Design

Yu Gui<sup>1</sup>, Yangyang Fan<sup>1</sup>, Jian Tian<sup>2</sup>, Zonghui Bo<sup>1</sup>, Yuxin Miao<sup>1</sup>, Yuxiao Fan<sup>1</sup>,  
Shunzi Ren<sup>1</sup>, and Xi Chen<sup>1</sup>(✉)

<sup>1</sup> School of Horticulture and Landscape Architecture, Tianjin Agricultural University, Tianjin,  
China

36361950@qq.com

<sup>2</sup> Tianjin Urban Renewal Construction Development Co., Ltd., Tianjin, China

**Abstract.** In modern living environment, lighting is no longer simply used as a lighting tool, but also plays an important role in creating space atmosphere. The reasonable design of the color, shape and position of the light source can effectively improve the indoor living experience. In this paper, the UVLED light source was used to shape the indoor space and environment. According to the three-step method of calculating illuminance requirements, completing lighting planning and improving indoor atmosphere, the interior lighting design of Vanke Guancheng Villa-Dajia Real Estate was carried out, and the UVLED lighting was compatible with the scheme of those audio-visual, rest, dining and reception scenes. The atmosphere creation was also analyzed and studied to explore the significance of light source design, which could be the reference experience for scholars who follow-up study to create indoor space atmosphere through lighting design.

**Keywords:** UVLED Light Source · Space Atmosphere · Light Source Design

## 1 Introduction

### 1.1 Research Background

In recent years, with the gradual improvement of people's living environment, the demand for spiritual life is getting higher and higher. Among them, people have newer and higher requirements and pursuits for indoor space environment. People have higher and higher requirements on the environment of necessary places for living, and the indoor space should meet the requirements of living, working, studying and entertainment. People have higher requirements for the living conditions and living environment of indoor space. With different needs of people, people need to meet not only material needs but also spiritual needs. The creation of indoor space atmosphere is one of the manifestations of people's demand for environmental diversity. Good atmosphere of interior design can not only create personalized artistic characteristics, but also improve the sense of space atmosphere, convey beautiful visual feelings to people, enhance people's living environment and feel the unique charm of space.

Light is an important natural element in life and plays an important role in the creation of atmosphere. The modern American architect Louis once said: "Light gives life to the characteristics of space and architecture". However, in interior space design, most people stay on the level of practical function of light source, which has a good role in creating indoor atmosphere. In recent years, the human-centered design concept has gained popularity, and the same human-centered design concept is needed in the design of light sources to maximize the charm of light, release the vitality of light, and bring people a relaxed and pleasant spiritual enjoyment.

## 1.2 Concept of UVLED Light Source

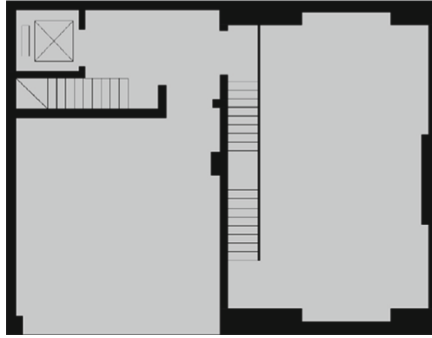
UVLED lights to smaller size, can be flexible custom combination of decorative, a combination of small lamp beads, to do from the point to expand to the surface of the light source irradiation. UVLED lights have fewer connecting wires, so that you can reduce the damage to the lamp strip beads in the interior design, and UV LED lights can be recycled twice. UVLED lights compared with other light sources, UVLED light source with a longer UVLED and LED luminous rate of the same case, UVLED light source to heat, power and other advantages to achieve a higher level of safety and environmental protection, eye protection is also better effect. UVLED light source can be achieved from the point to the expansion of the surface, the emission of light source uniform, soft, flexible combination, can be personalized Light source, to achieve the effect of beautifying the indoor environment.

## 2 Project Summary

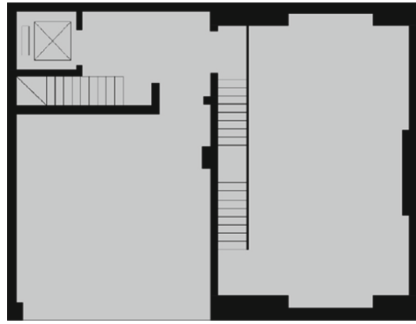
The design project is Vanke Guancheng Villa - Everyone, located in the central villa area of Shunyi District, Beijing, with an area of 300 square meters, five floors, including two underground floors and three above ground floors, which is a villa residence built by Vanke Group with the concept of "inheritance and innovation".

As shown in Figs. 1, 2, 3, 4 and 5, the original house plan of the villa, the interior space needs lighting distribution, lighting atmosphere design and other lighting design, the main lighting transformation area for the first and second basement space, the rest is part of the transformation space. For example, the recreation area next to the garage on the second basement level needs sufficient illumination to meet the entertainment demand; the audio-visual area and reading area on the first basement level need to set reasonable lighting according to the space atmosphere demand; in the living room area on the first floor, lighting is set according to the reception demand.





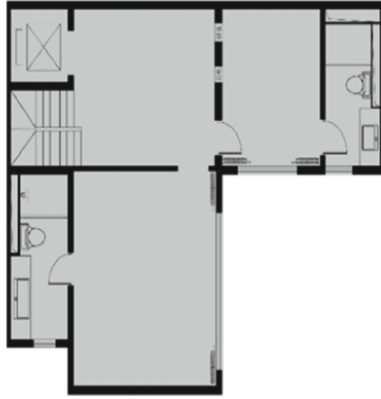
**Fig. 1.** Ground floor plan of the second floor



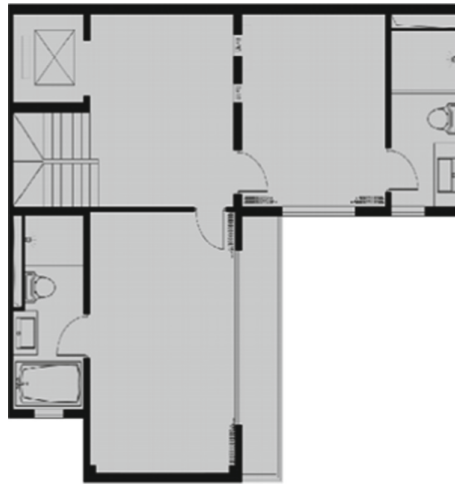
**Fig. 2.** Ground floor plan



**Fig. 3.** First floor plan



**Fig. 4.** Second floor plan



**Fig. 5.** Three-layer floor plan

In order to meet the functional requirements of each interior space and create the required atmosphere, the design of UVLED interior light source is designed for the garage area on the second floor, the audio-visual, reading and children's activity area on the first floor, the living room and dining room area on the first floor, the bedroom area for the elderly and girls on the second floor, and the master bedroom and children's room on the third floor.

### **3 Analysis of the Elements of UVLED Light Source Design**

#### **3.1 Calculating Illumination Requirements**

First, the UVLED lamp irradiation range, irradiation brightness, irradiation function of the actual effect will be affected by a number of factors, in the interior design of

the building's structure, functional requirements, wall and floor color, etc. are factors that should be taken into account when calculating the illumination demand, need to complete the UVLED lamp illumination demand value setting according to different scenes. In determining the UVLED lamp illuminance requirements, the most common method is to use the coefficient method to analyze the baseline values for the application of architectural scene lighting systems, and then consider the intensity of the light source in the scene, the reflection or refraction of light caused by the indoor environment and wall pollution, and to comprehensively determine the lighting effects of UVLED lamp sources in different locations and with different properties to ensure the reasonableness of the lighting settings to meet the interior lighting requirements. Purpose [1].

### **3.2 Perform Lighting Planning**

UVLED lamps and lanterns of various colors and adjustable shapes, but due to the strong variability of the product, the different technical composition and transport forms between different manufacturers, have caused the phenomenon of uneven product quality and performance, so in the lighting planning, the need to combine the actual lighting scene, the total power of other types of light sources to measure the statistics, and then based on the UVLED lamp settings demand, complete Indoor space all light source applications, to ensure that UVLED lighting design can meet the needs of the scene [2]. In the process of lighting planning, the brightness and color rendering index of the UVLED lamps used need to be considered comprehensively, and through multiple settings, both to protect the lighting needs and to meet the environmental protection and energy saving.

### **3.3 Enhances Indoor Ambiance**

Good lighting design can bring people a comfortable experience, improve the fashionability, beauty and fun of the whole space, and has an important significance to the creation of the atmosphere of the space [3]. Before the lighting design, the architectural structure, space color and other factors are taken into account, which can better complete the design of UVLED light source. In the space that needs bright atmosphere, the use of ordinary incandescent lamps and adjustable UVLED combination, can meet different brightness needs at different times and environments [4]; in the space that needs a soft atmosphere to apply chandeliers, color rendering and warm light UVLED combination, can make the space soft and natural, enhance the sense of environmental experience; in a warm and quiet atmosphere, it should avoid setting a high brightness light source that shines directly into the eyes. Adopt UVLEDs with higher color rendering, lower brightness, and shading or soft light treatment, with multi-control switches or voice-activated switches to enhance the use experience [5].

## **4 UVLED Light Source Design in the Specific Application of Indoor Space Atmosphere**

UVLED light source is small in size, small footprint, and can be flexibly applied to architectural design. Combining interior space with light source design, reasonable design

can achieve the effect of hidden lighting. The overall light source hue is the main color of indoor light, and the light source design can be carried out based on the three-step method of calculating illumination demand, completing lighting planning, and enhancing indoor atmosphere [6].



**Fig. 6.** Ground floor audio-visual scene lighting effects



**Fig. 7.** Ground floor audio-visual scene lighting with atmosphere effect

#### **4.1 Video and Audio Scenes Create a Romantic Atmosphere**

For the audio-visual scene, the demand for illumination is low, only a small number of lighting sources need to be set, and the location of the lighting source should also be set away from the audio-visual screen, not directly on the human eye, considering that most of this scene is a joint family viewing, the warm atmosphere is more suitable, so in such scenes for UVLED light source design, should pay attention to the light source set in the back of the sofa or wall compartments and other dark places that will not directly illuminate the entire space, using warm colors with low color rendering index, as well as high lighting uniformity and low illumination UVLED light source, to create a romantic atmosphere for the scene and enhance the effect of the viewing process [7]. (shown in Figs. 6 and 7).

## 4.2 Resting Space to Create a Tranquil Atmosphere

For the rest space, most of them belong to private space, the demand for illumination is low, and the light source design of the rest space can use warm and practical light sources [8]. On the one hand, in the arrangement of the light can not have a sense of oppression, should try to avoid direct light source of high brightness in the rest area, and use more irradiation to the walls, roof, confined areas of the panel light source or light belt; on the other hand, in the arrangement of the light can be appropriate in the end of the bed, sofa, bedside table and other areas to increase the creation of a sense of romance of the small light source, so that the whole space is more quiet. (shown in Fig. 8).



Fig. 8. Second floor master bedroom lighting effect

## 4.3 Dining Space to Create a Warm Atmosphere

This space needs to meet the needs of family gathering and communication, good lighting design has an important role in creating space atmosphere, dining space atmosphere is created from the color matching and lighting effects to consider two aspects: on the one hand, the use of ceiling-mounted high illumination UVLED to enhance the lighting needs of the whole scene; on the other hand, the use of warm color low color rendering index and low illumination UVLED strip light source to enhance the scene of warm atmosphere [9]. The combination of main lighting and auxiliary lighting with reasonable collocation makes the whole space atmosphere warm while bringing lamp changes and light levels to the whole space [11].

## 4.4 Living Room to Create a Welcoming Atmosphere

The living room space has more requirements for UVLED light source, one is to create a comfortable and intimate atmosphere for reception to enhance the reception experience; the second is to meet the higher illumination demand of the living room space, reduce the visual burden of personnel in the environment, maintain a good mood, and ensure the quality of reception; the third is to create an atmosphere of trust in the scene, reduce the wariness of the scene guests, and achieve the purpose of enhancing the reception experience. Therefore, it is a more feasible solution to apply UVLED light source with high illumination, cool color, high color rendering index, strong concentrated lighting performance, and no dazzle hazard in this scene [11].

## 5 Conclusions

Indoor light source design is an important link in the creation of indoor space atmosphere, to meet the “people-oriented” design concept, the use of UVLED light source to shape the indoor space and environment, and analysis and research on the creation of the atmosphere, through the design of light sources, can achieve a combination of practicality and artistry. The design of indoor light source will interconnect and integrate various factors to create space atmosphere, visually and psychologically bring different feelings, or immersion, or pleasure, or beauty, or tranquillity.

In this paper, we explore the UVLED light source to create atmosphere for indoor space atmosphere, analyze the elements of UVLED light source design, and meet people’s needs for light in different space activity scenes by combining with different light source modes. Through the design, it is found that UVLED light source has a vital role in creating the atmosphere of indoor space and people’s psychological feelings.

## References

1. Yu, S.: Research on shaping the mood of indoor lighting environment. *Beauty Times:Creativity* (on) (1), 3 (2019)
2. Xu, G., Zhu, R.: The use of light and shadow in display design. *Beauty Times: City* (3), 1 (2014)
3. Li, X.: The role of lighting in interior design. *Doors windows* (9), 1 (2017)
4. Long, B.: Lighting up living space with light - a discussion on interior lighting design. *Footwear Craft Des.* (20), 3 (2021)
5. Dong, J., Huang, M., Guo, Z., et al.: Dot design and simulation based on a circular LED surface light source. *Appl. Light Source* **31**(05), 701–704 (2010)
6. Gong, S.: Research on the Development Factors of Contemporary Interior Lighting Design. Southeast University, Nanjing (2004)
7. Liu, Y.: *The Language of Light—Analysis of Light Environment Design in Interior Space*. Northern Polytechnic University, Beijing (2015)
8. Quan, L.H.: On the importance of interior lighting design. *Da Guan: Tan* (8), 2 (2019)
9. Teng, Y.: On the importance of interior lighting design. *Art Sci. Technol.* (4), 1 (2019)
10. Gao, Y.: The art of lighting in interior decoration. *Art Appreciation* (21), 2 (2018)
11. Xiao, Y., Lu, Z.: Interior light source design for the creation of space atmosphere. *Light Sources Lighting* (12), 3 (2021)



# Seismic Stability Assessment of Civil Building Projects Based on BIM Technology

Rongguo Zhao<sup>(✉)</sup>

Shandong Institute of Commerce and Technology, Jinan, Shandong, China  
w1727502128@qq.com

**Abstract.** Building Information Modelling (BIM) technology is in a rapid stage of development in China. BIM technology is the use of computer information means to simulate information on the actual characteristics of a building, keeping information updated and shared throughout the life cycle of a project, and promoting efficient operation of project management. The main objective of this paper is to carry out research on the assessment of seismic stability of civil building projects based on BIM technology. To accelerate the development of BIM technology application and give full play to the advantages of its features such as full information, real-time, visibility, virtualisation and sharing, to enhance the comprehensive management technology level of pre-planning, submission, design, construction and operation and maintenance of construction projects, saving materials, saving time, reducing investment, improving quality and increasing efficiency, among other indicators. Further improve the whole information data of construction projects, and promote the process information of construction projects to be transparent, true, reliable and other safety performance, as the basis data basis for market behaviour regulation and integrity of the building intelligence sector.

**Keywords:** BIM Technology · Civil Construction · Building Engineering · Seismic Assessment

## 1 Introduction

Engineering design is the key to engineering construction, and the development and application of computer design software has greatly improved the efficiency of professional designers in architecture, structure, electrical and HVAC by changing the way of drawing from cardboard to computer. The emergence of complex projects such as super high-rise and commercial complexes, the lack of information and contradictory problems in design technology, construction management, operation and maintenance are increasing [1, 2].

In a related study, Achmad et al. proposed a design procedure based on displacement damage and implemented by the computer-aided tool DAMPERS (Damping Protection of Earthquake Resistant Structures) [3]. The residual capacity that a building may experience when subjected to multiple earthquakes during its nominal life is considered. A

promising approach to modelling building resistance during earthquake recovery has been proposed by Pernille et al. [4]. The Damage Index Model (DIM) can be applied interactively in a Geographic Information System (GIS) in order to calculate parameters and analyse data for quality management purposes. It allows easy access to data to assess the seismic mitigation and preparedness of buildings.

Concrete structure civil building engineering is a key area of concern in both design and construction. The constructability of the reinforcement arrangement of the civil building works, the pouring of concrete in the narrow space of the nodes and the choice of the node formwork support scheme in the design directly affect the construction quality of the civil building works and deviate from the force calculation results of the nodes. In this paper, constructability analysis under BIM model is carried out for civil building projects to refine the analysis and optimise the structural design. The on-site application of AR technology for complex civil building projects can achieve the demand for efficient functions such as guiding construction, assisting acceptance and real-time information management.

## 2 Design Studies

### 2.1 Characteristics of BIM

- (1) Visualization: BIM technology is able to show the construction condition in the form of three-dimensional diagrams, while information on all aspects of construction can be seen, providing a platform for communication, coordination and exchange between the various parties involved in the project, where they can communicate, change and provide a basis for the progress of the project [5, 6].
- (2) Simultiveness: BIM can carry out relevant simulations at multiple stages of the project, which is more convenient for the participants to understand the characteristics of the project, predict the possible risks in advance, and prepare for risk prevention and control, so as to achieve higher efficiency, cost saving and shorter construction period [7, 8].
- (3) Optimality: The optimisation of BIM is mainly reflected in two aspects: programme optimisation and design optimisation: programme optimisation is mainly to combine the construction organisation and design programme with the management of schedule, quality and cost, so as to select the most suitable programme.
- (4) Chartability: In the project construction, the technical delivery work is usually expressed in the drawings, BIM software through the establishment of the project's three-dimensional model, you can clearly see the process of the relevant technical indicators, and through the software's powerful simulation, composition, calculation function, make for the relevant technical indicators to modify more convenient, save the technical personnel due to change, check, review the contents of the drawings, a lot of time and energy spent.

### 2.2 Feasibility Analysis

- (1) Visual simulation of the building based on BIM and its associated software. The overall posture and seismic capacity of a building can be visualised by inputting



seismic waves. Simulations of the seismic capacity of existing buildings can be carried out before the earthquake to understand the seismic capacity of the building [9, 10].

- (2) BIM technology has good information sharing capability [11, 12]. Different models are interconnected and influenced by model components and semantic rules. Good information sharing is the guarantee of the reliability, flexibility and wide application of BIM technology. Information sharing with the help of BIM technology enables communication between various departments and links in building construction, ensures the construction quality of the building and improves the building's ability to resist earthquakes.
- (3) BIM technology contains a lot of information related to the construction of structural elements, which can be calculated, and can also provide comprehensive and accurate data support for the seismic capacity of housing buildings.
- (4) BIM technology is a parametric, process-oriented application technology that can visually express the impact of the building's components and the whole on the seismic capacity of housing buildings.

### 2.3 AHP Method to Determine Indicator Weights

The use of AHP to categorise the identified risk indicators, in a hierarchy, using only a qualitative approach, is too subjective and unconvincing. For the two compared risk indicators to score, the size of the number to judge the level of importance of their corresponding indicators, where  $i$  and  $j$  represent for the comparison of Two factors,  $a_{ij}$  represents the scale in which, the definition of the scale of each indicator is detailed in Table 1.

**Table 1.** Definition of scales for each indicator

Scale ( $a_{ij}$ )	Definition
1	indicates that two indicators $i, j$ are of equal importance compared to each other
3	Indicator $i$ is slightly more important than indicator $j$
5	Indicator $i$ is more important than indicator $j$
7	Indicator $i$ is more important than indicator $j$
9	Indicator $i$ is particularly important than indicator $j$
2, 4, 6, 8	Importance is in the middle of the above two adjacent judgements
Inverse of the above scale	If indicator $i$ is compared with indicator $j$ , the scale is $a_{ij} = 1/a_{ji}$

Based on the survey data in this paper, the eigen root method is proposed to be used to determine the weights of the judgement matrix. Firstly, the hierarchical single ordering is determined, and according to the judgment matrix  $A$ , the maximum eigenvalue  $\lambda_{max}$ , then the corresponding eigenvector is calculated by the corresponding formula, and finally normalisation is carried out.

The judgment matrix A constructed is.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad (1)$$

Calculate the product of the elements of each row of the judgment matrix as shown in Eq. (2)

$$M_i = \prod_{j=1}^n f_{ij} \quad (i = 1, 2, 3, \dots, n) \quad (2)$$

Then calculate the  $n$ th root  $w_i$  of the product  $M_i$ , as shown in Eq. (3)

$$\bar{w} = \sqrt[n]{M_i} \quad (3)$$

In the next step, the vectors of 3 are normalized as shown in Eqs. (4) and (5)

$$w_i = [\bar{w}_1, \bar{w}_2, \bar{w}_3, \dots, \bar{w}_n]^T \quad (4)$$

$$W_i = \frac{\bar{w}_i}{\sum_{i=1}^n \bar{w}_i} \quad (5)$$

### 3 Experimental Research

#### 3.1 Principles for the Selection of Comprehensive Evaluation Indicators

- (1) **Scientific:** The indicators selected for the comprehensive evaluation should have information that is objective and representative of the main factors of the problem under study, so that the results reflected are in line with objective reality; the evaluation methods are well integrated with the evaluation indicator system and can complement each other, so that the evaluation indicator system can reflect the actual situation scientifically through digital representation.
- (2) **Feasibility:** The evaluation system established should be easy to implement, the indicators of the evaluation system should be simple and clear, but not too little and too simple, to avoid the omission of information, to prevent the emergence of false results, untrue phenomenon.
- (3) **Integrity:** It should start from the whole of the object under study, from all aspects, to ensure that there is a certain logical relationship between each indicator, to reflect the structural characteristics of the object under study and the degree of influence of each part from all aspects, to achieve a comprehensive overview, in order to ensure that the selected indicators can objectively and truly reflect the actual phenomenon.
- (4) **Independence:** The selected indicators should be typical and representative, avoiding the selection of indicators that are too detailed and too cumbersome, reducing this functional overlap, and ensuring that the indicators are independent individuals of each other.
- (5) **Combination of qualitative and quantitative:** A combination of qualitative and quantitative methods should be used to assign values according to the specifications and assignment principles.

### 3.2 Process for Evaluating the Seismic Capacity of Masonry Houses

To evaluate the seismic capacity of a building, firstly, the building is analysed according to the set indicators, a suitable evaluation level is selected for each indicator, then the weights of the indicators in the factor and indicator layers are called up for calculation, and finally the results of the calculation are obtained, according to which the strength of the seismic capacity of the building is judged.

Before an earthquake, an evaluation of the seismic capacity of a building can be used for forecasting, prevention and emergency response, so that if the house is highly resistant to earthquakes, you can be aware of the situation; if the building is weak, the building will need to be reinforced to prevent future damage. After the earthquake, the evaluation of the seismic capacity of the housing building can be done for rescue, relief and reconstruction, and it can be judged whether the house needs to be reinforced for further use or needs to be rebuilt to protect the lives and property of the occupants.

## 4 Experimental Analysis

### 4.1 Determination of Weights of Evaluation System Indicators

The questionnaire was mainly paper-based and supplemented by an online questionnaire, of which 200 paper-based questionnaires were distributed and 156 were returned, 124 valid, accounting for 62.0%; 113 online questionnaires were returned, 97 valid, accounting for 85.84%. Among them were analysed in terms of work, job title, concern for the seismic resistance of housing buildings and the need to improve the seismic resistance of housing buildings. As shown in Table 2.

**Table 2.** Relevant statistical analysis

Results of personnel engagement statistics	Construction unit		Designers	Supervisors		Research Unit	Other
Percentage	36.1%		13.6%	10.4%		29.9%	10%
Personnel title statistics results	Senior Engineer	Engineer	Technician	Professor	Associate Professor	Lecturer	Other
Percentage	5%	32.6%	22.2%	5.4%	10.4%	14.5%	10%
Personnel are concerned about the statistical results of the seismic capacity of buildings							
	Very concerned	Fairly concerned		Generally concerned		Rarely	Not at all
Percentage	6.3%	32.1%		29.4%		22.2%	10%
Statistical results on the need for personnel to improve the seismic resistance of buildings							
	Very much needed		relatively necessary	Generally needed	May need	Don't need	Don't know
Percentage	4.5%		36.2%	30.8%	14.9%	7.2%	6.3%

The analysis of the results shows that most of the participants in the survey are in the construction industry, with the largest number of them working in construction units and the second largest number of them working in scientific research at 29.86%; among the participants in the survey, more than 50% are engineers and technicians, followed by lecturers and associate professors; in the part on whether to pay attention to the seismic capacity of buildings.

#### 4.2 Weighting Values for System Level Indicators

The system layer consists of four parts: seismic source, foundation, base and superstructure, to which values are assigned and obtained as shown in Table 3.

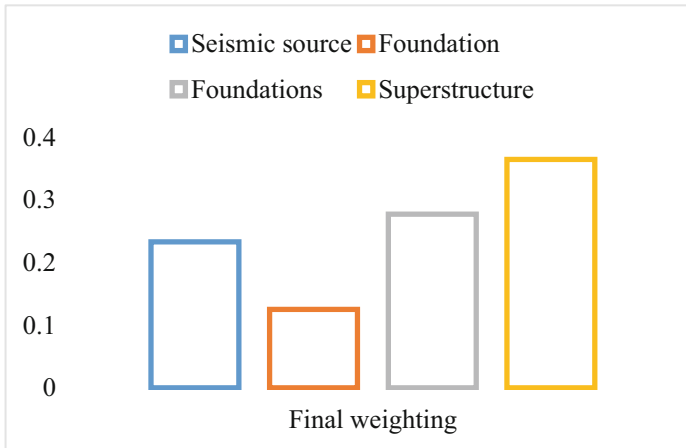
**Table 3.** Weighting relationships between the indicators at the system level

Systematic layers	Seismic source	Foundation	Foundations	Superstructure
Seismic source	1	2	1	1/2
Foundation	1/2	1	1/2	1/3
Foundations	1	2	1	1
Superstructure	2	3	1	1

The judgment matrix can be obtained according to the weight relationship in the table, and the eigenvectors and maximum eigenvalues can be obtained according to the formula. The ranking and histogram of the eigenvector weights of the system layer are shown in Table 4 and Fig. 1.

**Table 4.** Calculation results and ranking of the weights among the various indicators of earthquake

Systematic layers	Seismic source	Foundation	Foundations	Superstructure
Final weighting	0.2330	0.1252	0.2771	0.3647
Overall ranking	3	4	2	1



**Fig. 1.** Histogram of system level indicator weights

## 5 Conclusions

The BIM model contains the overall building data information of the building, and through different professional functional software, the relevant information in the model is analysed and used to carry out analysis of building layout, pipeline inspection, fire protection facilities, sunlight and energy saving analysis, code standard checking, etc. The application of BIM technology makes the data information of various professions share and linkage, timely reflecting the arrangement between various professions and systems Reasonableness, effective collision check. Accelerate the development of BIM technology application, play its full information, real-time, visual, virtual, sharing and other characteristics advantages, enhance the comprehensive management technology level of pre-planning, submission, design, construction and operation and maintenance of construction projects, save materials, save time, reduce investment, improve quality and increase efficiency and other indicators.

## References

1. Mesáros, P., Mandicák, T., Behúnová, A.: Use of BIM technology and impact on productivity in construction project management. *Wirel. Networks* **28**(2), 855–862 (2022)
2. Nath, B.K., Sen, N.: A partially backlogged two-warehouse EOQ model with non-instantaneous deteriorating items, price and time dependent demand and preservation technology using interval number. *Int. J. Math. Oper. Res.* **20**(2), 149–181 (2021)
3. Utomo, B., et al.: Enhanced of attendance records technology used geospatial retrieval based on crossing number. *Int. J. Interact. Mob. Technol.* **14**(16), 101–116 (2020)
4. Bjørn, P., Wulff, M., Petráus, M.S., Møller, N.H.: Immersive Cooperative Work Environments (CWE): designing human-building interaction in virtual reality. *Computer Supported Cooperative Work (CSCW)* **30**(3), 351–391 (2021). <https://doi.org/10.1007/s10606-021-09395-3>

5. Rorabaugh, A.K., Caño-Lores, S., Johnston, T., Taufer, M.: Building high-throughput neural architecture search workflows via a decoupled fitness prediction engine. *IEEE Trans. Parallel Distributed Syst.* **33**(11), 2913–2926 (2022)
6. Santos, J.P.G., et al.: A modular workflow for model building, analysis, and parameter estimation in systems biology and neuroscience. *Neuroinformatics* **20**(1), 241–259 (2022)
7. Satapathy, S.: Work place discomfort and risk factors for construction site workers. *Int. J. Syst. Assur. Eng. Manag.* **13**(2), 668–680 (2022)
8. Lin, Y., Matsumi, N.: Visuospatial working memory and the construction of a spatial situation model in listening comprehension: an examination using a spatial tapping task. *Cogn. Process.* **23**(1), 41–54 (2022)
9. Constantinou, S., Konstantinidis, A., Chrysanthis, P.K., Zeinalipour-Yazti, D.: Green planning of IoT home automation workflows in smart buildings. *ACM Trans. Internet Things* **3**(4), 29:1–29:30 (2022)
10. García, A.J., González-Rodrigo, B., Martínez, S., Martínez, R., Marchamalo, M.: Building health monitoring in the old town of Madrid: applicability of SAR Imagery to the monitoring of underground works through classification indexes. *Int. J. Digit. Earth* **14**(3), 271–287 (2021)
11. Traboulsi, S., Knauth, S.: IOT analysis and management system for improving work performance with an IOT open software in smart buildings. *J. Ubiquitous Syst. Pervasive Networks* **14**(01), 1–6 (2021)
12. Cornwell, C., Schmutte, I.M., Scur, D.: Building a productive workforce: the role of structured management practices. *Manag. Sci.* **67**(12), 7308–7321 (2021)



# Design and Implementation of Virtual Simulation Teaching Platform of Sports Action Technology Based on Maya

Sicong Shan and Shuang Sun<sup>(✉)</sup>

Jilin Engineering Vocational College, Siping, Jilin, China  
tujiuruijun@163.com

**Abstract.** The application of virtual simulation technology in physical education teaching conforms to the development trend of education information and can improve the effect of physical education teaching. An effective virtual simulation teaching platform construction scheme of sports action technology was proposed. The motion capture system and surface EMG equipment were used to synchronize sports action acquisition, and the motion capture data file was bound to the drawn human muscle model for rendering, which was integrated into the built platform. The platform focuses on the combination of virtual and real, optimizes cognition, improves teaching efficiency and quality, and provides convenience for researchers with the stored movement-related data. The platform is extensibility, which can load the reserved interface for other sports movement data and pave the way for the construction of sports movement database, and has broad application prospects.

**Keywords:** Virtual Simulation · Physical Education · Motion Capture · Surface Electromyography · Sports Action Library

## 1 Introduction

In recent years, virtual simulation technology has developed rapidly at home and abroad, and its application in the field of teaching has been praised by many professionals. In foreign countries, the technology has been widely used in college classroom teaching, and in China, it has also been preliminarily applied in physics and chemistry experiments and other classes. Virtual simulation teaching conforms to the development of The Times, meets the needs of educational modernization, and provides technical support for the construction of characteristic and advantageous disciplines [1].

Virtual simulation teaching meets the needs of education modernization. The so-called teaching process mainly includes teaching content, teaching methods, equipment and other contents, which have a profound impact on the formation and development of students' way of thinking and ability. With the development of The Times, the teaching form is constantly updated. In the face of the development trend of the application of virtual simulation teaching abroad, which has gradually become resource networking, remote activities and management information, our country has introduced the 10-year

Development Plan of Education information, Education Information 2.0 Action Plan and other planning policies successively [2]. The virtual simulation technology further permeates into the teaching fields of different levels and various disciplines, stimulates students' learning interest and cultivates students' innovative thinking. On the whole, it promotes the application of virtual simulation technology in the teaching field and speeds up the implementation of educational modernization [3].

Virtual simulation teaching meets the needs of the development of physical education. The basic means of sports is exercise, the ultimate goal is to enhance physical fitness and improve the level of sports technology, not only involving biomechanics, psychology, physiology and other natural disciplines, but also involving sports education, training, sociology and other social sciences. The demand for highly qualified sports specialized talents grows continuously in our country, hoping that the college physical education and sports related majors can cultivate multi-competency sports professionals who are capable of engaging in sports teaching, training, competitions, management and social sports guidance. Therefore, colleges and universities must reform the traditional physical education form, and the application of virtual simulation technology makes up for the deficiency in the traditional teaching process to a certain extent. But at present, virtual simulation physical education mainly focuses on the sports theory level, in the sports application level and development level is less, and almost blank for sports rehabilitation, sports engineering and other directions of talent training [4].

Under the sharing mechanism, universities can cooperate and exchange with each other to build a series of high-level virtual simulation teaching resources, avoid repeated construction, and realize cross-specialty and cross-department comprehensive sharing, integration and optimization with the help of Internet technology, cloud storage and other technologies, so as to finally realize the good wish of benign development. Following the wave of education information reform in colleges and universities, it is necessary to build a highly integrated, interactive and completely open virtual simulation teaching platform of sports action technology and its application to sports colleges and related disciplines to adapt to the increasing trend of the social supply of high-quality compound sports talents [5].

## 2 Research Purpose and Significance

Following the wave of education information reform, the construction of highly integrated, interactive, fully open virtual simulation teaching platform of sports movement technology, the combination of virtual and real, optimize cognition, comprehensive application [6]. The 3D human movement technology movement material library is established to assist the teaching staff to use the virtual simulation teaching system to visually show the key points of movement technology, enhance the interaction, break through the traditional physical education mode, and improve the safety, systematization and scientific nature of physical education. Give play to the leading role of teachers and reflect the principal position of students in the learning process. It is convenient for teachers to teach, improve teaching efficiency, enhance the fun of teaching process, analyze and compare the differences in sports movements, so as to enhance students' learning enthusiasm and initiative, enhance the interaction between teachers and students, and



build a student-centered learning environment. Digital sports action resources can be shared on the platform to meet the needs of virtual simulation of sports body material database for different sports events and promote the interconnection of virtual simulation of sports body material database for different sports events [7]. The original data of biomechanics can be downloaded to facilitate further analysis by relevant researchers.

### **3 The Necessity of Applying Virtual Simulation to Physical Education Course Teaching**

In the context of national fitness, college students, as the main force of the country's future society, are bound to learn professional sports science knowledge, improve their physical quality, and develop correct sports ideology and ideology. In addition, the improvement of the competitive ability of professional sports athletes also needs the guidance of professional coaches and other professionals. Therefore, sports teaching are very important. Physical education teaching curriculum also needs to go deep into the essence of the attempt to reform [8].

With the rapid development of modern technology, there are more and more researches on adding new technology teaching into physical education curriculum. Teaching methods and means have been innovated, such as flipped classroom, Wechat platform assisted teaching, MOOCs, micro classes, etc. In short, the teaching situation of physical education major is not ideal, and there are deficiencies in teaching objectives, teaching content, teaching means and teaching evaluation [9]. Domestic scholars have been exploring new physical education teaching methods to improve the problems in the teaching process to a certain extent.

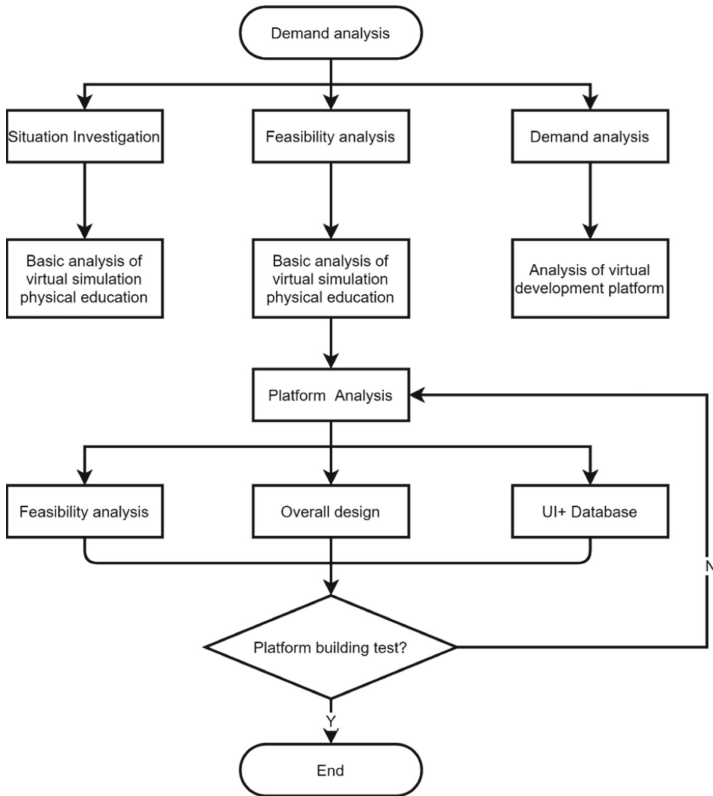
The application of virtual simulation in physical education teaching is an inevitable trend of development. Virtual simulation technology can change the traditional teaching methods, arouse the enthusiasm of students to learn sports in the teaching process, and fully show the technical details easy to miss in the traditional teaching, outstanding demonstration. Nowadays, the teaching fields of different disciplines are developing with cross-penetration, and the knowledge of the science of movement and human body in the application of virtual simulation teaching cannot be displayed intuitively, which makes it difficult for students to feel, which is not conducive to the training of compound sports talents [10].

The application of virtual simulation teaching is mainly focused on physical education and training majors, and the object oriented is physical education teachers and students, coaches and athletes. There are few researches on sports engineering and human science talent training, and the covered subject categories are not complete enough. The original resources of virtual simulation teaching have the problems of safety and repeated construction.

### **4 Research Content and Scheme**

According to the current situation and requirements, the framework and functions of the design platform are studied. Through the comprehensive application of key technologies

of the motion data acquisition system and the human motion model system, a feasible and scientific motion acquisition scheme is designed, the model is built and integrated on the platform to realize the synchronous presentation of multi-angle 3D simulation animation of sports movements and the visualization of biomechanical data, and the key points of motion technology are visually displayed. Enhance interactivity. To realize the sharing of digital sports action resources, meet the needs of different users, including students, teachers and researchers, for virtual sports simulation information, and promote the establishment of virtual physical simulation physical action technical material library. The flow chart of the research content and scheme is shown in Fig. 1.



**Fig. 1.** The flow chart of the research content and scheme

Determine the application status, existing problems and development trend of virtual simulation physical education teaching, systematically comb out the key technologies and related teaching theories required in the process of building a virtual simulation teaching platform for human movement, conduct feasibility analysis, and then sort out the needs of the platform through interviews. Overall design of the platform based on design principles, followed by functional design, database design and UI interface design; In terms of specific content, the case sports action is selected as the representative, the data collection scheme is designed, the model and platform are drawn, and the

difficulties and solutions existing in the construction process of sports action technology virtual simulation teaching project are discussed, providing technical support for the development of sports virtual simulation teaching in the future.

## 5 The Core Code

```

using SYS;
using SYS.Drawing;
using SYS.Collections;
using SYS.ComponentModel;
using SYS.Bat.Forms;
using SYS.Data;
using SYS.Net;
using SYS.IO;
using SYS.NPOI;
namespace ExcelFileExporter
{
public class Form1.SYS.Bat.Forms.Form
{
private SYS.Bat.Forms.Label label1;
private SYS.Bat.Forms.Label label2;
private SYS.Bat.Forms.TextBox srcAddress;
private SYS.Bat.Forms.TextBox tarAddress;
private SYS.Bat.Forms.StatusBar statusBar;
private SYS.Bat.Forms.Button Start;
private WebClient client=new WebClient();
statusBar.Text =;
client.DownloadFile(URLAddress,fileName);
Stream str=client.OpenRead(URLAddress);
StreamReader reader=new StreamReader(str);
while(allmybyte>0)
{
int m=str.Read(mbyte,startmbyte,allmybyte);

```

## 6 Key Technologies

### 6.1 Motion Capture Notion

Motion capture technology involves size measurement, can use the dynamic capture system (sensor, signal capture equipment, physical space object positioning and azimuth-determination, data transmission equipment) to track the key parts of the moving object,

and then through computer processing to obtain three-dimensional trajectory and three-dimensional spatial coordinates of the specific data, in animation production, Gait analysis, biomechanics, ergonomics and other fields are widely used. In terms of animation production, the processing of human movement data obtained from motion capture can drive the 3D model to virtualize and restore the physical action, rather than manually simulate the formation, with realistic effect. There are many kinds of motion capture technology, according to the different principles, can be divided into mechanical, acoustic, electromagnetic and optical. Users choose the most suitable dynamic capture system from the multi-dimensional analysis of the advantages and disadvantages of each capture system, such as convenience, accuracy and real-time.

## 7 Digital Modeling Maya

Autodesk Maya is the world's top 3D animation software produced by Autodesk, which integrates the mainstream digital effects technology of Alias and Wave front. It has the features of perfect function, flexible work, easy to learn and use, high production efficiency, strong rendering sense and so on. It can run on Windows NT, SGI IRIX and a variety of operating systems. With its ability to combine state-of-the-art modeling, digital fabric simulation, hair rendering, and motion matching technologies to raise the standard of virtual graphics, Maya is the preferred solution for many developers. Maya itself contains curve editor, time editor, Tax editor and other editors that can achieve many special effects. The specific modeling methods are NURB S modeling, polygon modeling, subdivision modeling, sculpting modeling; At the same time, the use of posture editor, deformation editor and Human IK tool can realize the bone binding of the model in Maya software, which lays a foundation for the virtual simulation of the real movement of human body and improves the practical feeling of users.

## 8 Conclusions

It is necessary to build a highly integrated, interactive and completely open virtual simulation teaching platform of sports movement technology. This paper studies the overall demand framework of the virtual simulation teaching platform of sports movement technology based on simulation technology, and puts forward a set of effective simulation scheme, and puts forward the corresponding solutions to the problems encountered in the realization process. Major contributions are as follows.

Based on the analysis of the current situation and research results, the embryonic form of virtual simulation teaching platform of sports movement technology is proposed. Analyze the user group, simulation content and system objectives, propose the overall demand of the platform, and then carry out functional and non-functional demand analysis; Design the platform framework, each function module, database and UI interface.

Viton motion capture system and surface EMG equipment are used to synchronize the collection of functional training actions, Maya and Motion Builde.

Unity3D script framework combined with C# programming technology was used to build a virtual simulation teaching platform, which realized the storage and management

of kinematic data, surface EMG data and video shooting; The synchronous display of virtual simulation animation and EMG signal was realized, and the difference between the advantages and disadvantages of functional training actions was presented.

## References

1. Avila-Pesantez, Rivera, L.A., Iban, M.S.: Approaches for serious game design. *Comput. Educ. J.* **10**, 2–13 (2017)
2. Vallance, M., Schaik, P.: LEGO Mindstorms Proceedings of the 2019 Computer Society Conference on Human-Computer Interaction. 9, 150–160 (2019)
3. Nguyen, T., Dao, T.: Applications of motion optimization algorithm in the travelling salesman problem. *J. Comput. Eng.* **22**, 61–67 (2019)
4. Trong-The, N., Tsu-Yang, W., Jeng-Shyang, P.: Node coverage optimization strategy based on motion optimization. *J. Network Intell.* **5**, 2–10 (2019)
5. Deussom, M., Tonye, T.: New propagation model optimization approach based on particles swarm algorithm. *Int. J. Comput. Appl.* **12**, 37–45 (2015)
6. Deussom, B., Tonye, E., Kabiena, E.: Propagation model optimization based on artificial colony algorithm. *J. Electr. Electron. Eng.* **16**, 15–27 (2020)
7. Deussom, B., Souop, K.D.D., Tonye, E., Michael, E.: Social spider algorithm-based approach for propagation optimization. *Int. J. Eng. Res. Technol.* **12**, 528–534 (2022)
8. Tonye, B., Deussom, K.: Optimisation de modèles de propagation à partir données de mesures radio de ville de Yaoundé. *J. Cameroon Acad. Sci.* **15**, 181–200 (2016)
9. Wan, T., et al.: RGB-D point cloud registration via infrared camera. *Multimed. Tools Appl.* **23**, 32–40 (2019)
10. Yutai, R., Fan, Y.: Research on path tracking algorithm of vehicle based on image processing. *Int. J. Pattern Recogn. Artif. Intell.* (05), 58–69 (2019)



# Industrial Economy and Regional Growth Strategy Based on Particle Swarm Optimization Algorithm

Hengran Bian<sup>1</sup>(✉) and Yi Liu<sup>1,2</sup>

<sup>1</sup> Institute of Strategy Research for the Guangdong-Hong Kong-Macao Greater Bay Area, Guangdong Academy of Sciences, Guangzhou 510070, Guangdong, China  
b18092100321@cityu.mo

<sup>2</sup> Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

**Abstract.** With the rapid growth of the global economy, it has become an inevitable trend for Chinese enterprises to transform and upgrade their industries. This paper starts with the research status at home and abroad, and analyzes the problems existing in the current industrial growth of China in combination with relevant literature and cases in the industry. Then, through the research of particle swarm optimization (PSO) algorithm, the growth trajectory of major cities in China and the regional competition status of each region are summarized. Finally, the PSO algorithm is applied to the optimization solution model of the location of industries in various provinces and cities in China based on the regional economic and social growth strategy, and the final results are obtained. The results show that the coordinated growth coefficient and comprehensive growth level are about 5 and 7 in this model. The level of coordinated growth is shown in 5 test results. Therefore, the industrial economy and regional growth strategy model based on PSO algorithm meets the needs of users.

**Keywords:** Particle Swarm Optimization · Industrial Economy · Regional growth · Strategic Research

## 1 Introduction

The growth of enterprises cannot be separated from the regional economy, and the industrial economy has an inseparable relationship with the country. Therefore, the industrial strategy and regional layout should be taken into account in regional planning [1, 2]. Based on the analysis of relevant literature, periodical materials and industry data at home and abroad, this paper finds that China is currently in a period of transformation from manufacturing to “Internet plus”. However, due to the late start and rapid growth of this field in China, people generally lack this knowledge, so it is necessary to combine it with China’s specific national conditions to formulate corresponding economic growth strategies and regional layout strategies [3, 4].

Many scholars have conducted relevant studies on regional growth strategies. In the research of industrial economy and regional growth strategy, foreign scholars mainly take technical economics, management and mathematics as the theoretical basis, and use a variety of disciplinary methods to explore. Domestic scholars have proposed a new idea “black box method” [5, 6]. He analyzed the enterprise’s competitive advantage and its influencing factors through the internal structure and production process of the enterprise. At the same time, he compared the specific situation of China with the relevant foreign fields and drew conclusions. Some domestic scholars also put forward their own research models on industrial economy and regional growth strategy. Some scholars used mathematical methods to analyze overcapacity in China’s manufacturing industry and changes in industrial structure [7, 8]. Some scholars use the “enterprise” internal governance theory to discuss the problems in the growth of China’s manufacturing industry. They believe that the company should start from improving production efficiency and constantly innovate the technical level. Therefore, based on the PSO algorithm, this paper studies the industrial economy and regional growth strategy.

Under the background of global economic integration, industrial growth strategy has increasingly become a topic of academic concern. This paper mainly introduces the interdependence and coordinated growth between industry economy and regional construction. Based on the “point axis” structure, the mathematical model and algorithm analysis technology are combined to build an indicator system for planning the growth rate of regional GDP and the growth rate of enterprise output value based on the growth of urban industrial gross output value. Then, the connection between the nodes in the system is used to carry out optimization solutions such as social capital flow, so that the industrial economy can develop more stably and rapidly.

## **2 Discussion on Industrial Economy and Regional Growth Strategy Based on PSO**

### **2.1 Regional Growth of Industrial Economy**

Industrial economy refers to the sum of material capital investment and commodity exchange activities among all enterprises in a country or region to achieve certain goals. It is also known as the constituent elements of the productive national economy - namely, various resources. At the same time, it is a dynamic balance process formed by the inter-relationship and combination of the production factors of various sectors in the national economy. It is divided by regions. It has obvious regional and industrial characteristics [9, 10]. It mainly includes: first, resource allocation and product structure; The second includes technical level and market scale. From a macro point of view, “region” should include a regional growth model within the social scope, market-oriented and resource allocation. In a narrow sense, it generally refers to a special form arising from the interconnection between the means of production or service departments. Economy and society are complementary and mutually reinforcing entities, which interact and jointly promote economic growth. Therefore, in real life, industry is often referred to as “national economy”. It includes three levels: production factors, intermediate products and basic services. The most important one is the production factor, that is, the basic

technical level. Intermediate products refer to the provision of special tools or other auxiliary materials needed for raw materials and parts, as well as various physical entities. It also includes the characteristics of the strength of the enterprise's own ownership or control ability and its sensitivity to changes in the market environment [11, 12]. In today's society, the competition for enterprise growth is becoming increasingly fierce, and industrial economy is an important condition for a country or region to achieve its long-term sustainable growth strategic objectives. At the present stage, China's industrial economy and real estate industry are in an underdeveloped state. Therefore, if we want to better promote the optimization of social resource allocation, drive the improvement of production factors and technological innovation capabilities in the whole region, we need to make rational and effective use of the industry and urban agglomeration, so as to promote enterprises to obtain an advantageous position in the fierce competition, and then obtain greater benefits.

## **2.2 Importance of Regional Growth of Industrial Economy**

The industrial growth strategy has a very important and profound impact on the national economic growth of a country. It is not only related to the gross national product and the domestic employment rate indicators, but also directly reflects the coordination between government policies and regional economy, as well as the role of various factors in sustainable growth, and determines its future role in the macro-economy, extent of play and prospects. With the continuous growth of science and technology and productivity, the market demand is changing day by day. In this case, the product output is also increasing. This requires the combination of industrial economy and regionalization strategy. Through the improvement of technical innovation, management mode and other aspects, the whole industry has been promoted to develop in a benign direction. At the same time, the enterprise itself can get better growth space and improve efficiency faster to obtain more efficient production capacity and stronger competitiveness. Industrial economy is an important basis for the growth of a country. In the process of globalization, China, as the largest developing country in the world, needs a large number of foreign enterprises to carry out production and business activities in its domestic market and international market. Regional economy is based on geographical space and natural resources, and within a certain region, it forms the increasing scale benefit, agglomeration effect and comprehensive production capacity through market mechanism. With the improvement of social productivity and the improvement of people's quality of life and other factors, the level of consumer demand has been rising, and consumption habits have also changed. Therefore, if enterprises want to obtain more profits, they need to deeply process and sell the products they produce to increase their income sources to obtain greater growth space and opportunities. At the same time, as a new thing, industrial economy is of great significance in regional growth.

## **2.3 PSO Algorithm**

Through the study of particle swarm, a series of random time-varying functions are formed based on probability distribution, wavelet transform and fuzzy theory, and the global optimal solution is found in non-uniform, discrete and multi-dimensional state



space. Apply it to system theory. PSO is an unsupervised optimization method. So it can solve the uncertainty problem and has high robustness. In addition, particle swarm can also be used to optimize the search process. In PSO (PSO), particles are randomly generated, that is, a new individual is constructed according to different probability distributions. Then the population is updated to form new individuals. The iteration tree algorithm is a global optimization method based on this idea. It can also be called “black box” strategy or “fuzzy control technology”, etc. “PSO” algorithm is that ants fly randomly on different nodes, search all particles according to certain rules, and then minimize the whole world according to these special information. This method can realize resource sharing and improve system performance. First, choose an optimal operation path to find the optimal solution point and optimal scheme; Finally, the maximum possibility value of the final result is obtained by iteratively updating from the best solution (i.e. “meeting the maximum value to the maximum extent” or “minimizing the delay time”). In the PSO algorithm, all particles have the same or similar trajectory, and each individual is independent and can not be ignored. So we can use this principle to solve a new problem. Its basic idea is to obtain the optimal solution or global optimal value by analyzing the location function relationship and path structure formed under all initial states, and then select the corresponding algorithm according to different types of special parameters (such as space, time, etc.), and design a particle swarm model to obtain the ultimate goal of the optimization results, so as to achieve the strategic layout of economic growth. Suppose that the dimension of the search space is  $D$ , representing the position of the  $i$ th particle in PSO, expressed by the variable  $x = x_i$ , and the velocity of the  $i$ th particle expressed by  $v$ . The optimal position of the  $i$ th particle searched until the current iteration is represented by the variable  $P$ , and the optimal position searched until the current iteration is represented by the variable  $p_g$ . For each particle, its  $d$ -th dimension ( $1 \leq d \leq D$ ) varies according to the following equation:

$$v_{id}(t + 1) = \omega v_{id}(t) + c_1 rand_1() (P_{id} - x_{id}(t)) + c_2 rand_2() (P_{gd} - x_{id}(t)) \quad (1)$$

$$\begin{cases} v_{id} = V_{max}, \text{ if } v_{id} > V_{max} \\ v_{id} = -V_{max}, \text{ if } v_{id} < -V_{max} \end{cases} \quad (2)$$

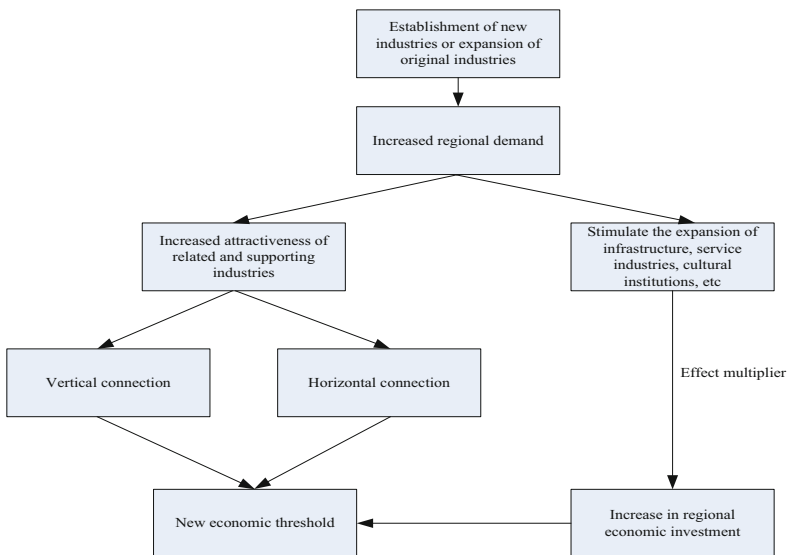
Among them,  $C_1$  and  $C_2$  are non negative acceleration constant factors, which respectively represent the weights of statistical acceleration terms that push particles to the  $P_{id}$  and  $P_{gd}$  positions, and the values have different meanings: a large value indicates that particles will suddenly rush or cross the target area, while a small value indicates that particles can wander  $rand_1()$  and  $rand_2()$  outside the target area before being pulled back.

### 3 Experimental Process of Industrial Economy and Regional Growth Strategy Based on PSO Algorithm

#### 3.1 Industrial Economy and Regional Growth Strategy Based on PSO Algorithm

From the perspective of industry (as shown in Fig. 1), enterprises often have overcapacity, product backlog and other problems in the production process. Therefore, in

order to solve these problems, we need to seek new methods. First, we need to improve productivity. Industry is one of the pillars of the national economy and an important symbol of national economic growth and social progress. Second, we need to optimize and upgrade the existing technology level and technical structure. Third, we can enhance the management control and supervision of the industry through innovation and introduction of advanced talents. The industrial economy and regional growth strategy based on PSO algorithm is to apply it to the production and operation of enterprises, and how to make these resources play the maximum role, so as to obtain the maximum benefits. There is a basic, progressive and phased industrial system in the whole socio-economic structure. It includes a series of product systems with technology as the support point and consumption as the main body. At the same time, it is also necessary to consider social needs, market needs and other factors to optimize the combination and adjustment of various links involved in the enterprise's production and operation, so as to achieve the optimal overall function.



**Fig. 1.** Strategic Process of Industrial Economy and Regional growth Based on Microgroup algorithm

### 3.2 Test of Industrial Economy and Regional Growth Strategy Based on PSO Algorithm

In the research process of this paper, we will find out their problems by comparing domestic and foreign enterprise growth strategies, industrial economy and regional growth strategies, and then put forward targeted suggestions. It is mainly to calculate the product price level of each production unit and the degree of product difference between each production unit according to the average profit rate and cost cost ratio of the industry, and

use the “market barrier” method to determine the basic marginal contribution coefficient, unit value density index and other indicators within each company to build a comprehensive evaluation model for the enterprise’s competitive strength. It is necessary to conduct a comprehensive investigation on the enterprise’s operation, production capacity and market demand, Then compare all the sample data with the samples required in the model, and determine whether it is feasible by comparing the experimental values obtained under different schemes.

## 4 Experimental Analysis of Industrial Economy and Regional Growth Strategy Based on PSO Algorithm

### 4.1 Test and Analysis of Industrial Economy and Regional Growth Strategy Based on PSO Algorithm

The research on industrial economy and regional growth strategy based on PSO algorithm is a multi-disciplinary, multi domain knowledge, and has strong operability and feasibility, so it is more complex in application. In order to better achieve the objectives of evaluating and analyzing the production capacity of enterprises and promoting the economic and technological level of the industry. This paper tested the industrial capacity system model of PSO algorithm. Table 1 shows the test data of industrial economy and regional growth strategy.

**Table 1.** Industrial economy and regional growth strategy test data

Test times	Coordinated growth level	Coordinated growth coefficient	Comprehensive growth level
1	Coordinate	5.56	7.34
2	Coordinate	5.32	7.97
3	Basic coordination	5.76	7.86
4	Coordinate	5.98	7.23
5	Basic coordination	5.82	7.63

In the research of industrial economy and regional growth strategy based on PSO algorithm, the mathematical model is established by calculating the unit output value, GDP per 10000 yuan and gross social output value of each enterprise. This method can not only solve the collinearity problem between single enterprises, but also effectively avoid the competition among multiple industries. In the solution process, we can also consider the impact of multiple independent operating systems on their performance and the relationship with other related systems for comprehensive analysis, and finally calculate the optimal solution according to the matrix algorithm. As shown in Fig. 2, the coordinated growth coefficient and comprehensive growth level in the model are about 5 and 7, and the coordinated growth level is consistent in five test results. Therefore, the industrial economy and regional growth strategy model based on PSO algorithm meets the needs of users.

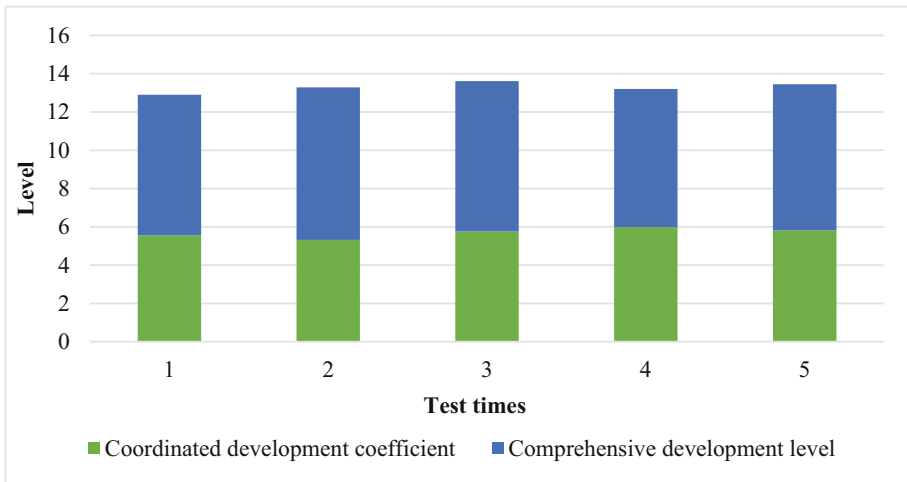


Fig. 2. Test of Industrial Economy and Regional growth Strategy Based on Microgroup algorithm

## 5 Conclusions

With the growth of economy, industrial economy has gradually become a topic of concern, and it is also one of the most competitive areas for enterprises in the world. This paper mainly introduces the research on regional growth strategy based on the technical route of “market location selection”. First, it explains the influencing factors from the three aspects of theoretical basis, application model and optimal path, and how to use mathematical tools to solve relevant practical problems in reality. Then, it draws a conclusion that it is feasible to form an industrial chain with certain economic value in space by using PSO algorithm to establish experimental examples and analyze its effect map.

**Funding.** This work was supported by Special Fund project of National Natural Science Foundation of China (42130712); GDAS Special Project of Science and Technology Development (2020GDASYL–20200102002).

## References

1. Kumar, R., Bansal, H.O., Gautam, A.R., Mahela, O.P., Khan, B.: Experimental investigations on particle swarm optimization based control algorithm for shunt active power filter to enhance electric power quality. *IEEE Access* **10**, 54878–54890 (2022)
2. Lakshmi, Y.V., Singh, P., Abouhawwash, M., Mahajan, S., Pandit, A.K., Ahmed, A.B.: Improved chan algorithm based optimum UWB sensor node localization using hybrid particle swarm optimization. *IEEE Access* **10**, 32546–32565 (2022)
3. Langazane, S.N., Saha, A.K.: Effects of particle swarm optimization and genetic algorithm control parameters on overcurrent relay selectivity and speed. *IEEE Access* **10**, 4550–4567 (2022)

4. Sharma, S.K., Khambampati, A.K., Kim, K.Y.: Hybrid particle swarm optimization-gravitational search algorithm based detection of graphene defects with electrical impedance tomography. *IEEE Access* **10**, 105744–105757 (2022)
5. Ahandani, M.A., Abbasfam, J., Kharrati, H.: Parameter identification of permanent magnet synchronous motors using quasi-opposition-based particle swarm optimization and hybrid chaotic particle swarm optimization algorithms. *Appl. Intell.* **52**(11), 13082–13096 (2022)
6. Dadvar, M., Navidi, H., Javadi, H.H.S., Mirzarezaee, M.: A cooperative approach for combining particle swarm optimization and differential evolution algorithms to solve single-objective optimization problems. *Appl. Intell.* **52**(4), 4089–4108 (2022)
7. Kumar, D., Pandey, M.: An optimal and secure resource searching algorithm for unstructured mobile peer-to-peer network using particle swarm optimization. *Appl. Intell.* **52**(13), 14988–15005 (2022)
8. Dhruv, B., Mittal, N., Modi, M.: Improved particle swarm optimization for detection of pancreatic tumor using split and merge algorithm. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **10**(1), 38–47 (2022)
9. Nartey, C., et al.: Blockchain-IoT peer device storage optimization using an advanced time-variant multi-objective particle swarm optimization algorithm. *EURASIP J. Wirel. Commun. Netw.* **2022**(1), 1–27 (2022)
10. Dixit, A., Mani, A., Bansal, R.: An adaptive mutation strategy for differential evolution algorithm based on particle swarm optimization. *Evol. Intell.* **15**(3), 1571–1585 (2022)
11. Singh, N., Singh, S.B., Houssein, E.H.: Hybridizing salp swarm algorithm with particle swarm optimization algorithm for recent optimization functions. *Evol. Intell.* **15**(1), 23–56 (2022)
12. Teraiya, J., Shah, A.: Optimized scheduling algorithm for soft Real-Time System using particle swarm optimization technique. *Evol. Intell.* **15**(3), 1935–1945 (2022)



# Design and Implementation of Online Adaptive Teaching System Based on Data Clustering Algorithm

Jie Ding<sup>(✉)</sup>

Jilin Agricultural Science and Technology University, Jilin, Jilin, China  
1553859608@qq.com

**Abstract.** Based on the two adaptive modules of automatic grouping of students and automatic grading of subjective answers, the evolutionary grouping algorithm is used to keep the relative consistency between groups at different times. According to the automatic grouping function of students, based on the behavioral sequences of students in different periods, the implicit behavioral patterns in the sequences are mined, and the grouping results are generated once for each period, so that the same group of students have similar behavioral patterns, which can be used as a reference for teachers to make teaching plans for each group. The system performs better than the traditional clustering algorithm in English translation teaching test.

**Keywords:** Adaptive teaching · Data analysis · Clustering algorithm · English Teaching

## 1 Introduction

Adaptive Learning is a research direction combining computer technology and online education. As the name suggests, it is a learning style that varies from person to person. Everyone's qualifications and background are different, and it is difficult to truly teach students according to their aptitude in traditional education. Through the analysis and mining of educational data, adaptive learning recommends personalized learning paths to students, so as to maximize the learning efficiency of each student as much as possible, providing the possibility for the realization of individualized education. One of the key benefits of adaptive learning is its ability to personalize learning. A powerful adaptive learning system can help students personalize learning while evolving to understand the learner better [1].

Adaptive Learning system is a computer-based education system, which can modify the presentation form of textbooks according to the performance of students. An ideal adaptive learning system captures fine-grained data and uses learning behavior analysis to give learners personalized system responses. In this system, students are able to track their own learning, which means they can develop valuable self-management skills and participate in their personal learning progress [2]. The matching learning management system can provide comprehensive administration, documentation, tracking and

reporting progress, and user management. The adaptive learning system can be generally divided into two categories: one is to provide students with the recommendation of learning progress and knowledge points, and the other is to provide teachers with the status inquiry and group recommendation of students [3]. Some teacher-oriented tools, such as Coursera's weekly data report, simply present first-hand data, such as the number of active students and the number of false questions, without deeper analysis and mining. The research of this topic will be carried out on the basis of first-hand data, trying to obtain more meaningful information to help teachers grasp the teaching dynamics [4].

In the teacher side of adaptive learning system, information and data beneficial to students' learning efficiency are also beneficial to teachers. Mastery of current student data allows teachers to see how each student is performing and to keep track of how students are learning in a timely manner, which helps teachers identify students who are not keeping up and are at risk of poor learning outcomes, and provides interventions that are tailored to students' needs and response levels. When teachers query students' learning progress, they can classify students according to knowledge points, which requires the use of knowledge graph. Knowledge graph is a directed graph reflecting the correlation between knowledge points [5]. It can be marked manually or constructed automatically by computer. If the latter is adopted, relevant algorithms involving artificial intelligence and machine learning are needed to mine the knowledge points corresponding to the course content and the relationship between them mainly through natural language processing.

## **2 Structure Analysis of Clustering Algorithm**

### **2.1 Sequence of Behavior**

Action Sequence, as the input of the whole algorithm, usually comes from the students' learning log information collected by online education system. A student's behavior sequence consists of a sequence of events that occur in chronological order. All events in the sequence of any student's behavior belong to a specific finite set of events. According to the tendency of algorithm users to focus on classification, different event sets can be customized and corresponding behavior sequences can be extracted from students' log information [6]. At every moment, each student has a sequence of behaviors, which contains their own behavior pattern. The moment here is artificially defined and generally refers to a complete learning process. In this way, the sequence of behaviors at all times includes all the behavior patterns of the student during the course of the semester [7].

### **2.2 Markov Chain**

In order to mine the behavior pattern contained in the behavior sequence, Markov chain is used to represent the behavior sequence indirectly. Markov Chain is a stochastic probability model, which describes the transition probability of each event in the event set, and the random process composed of these events satisfies the Markov Property. Stochastic processes that satisfy Markov properties are called Markov processes. This random process satisfies the property that the probability of an event occurring at the

next time depends only on what happened at the present time, and has nothing to do with what happened at the past time. Thus, a Markov chain can be described simply by a fixed probabilistic transition matrix in which each element represents the conditional probability of occurrence between two events [7].

The presence of the Stationary Distribution is an important feature of Markov chains and will be used later for normalization. A stationary distribution of a Markov chain is a probability distribution of events that keeps the probability of each event constant. If the stationary distribution is set as the initial distribution of Markov chain, the probability distribution of each event occurring at the initial time can be obtained. With the advance of time, the state transition matrix of Markov chain will act on the probability distribution of each event to get the probability distribution of the later moment, and the probability distribution of each moment will remain the same as the initial moment. Of course, a stationary distribution satisfies all the conditions for being a probability distribution, including the sum of all the probabilities being 1 [8].

### 2.3 Inverse Similarity Matrix

Similarity is an index measuring the degree of similarity between two targets. The greater the similarity, the more similar the two targets are. The Reversed Similarity of the index is used to express the degree of similarity between Markov chains. In other words, the similarity of the two Markov chains is inversely proportional to the value [9].

Think of each Markov chain as a point in higher dimensional space, then the inverse similarity between them is actually the distance between the corresponding points. The most common way to measure the Distance between geometric points is Euclidean Distance. In addition, since the Markov chain can be approximately regarded as a probability distribution, the distance of the Markov chain can also be measured by the index measuring the distance of the probability distribution. There are some ways to measure the Distance of probability distribution including JS Jensen-Shannon Divergence and Hellinger Distance.

$$d_{KL}(P\|Q) = \sum_i p_i \log \frac{p_i}{q_i} \quad (1)$$

$$d_{KL}(P\|Q) = \frac{1}{\sqrt{2}} \sqrt{\sum_i (\sqrt{p_i} - \sqrt{q_i})^2} \quad (2)$$

## 3 k-Means Clustering Algorithm

In the AFFECT evolution algorithm, a static clustering algorithm needs to be selected to perform the actual clustering process. Mainstream clustering algorithms can be roughly divided into two categories: the number of clusters that need to be specified and the number of clusters that do not need to be specified. Algorithms that do not require a specific number of clusters, such as Hierarchical Clustering, cluster targets into different groups at different levels through a top-down or bottom-up clustering approach. In the grouping problem of students, the algorithm of specifying the number of clusters is



adopted uniformly, because the actual data without specifying the number of clusters is not good, and the actual data is noisy due to its complexity, so it is difficult for the algorithm to automatically find the appropriate number of clusters [10].

K-means clustering algorithm is an algorithm that is more suitable for clustering and grouping points in high-dimensional space among the clustering algorithms that specify the number of clusters. This algorithm is used for static clustering. K-means belongs to the unsupervised learning algorithm in the field of machine learning, that is, it does not need to mark the category of each point in advance to train the model. k in k-means means the number of clusters specified manually. The goal of the algorithm is to assign each point to be classified into one of the k categories, so that the distance between each point and the cluster center of the cluster it belongs to is the minimum distance between it and the cluster center of all the clusters. The architecture of the skip-gram model is shown in Fig. 1.

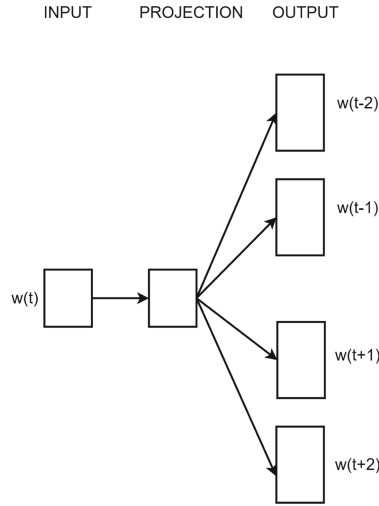


Fig. 1. Architecture diagram of skip-gram model

The CBOW model is very similar to the skip-gram model. It also needs to train a shallow neural network with only one hidden layer, but the assumed task is opposite to the skip-gram model: given the context word of a word, predict the probability distribution of the word as before, need to specify a window size to extract the word pair, the weight vector of the corresponding position in the hidden layer is represented as the word vector. This is also the case with skip-gram. Slightly different from the skip-gram model, the word vector of the CBOW model reflects the features of words as contexts, while the skip-gram model highlights the features of words as central words.

### 4 Implementation Method

Common methods for analyzing serialized data and detecting “similar behavior patterns include Sequence Mining, Hidden Markov Model (HMM) and so on. These methods

can effectively recognize and classify students' behavior sequences at a fixed moment, but they do not refer to the grouping results of the past moment as part of the grouping results of the present moment. Therefore, the grouping results obtained in this way tend to have strong inconsistency at different times, leading to the mutual isolation of grouping results in different periods and limited reference value. This is also the case with skip-gram. Slightly different from the skip-gram model, the word vector of the CBOW model reflects the features of words as contexts, while the skip-gram model highlights the features of words as central words.

The pipeline of evolutionary clustering algorithm is a combination of a series of algorithms. The input is the behavior sequence of students and the output is the grouping result of each moment. The core of the pipeline is the evolutionary clustering algorithm AFFECT. Different from the common clustering algorithm, AFFECT algorithm obtains the grouping result at every moment, referring to the grouping result at the past moment. Therefore, the resulting grouping result can be viewed as a weighted average of the historical grouping and the grouping based only on the current student behavior data. In this way, the grouping results at each moment are interrelated rather than isolated, so as to ensure the relative consistency of grouping results.

Generating Markov chain corresponding to each student's behavior sequence is the first step of the whole algorithm. A sequence of actions is a chronological list of all the actions of a student in a given class (moment). The category of behavior is determined by the user of the algorithm depending on its grouping emphasis. For example, if a user of the algorithm wants to group all students according to their learning ability, he can choose a sequence of behaviors such as: correct, correct, wrong, correct... The sample sequence, with only two categories of behavior, represents a sequence of right and wrong answers by a student on a quiz question in the course.

Markov chains are a common way to model serialized data because serialized data tends to be Markov, that is, the probability of an event occurring at the current moment depends only on what happened at the previous moment, not on what happened in history. This is also the case with skip-gram. Slightly different from the skip-gram model, the word vector of the CBOW model reflects the features of words as contexts, while the skip-gram model highlights the features of words as central words. In the student grouping problem, the behavior sequence of students also satisfies the Markov property. For example, in the sequence of students' answering behaviors mentioned above, there are only two kinds of events: [correct] and [wrong]. With the advance of time, the events are constantly transformed between [correct] and [wrong]. If a student has strong learning ability, it can be considered that the probability of [correct] event transferring to [correct] event is higher, while the probability of other event combination transferring is lower. On the contrary, if a student's learning ability is not strong, he is likely to give wrong answers frequently, so the probability of [wrong] event being transferred to [wrong] event is much greater than that of the student with strong ability, and the probability of [right] event being transferred to [right] event is significantly reduced. In this way, on the premise that the correct or wrong of the current answer has nothing to do with the historical correct or wrong events, but is only related to the correct or wrong events of the previous moment, that is, on the premise that the sequence of answer behavior meets the Markov property, different Markov chains can be constructed to distinguish

the strength of students' learning ability. Therefore, the constructed Markov chains can be further used in the subsequent clustering algorithm.

## 5 Conclusions

Starting from the teacher end of adaptive learning system, this paper discusses the method of mining education data by computer technology around two applications. In the test of "Grouping Students Based on Behavior Sequence", the method of grouping English translation students by using the behavior data easily obtained from online education system is proposed. Compared with the traditional static clustering algorithm, the evolutionary grouping algorithm used can explore the evolution of grouping conditions while maintaining the relative consistency of groups at different times. In order to verify the performance of the algorithm, two sets of experiments were conducted on simulated data and real data. In the simulation data experiment, the maximum power matching method of bipartite graph is proposed to find the corresponding relation of different time group numbers. In the real data experiment, a method is proposed to use the average static distribution of the group to express the characteristics of the group and judge the good or bad result of the group. Through these two groups of experiments, it is verified that the evolutionary grouping algorithm can indeed outperform the traditional clustering algorithm in grouping results.

**Acknowledgements.** Analysis on the Types and Meanings of Common Intercultural Information in English News. YouthFund Project of Jilin Agricultural Science and Technology College, JLASTU Contract No. 2017 [228].

## References

1. Ayanian, J.Z., Markel, H.: Donabedian's lasting framework for health care quality. *N. Engl. J. Med.* **375**, 205–207 (2016)
2. Hulton, L., Matthews, Z., Bandali, S., Izge, A., Daroda, R., Stones, W.: Accountability for quality of care: monitoring all aspects of quality across a framework adapted for action. *Int. J. Gynecol. Obstet.* **132**, 110–116 (2016)
3. Akachi, Y., Kruk, M.E.: Quality of care: measuring a neglected driver of improved health. *Bull. World Health Organ.* **95**, 465–472 (2017)
4. LeFevre, A., et al.: Program assessment of efforts to improve the quality of postpartum counselling in health centers in morogoro region, Tanzania. *BMC Pregnancy and Childbirth*, 18, Article No. 282 (2018)
5. Tripathi, V., Stanton, C., Strobino, D., Bartlett, L.: Measuring the Quality of Maternal and Care Processes at the Time of Delivery in Sub-Saharan Africa: Development and Validation of a Short Index. *BMC Pregnancy and Childbirth*, 19, Article No. 133 (2019)
6. Koblinsky, M., et al.: Quality maternity care for every woman, everywhere: a call to action. *The Lancet* **388**, 2307–2320 (2016)
7. Afsana, F., Kabir, M.A., Hassan, N., Paul, M.: Automatically assessing quality of online health articles. *IEEE J. Biomed. Health Inform.* **25**, 591–601 (2020)
8. Ali, J., Perumal, S., Shaari, H.: Application of the stimulus-organism-response model in the airline industry: examining mediating role of airline image in repurchase intention. *Int. J. Supply Chain Manage.* **9**, 981–989 (2020)

9. Perumal, J.A.D.S., Shaari, H.: Mediating role of airline image on repurchase intention: does word of mouth matter? *Hamdard Islamicus* **43**, 1919–1930 (2020)
10. Bhandari, P., Amponstira, F.: Model of entrepreneurial orientation, competitive advantage and performance of women-owned enterprises in Gandaki Province, Nepal. *Open J. Bus. Manage.* **9**, 2854–2865 (2021)



# Design and Implementation of Online Teaching System Based on J2EE

Qingfeng Guo<sup>(✉)</sup>

Teacher's College, Beihua University, Jilin, Jilin, China  
1547717173@qq.com

**Abstract.** In order to solve the problems of online teaching system failing to meet the requirements of online distance education for processing massive teaching data and limited network bandwidth, the existing system is optimized based on J2EE (Java 2 Platform, Enterprise Edition) by adopting mature cloud computing and load balancing algorithm. Create standard reusable module components and build a hierarchical structure that can automatically process programming to meet the needs of teachers and students for online learning systems. After the ideological and political course teaching test, the system operation meets the design requirements.

**Keywords:** Online Teaching · Load Balancing Algorithm · J2EE · Ideological and Political Education

## 1 Introduction

Online teaching system is a kind of long-distance teaching as the purpose, with the Internet communication ability to provide teaching services, in order to achieve the maximum application of education and teaching resources not affected by time and space. In terms of the practical significance of the development and application of online teaching system, the following performances are more outstanding [1].

As a new teaching mode, online teaching mode can be used as an effective complement to traditional teaching mode, and the ability to organic combination with traditional teaching mode, thus make up and solve in the traditional teaching pattern of mainly teachers teach such as low interactivity between teachers and students, cannot let the student to participate effectively in, is difficult to realize the personalized education, etc. [2].

Online teaching system is relatively flexible in use and less affected by time and space factors in teachers' teaching and students' learning. According to the requirements of the content of the course and the actual situation of the students' learning, besides the traditional one-to-many teaching mode, we can also choose one-to-one and many-to-many and other forms of course teaching [3]. At the same time, the network environment can allow at any time. The design and implementation of the system can not only meet the current social demand for the use of online teaching system, but also in the theoretical aspect of advanced computer and Internet technology education and teaching application, as well as the online teaching business in the mass data and information processing provides an effective solution direction, so it has a very positive theoretical significance.

## 1.1 Abundant Teaching Resources

With the help of the current relatively advanced computer and Internet technology, the current limited educational resources can be integrated to the maximum extent, and resources can be shared in the form of video, micro-class or live broadcast, and the current relatively cutting-edge big data analysis technology can be used to systematically and visually analyze the teaching situation of teachers, learning situation of students, and educational information and resources. And continue to improve and optimize the teaching forms and methods [4].

This paper takes the online teaching mode, which attracts more attention in the education field, as the research object, attempts to design and implement the online teaching system by using the current relatively mature software engineering technology, and uses the more advanced cloud computing technology and load balancing optimization algorithm to solve the problems of low data processing capacity of the traditional online teaching system and the failure of network bandwidth resources to meet the use requirements [5].

## 2 Analytical Modeling Tool

### 2.1 BPMN Business Process Modeling

BPMN is a standardized business process modeling language, and its application is an important step to reduce the many existing business modeling tools and markup gaps. In BPMN, pools and lanes are modeled. A business is represented as a pool, and organizational units within a business are represented as swim lanes. Interactions between pools are limited to the delivery of messages, regardless of each other's internal processes. Therefore, in the analysis of the business process of the online teaching system based on J2EE in this research, the business transfer and operation process between the user roles of the system can be established and optimized accurately and effectively, which is more intuitive and inclusive than other modeling tools such as UML activity diagram. Therefore, this paper adopts the BPMN method to model and analyze the system business process and function in the stage of system business process analysis [6, 7].

### 2.2 UML Unified Modeling Language

In the design and development of object-oriented software engineering system, it is necessary to analyze and design various objects effectively. Therefore, it is very important to establish different models for various objects, and then analyze and design the objects in the program. UML modeling technology is such a graphical universal modeling language. Its principle is to construct various models of the whole system in the form of model elements. These model elements include the various classes defined in the system and their relationships, the dynamic changes generated by the interaction between the instantiated classes, and so on. Among all kinds of system analysis methods, UML modeling technology has great advantages and good universal applicability. It has gradually become the mainstream system analysis method in software engineering [8].

### 3 J2EE Development Technology

The words Java will be familiar to any programmer or IT practitioner. It is an efficient, elegant, open-source programming language with excellent cross-platform properties. It is widely used in the realization and development of various applications, industrial control programs, website Web services and mobile apps. J2EE is Java enterprise-level distributed application development solution, is a set of widely recognized program development specifications and standards. J2EE is a relatively complete set of Service, API application interface and related protocols, mainly used in the development of Web program system. It mainly includes JSP dynamic web page technical standard; JDBC database access standard and protocol; EJB distributed application component technology and specification; XML, Extensible Markup language and other special technical specifications [9]. Typically, an application developed through J2EE consists of multiple layers of components, such as a client-oriented layer for user operations; A system layer for enterprise information or data; and enterprise-oriented business logic layer.

#### 3.1 SSH Framework Technology

In J2EE enterprise-level application development, framework technology is used to reduce the workload of programming development, the difficulty of late development and maintenance, and the ability of local system upgrade and expansion. In this paper, SSH framework (40) is used to implement the technical system [10]. The SSH framework is the integration of three open-source application frameworks, all of which are typical development frameworks of J2EE frameworks.

#### 3.2 Struts2 Framework

Struts2 framework is a Java EE network programming framework, developed from the Struts framework. Struts is based on the MVC programming idea and improves the efficiency of Internet application development with the MVC programming model. The emergence of Struts2 framework not only effectively inherits various advantages of Struts framework in design and use, but also puts forward a new MVC network application development framework. Through a large number of applications of interceptor mode, it realizes the framework module integration mode similar to Web Work framework [11].

#### 3.3 Spring Framework

The spring framework is a lightweight development framework. It is so powerful that you can use it to load Bean objects in server programs, which brings many advantages to server-side programming, such as not having to initialize the Bean class with the new keyword every time.

#### 3.4 Hibernate Framework

Hibernate is a new object relational mapping mechanism, which is designed to solve the matching problem between objects and relations in object-oriented programming.

Through coordination mechanism, the code written by programming language is completely isolated from the program database, so that code writers do not need to consider the types of background database when carrying out program development. Considering only the operation of data table can greatly improve the efficiency of programming development.

## **4 Cloud Computing Technology Application**

Cloud computing is favored by many people for its high-cost performance, scalability and flexibility. Compared with the general Internet application technology mode, the virtualization technology of cloud computing is the most significant one among all its features. It is not bound by time and space, and can virtualize applications and resources and then provide them to users. Through the virtual platform in the cloud computing service, users can operate the corresponding terminals to complete the expansion, migration, data backup and other actions without touching the physical space of cloud computing. In addition, cloud computing has strong distribution, sociality and diffusion. As a method of sharing infrastructure, cloud computing can connect massive systems together to form a system pool, and then provide various Internet application services.

### **4.1 Multi-service System Heterogeneity Support**

The heterogeneity of the business system means that the business is not unified, nor is it manually defined before use, but is set by the user on the basis of use needs. On most cloud computing platforms, multiple business types can run freely at the same time, which is one of the main differences between cloud computing and traditional grid computing.

### **4.2 Dynamic Scalability Support for Resources**

Cloud computing can support the dynamic scaling of its resources and realize the network redundancy of its basic resources. This feature ensures that the addition, modification, or deletion of any resource node in the cloud computing service does not adversely affect other nodes. At the same time, the resource node will not cause service interruption or user data loss in the entire cloud computing service environment if it encounters various exceptions such as hardware faults. The preceding resource nodes include storage nodes, computing nodes, and network nodes. In addition, the cloud computing service can support the dynamic transfer of its resources to realize the dynamic scheduling of its resources. According to the changes of the overall business requirements of the cloud computing service, the idle resources can be transferred to the location with high demand, thus improving the carrying capacity of the entire cloud computing platform. When the overall business demand of the cloud computing platform decreases, idle resources can be pooled to run in energy-saving mode, so as to reduce the operating costs and energy consumption of the platform and achieve certain environmental protection effects.



### 4.3 On-demand Billing Support

The cloud computing service platform supports the dynamic transfer of its resources. The external feature of this feature is the on-demand resource allocation support. Using graph virtualization technology, the computing resources in cloud computing services can be homogenized and quantified to realize the dynamic allocation of computing capacity and resource utilization. On the basis of on-demand distribution, quantified resources are priced and then charged according to the amount of users' usage, which has become a prominent feature of cloud computing service platform. The working flow chart of cloud computing service platform is shown in Fig. 1.

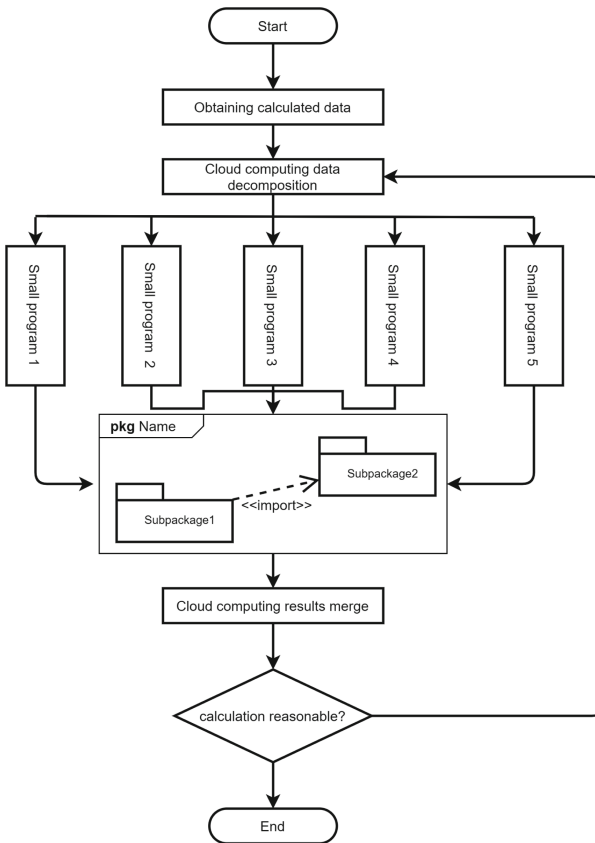


Fig. 1. Working flowchart of cloud computing service platform

Most online teaching work is carried out in a unified or relatively centralized time. Therefore, for online teaching system, the number of visits will surge in a short time, and the request for reading and calling teaching resources will be centralized. After the cloud computing distributed optimization design of the online teaching system based on J2EE, there may also be uneven load distribution among the nodes of the cloud computing

cluster, service request blocking, computing resource waste and other situations, so it is necessary to further improve the load balancing design.

The most commonly used load balancing algorithms are static and dynamic. These two load balancing algorithms are suitable for different situations, but overall dynamic load balancing algorithm has more outstanding resource allocation ability and good performance in terms of flexibility, so it is more respected by the majority of program developers, and relatively more applications. Among dynamic load balancing algorithms, Pick-KX algorithm is a representative one. The algorithm is implemented in a clustered server environment as shown in Eq. (1).

$$S = \{S_i, i = 1, 2, \dots, n\} \quad (1)$$

$$P_j = \frac{X_i}{\sum_{i=1}^i L_i} \quad (2)$$

$$L_{total} = \sum_{i=1}^i L_i \quad (3)$$

$$X_i = \sum_{i=1}^i L_i \quad (4)$$

The running cycle of a cloud computing cluster can be expressed. The load capacity of each node in the cluster can be expressed in terms of. On this basis, certain conditions can be met before the Pick-KX algorithm can be used to allocate server resources, as shown below.

## 5 Conclusions

Taking online distance teaching mode as the research object, this paper designs and implements an online teaching system based on J2EE based on popular software engineering technology. At the same time, in view of the phenomenon that the traditional online teaching system cannot meet the current online distance education business to deal with massive teaching data, and the network bandwidth resources cannot meet the increasing demand of online students, try to use the current popular cloud computing technology and load balancing algorithm to design and optimize the system. To the maximum extent to meet the ideological and political teachers and students on the online teaching system needs.

Analysis and modeling tools, system development technology and cloud computing technology related to the design and implementation of online teaching system based on J2EE are analyzed to provide technical support for the design and implementation of the system. In order to solve the problem that the traditional online teaching system cannot meet the current online distance education business to deal with massive teaching data, and the network bandwidth resources cannot meet the increasing demand of online students, the load balancing algorithm is optimized on the basis of cloud computing

technology, and the teaching service providing ability of the online teaching system based on J2EE is improved. On the basis of the above cloud computing framework and load balancing optimization algorithm, from the perspective of actual use, the requirements of the online teaching system based on J2EE are analyzed. On the basis of the analysis of system requirements, the overall architecture of the system, the system function module, and the database structure design.

## References

1. Yang, D., Xue, K., Zhai, H., et al.: A dual-band circularly polarized planar monopole antenna for Microwave Access applications. *Microwave Optical Technol. Lett.* **2019**(3), 172–180 (2019)
2. Kunwar, A., Kanaujia, B.K., et al.: Inverted L-slot triple-band antenna for WLAN and WiMAX applications. *Int. J. Microwave Wirel. Technol.* **2015**(5), 230–241 (2015)
3. Liu, Z.D., Weng, L.B., Fang, H.X., Wang, S.C.: A comparison of ionosonde measured predictions over China. *Adv. Space Res.* **65**, 1936–1946 (2019)
4. Xia, W.J., Zexin, L., Huang, Y.Q., et al.: Manifold and graph integrative convolutional network for CT reconstruction. *IEEE Trans. Med. Imaging* **2021**(3), 124–130 (2021)
5. Nanema, H., Zerbo, L.J., Konate, K., Ouattara, H.: Layer peak parameter effects on critical frequency. *J. Sci. Eng. Res.* **7**, 52–58 (2018)
6. Ao, Z., Hewei, G., Li, Z., Yuxiang, X.: A dual-domain deep learning-based reconstruction method for fully 3D data helical CT. *Phys. Med. Biol.* **13**, 25–31 (2020)
7. Ge, Y.H., Su, T., Zhu, J.G., Deng, X.L., et al.: Deep computed tomography reconstruction network with analytical knowledge. *Quantitative Imaging Med. Surg.* **2020**(2), 521–530 (2020)
8. Nanema, H., Zerbo, B.L., Konate, K., Ouattara, H.: Height of peak parameter effects on critical frequency. *J. Sci. Eng. Res.* **5**, 61–67 (2018)
9. Fu, J., Dong, J.B., Zhao, F.: A deep learning reconstruction framework for differential phase-contrast tomography with incomplete data. *A Publication IEEE Signal Process. Soc.* **1**, 152–160 (2020)
10. Nguyen, T.T., Nguyen, N.D., Nahavandi, S.: Deep reinforcement learning for multiagent systems. *IEEE Trans. Cybernetics* **9**, 350–360 (2020)



# Design and Implementation of Classroom Teaching Line Analysis Platform Based on Pose Recognition Algorithm (MCG-CPR)

Xiaobo Sun<sup>(✉)</sup>

Art College, Northeast Electric Power University, Jilin, China

895442004@qq.com

**Abstract.** A teaching behavior analysis model based on dual video streams-DVS-TBA is proposed, which realizes information integrity by constructing non-switching dual-channel video, and realizes classroom behavior recognition based on body features and image features, so as to effectively improve the accuracy of classroom teaching behavior recognition, and realize the description and analysis of classroom teachers and students' behaviors in the whole time series and samples. The art history teaching process test shows that the model can evaluate the teaching effect effectively.

**Keywords:** Classroom teaching · Behavior analysis · MCG-CPR · Teaching effectiveness evaluation

## 1 Introduction

With the continuous breakthrough of artificial intelligence theory and technology in the field of computer vision, image analysis using computer vision model becomes faster and more accurate, and many computer vision models have reached the efficiency of “second results” [1]. It is found that some computer vision technology can be used in classroom teaching behavior analysis. The object detection algorithm can help the system quickly and accurately locate the location of the object in the classroom video. The human body pose estimation technology can help the system to obtain the human body key points in the video, and provide the human body structure map features for the classification of human behavior pose. Image classification technology can help to systematically infer the classroom behavior of the subject in the classroom video; Face recognition technology can help the system to determine the identity of the current subject [2].

In order to build a classroom teaching behavior analysis system that can completely use teaching information to analyze classroom and individual teaching behaviors without personnel intervention, target detection technology, human body pose estimation technology, image classification technology and face recognition technology are integrated. Firstly, the above computer vision technology is used to obtain the behavior and posture information of each person in the current sample image [3]. Then, the S-T teaching analysis method is used to summarize the behavior information of all subjects, deduce

the current classroom teaching behavior, and construct the S-T teaching behavior code. Finally, the model of classroom teaching is obtained through S-T teaching behavior coding. This method can not only achieve objective and effective classroom teaching behavior analysis, but also provide different benefits for students, teachers and schools [4]. In the classroom scene, there is a big difference between the number of teachers and the number of students. As a result, during the teaching process, teachers cannot pay full attention to the classroom and cannot correctly judge the learning quality of students. Such a situation will not only reduce students' interest in their studies and discourage their learning motivation, but also make teachers misestimate the quality of class, which is not conducive to teachers' reflection on classroom teaching. For students, through the system, teachers can obtain the classroom behavior of each student in class, which provides part of the basis for teachers to provide personalized guidance to students and help improve the learning state of students [5].

For teachers, it can help teachers to reflect on teaching. In classroom teaching, no one pays attention to teachers' teaching behavior. Teachers can only analyze themselves according to their self-cognition, which is highly subjective. This will lead to the teacher's reflection in classroom teaching is blocked, teaching ability improvement is limited [6]. Therefore, for teachers, statistics, induction and analysis of behaviors in classroom teaching scenarios and analysis of classroom teaching models can help teachers obtain real classroom teaching behaviors, conduct classroom reflection, improve teachers' professional quality and optimize the level of classroom teaching.

For the school, it can help the school to obtain the teaching evaluation results conveniently. In the teaching evaluation, the teaching evaluation of teachers in traditional schools mainly consists of questionnaires and expert evaluation. This method is time-consuming and laborious. It is a rational calculation under the guidance of interests, and the evaluation of teachers is not objective, true and comprehensive. Using computer instead of manual to make statistics on teachers' usual performance saves time and effort, solves the problem of strong subjectivity of traditional evaluation methods, and improves the objectivity of teaching evaluation [7].

## **2 Analysis of Classroom Teaching Behavior Based on Video**

With the popularization of video equipment in school classrooms, the cost of obtaining classroom teaching videos has been effectively reduced, which provides a friendly way for education and teaching researchers to obtain data. At the same time, with the rapid development of artificial intelligence technology, great progress has been made in image and video processing and classification, and video-based classroom teaching behavior analysis research has been effectively carried out. At present, video-based teaching behavior analysis general process. According to the teaching behavior standard of expert knowledge, the automatic discriminant function of classroom teaching behavior is constructed. The artificial intelligence technology is used to sample and analyze the teaching videos and extract the effective teaching behavior features, and the discriminant function of classroom teaching behavior is used to code the current sample of classroom teaching behavior. Through the induction statistics of all the classroom teaching behavior code to construct the teaching behavior sequence, deduce the classroom teaching model [8].

The video-based classroom teaching behavior analysis method solves the interference of traditional teaching behavior analysis method to teaching and improves the objectivity of teaching behavior analysis. However, the existing teaching behavior analysis system mostly uses multi-channel camera to switch videos, which leads to some loss of video information, resulting in incomplete analysis of classroom teaching behavior, and the accuracy of classroom teaching behavior identification needs to be improved. Aiming at the shortcomings of the existing teaching behavior analysis system, a teaching behavior analysis model based on double video streams is proposed. This model collects classroom teaching scenes in a non-switching way through the dual cameras located at the front and back of the classroom [9]. At the same time, according to the frame rate sampling of the camera video stream, the characteristics contained in the teaching video are fully used to analyze the classroom teaching behavior, so as to improve the accuracy of the analysis of classroom teaching behavior. The human body video conversion formula of classroom teaching is shown as (1) and (2).

$$S^t = \rho^t(F, S^{t-1}, L^{t-1}), \forall t \geq 2 \quad (1)$$

$$L^t = \Phi^t(F, S^{t-1}, L^{t-1}), \forall t \geq 2 \quad (2)$$

### 3 Face Recognition Algorithm

In order to determine the behavior of each subject in the classroom, the identity of the subject needs to be confirmed, therefore, the use of face recognition as the subject identity confirmation tool. The general step of face recognition consists of face detection, image preprocessing, face representation algorithm and face feature comparison algorithm. When face registration, the face region in the image is divided by face detection, and the face in the image is adjusted to the unified standard image by image transformation using image preprocessing [10]. Finally, the discriminative features in the image face are extracted and saved by the face representation algorithm, and the face feature database is constructed for matching recognition. When face recognition is needed, the face features in the image to be verified are extracted through face detection, face preprocessing and face representation algorithm, and then the similarity between the face features to be verified and the features in the face feature database is calculated using the face feature comparison algorithm, that is, the probability that the face to be verified belongs to the subject of each face. The main identity of the feature with the greatest similarity is the identity of the face to be verified. When the maximum probability is greater than the Kan value, it indicates that the face belongs to the subject; otherwise, it indicates that the subject to which the face belongs is not registered.

### 4 Real-Time Video Streaming Protocol

Digital video data is transmitted in the network mainly through downloading and streaming. To download a video, you need to copy the complete video data to the local computer and play it through the video decoder. Streaming is the core technology of streaming

media application, which can broadcast video, audio and other data side by side. The video streaming analysis application is a kind of streaming media application. At present, streaming media applications are divided into client and server according to different application functions. The client is used to request and display the streaming media data; the server is used to store and manipulate streaming data and respond to client requests. Among them, the transmission of streaming media data between client and server needs to rely on streaming transmission. At present, streaming transmission generally uses the real-time transport protocol RTP of the transport layer.

RTP is a transport layer protocol used by IETF's Audio and Video Transmission Working Group to transmit streaming media data over IP networks. It is used to solve the problem that the arrival time of streaming data cannot be predicted during data transmission. Generally, end-to-end systems involving streaming media need to use RTP protocol, such as telephone communication and live broadcast. RTP is usually run based on UDP to ensure speed. RTP is generally divided into RTP data Protocol and RTCP (RTP Control Protocol), which is used to ensure the quality of streaming media data transmission. The Bayesian function discrimination model is shown in Fig. 1.

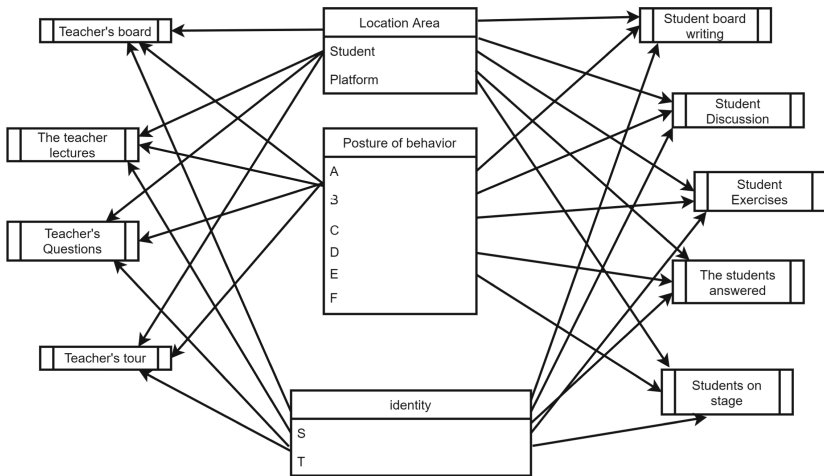


Fig. 1. Bayesian function discriminant model diagram

### 5 Multi-feature Classroom Behavior Posture Classification Based on Attention Mechanism

In order to encode classroom teaching behavior through the non-switching dual-channel video stream, it is necessary to obtain the behavioral posture features contained in the classroom teaching video stream and confirm the identity of the subject corresponding to the behavioral features, that is, the classroom subject identity - behavioral posture features. How to accurately judge the behavior of teachers and students in the video

when the background of teaching environment video is complex and the main body of the classroom is blocked.

To solve these two problems, an attention-mechanism based multi-feature classroom gesture recognition algorithm (MCG-CPR) is proposed to replace the traditional image classification algorithm. The algorithm solves the problems such as the complex background of teaching environment videos and improves the accuracy of human behavior posture recognition in class. Meanwhile, a multi-model classroom subject identity-behavioral posture representation model (M-CPIE) is proposed to quickly and effectively acquire the characteristics of classroom subject identity-behavioral posture. This method can be divided into two stages: image segmentation stage and image analysis stage. In the image segmentation stage, the classroom subject identity object detection algorithm (FP-YOLO) is used to locate the human body and face area in the sample image. In the image analysis stage, firstly, the segmented human body region image and face region image are transferred to the classroom behavior gesture recognition model and the subject identity recognition model respectively to obtain the regional human body action and regional face identity. Finally, based on the positioning information of image segmentation, through the multi-model feature fusion method, the face identity features and human action features are fused to output the class subject identity and behavior posture features.

## **6 Teaching Behavior Analysis Model Based on Double Video Streams**

A multi-feature classroom subject identity-behavior and pose representation model (M-CPIE) was constructed with FP-yoke detection algorithm, face recognition algorithm and MCG-CPR multi-feature classroom behavior and pose classification algorithm based on attention mechanism. The model can quickly and accurately extract the information of subject identity and behavior posture used for teaching behavior analysis in classroom teaching videos. Looking at most instructional behavior analysis methods, they can analyze and encode instructional videos such as open class, lecture and micro class. The video resources analyzed above need to be recorded by professionals, which costs human and material resources. At the same time, they have a relatively obvious emphasis on the subject, which generally focuses on teachers, that is, teachers and students cannot appear in the teaching video completely at the same time. The behavior information of students and teachers in the classroom teaching video is lost, which leads to the deviation of the analysis results of classroom teaching behavior. The teaching behavior analysis model locates in the simple and complete analysis of classroom teaching behavior, and constructs the teaching behavior analysis model based on dual video streams (DDS-TBA). It uses dual cameras at the front and back of the class to record the classroom instruction, without the need for professional recording (just make sure the cameras are working). At the same time, all scenes of the classroom can be obtained at the same time, which solves the problem of serious loss of video recording information. In the process of analysis, the model firstly obtains the double video stream samples by analyzing the video stream and preprocesses the sample data. Then, M-CPIE model was used to analyze the dual video streams to obtain the complete information of class



subject identity and behavior posture. Finally, through the statistical induction of the information of subject identity and behavior posture, based on S-T analysis method, the classroom teaching behavior is coded, the sequence of classroom teaching behavior is constructed, and the model of classroom teaching behavior is inferred.

At present, the recording of classroom teaching videos requires the intervention of professionals. At the same time, the recorded videos have obvious information loss due to different concerns. Therefore, it is proposed to use the surveillance cameras located at the front and back of the classroom to record classroom teaching, so as to achieve objective and complete teaching video recording with almost no manpower cost. In order to reasonably use the information in the dual video stream, the video stream of the camera is firstly analyzed to obtain the time information and image data contained in the video stream. Then, using the time information and image data through time alignment, sample fusion to build the fusion image features. After using M-CPIE to extract the features, the subject-behavior posture data of S-T teaching analysis was generated by the class-subject identity-behavior posture information segmentation algorithm of dual video streams.

## 7 Conclusions

The classroom teaching behavior analysis system based on artificial intelligence technology samples the teaching video with a fixed frequency, analyzes the sample image to obtain the way of classroom teaching behavior, codes the classroom behavior, and determines the classroom teaching mode. This method makes the analysis of classroom teaching behavior simple and objective. However, most of the existing analysis methods use multi-channel video switching, which leads to some loss of video information, resulting in incomplete analysis of classroom teaching behavior, and the identification accuracy of classroom teaching behavior needs to be improved. To solve the above problems, this paper puts forward the DVS-TBA model, and constructs the teaching behavior analysis system based on double video streams. This paper solves the problems of the current teaching behavior analysis system in data acquisition and video sampling, integrates the current excellent artificial intelligence technology, proposes a convenient and efficient teaching behavior analysis method based on double video streams, and realizes the teaching behavior analysis system based on double video streams by using the middleware such as message queue.

## References

1. Li, L., Wang, W., Wang, B., Wang, Z.Y., M.X., Han, J.G.: Active modulation of refractive index by stress. *Appl. Optics* **53**, 6335–6351 (2013)
2. Ledig, C., Huszar, F., Caballero, J., Acosta, K., Aitken, A.: Photo-realistic single image super-resolution using a generative adversarial network. In: *Conference on Computer Vision and Pattern Recognition*, pp. 107–115 (2017)
3. He, P., Wang, L., Du, F., Kiran, R., Yang, M.: Additive manufacturing and mechanical performance of trifurcated steel joints. *Mater.* **2020**(9), 123–131 (2020)

4. Shi, W., Caballero, K., Huszar, H., et al.: Real-time single image and video super-resolution based on convolutional neural network. In: *Computer Vision and Pattern Recognition*, 27–30 June 2016, pp. 184–188 (2016)
5. Du, W., Wang, Y., Wang, H., Yang, J.: Novel machine learning approach for shape-finding design of tree structures. *Comput. Struct.* (2022), 7, 45–51
6. Long, Z.Y., Wang, Y., et al.: Image super-resolution based on a deep convolutional neural network. *Appl. Opt.* **58**, 271–278 (2019)
7. Cornia, M., Cucchiara, R.: Meshed-memory transformer for image captioning. *IEEE Computer Society Conference*. **13**, 251–260 (2020)
8. Lu, H., Mao, B., Liu, B.: Mathematical degradation model learning for image super-resolution. *IEEE Access* **7**, 128–139 (2021)
9. Brown, B., Mann, T., Ryder, M., et al.: Language models are few-shot learners. *Neural Inf. Process. Syst.* **13**, 531–543 (2020)
10. Hu, M., Su, T., Li, H., et al.: Adaptive cross-stitch graph convolutional networks. *ACM Multimedia Asia* **7**, 142–150 (2021)



# Double Attention Mechanism Text Detection and Recognition Based on Neural Network Algorithm

Yongjun Qi<sup>1,2</sup>, Hailin Tang<sup>1,2(✉)</sup>, and Li Huang<sup>1</sup>

<sup>1</sup> Faculty of Megadata and Computing, Guangdong Baiyun University,  
Guangzhou 510450, Guangdong, China  
linht88@163.com

<sup>2</sup> Mongolian University of Science and Technology, Bayanzurkh District, Ulaanbaatar 13341,  
Mongolia  
qyj120040878@126.com

**Abstract.** How to effectively identify these signals and data has become an urgent topic. The neural network model is a stochastic system composed of nonlinear neurons. Therefore, it has strong self adaptability and controllability. This paper proposes a method based on training samples. It classifies the original continuous text through the artificial neural network algorithm. This paper mainly uses experimental method and comparative method to analyze the accuracy, precision, recall rate, F value and its trend in training and the results under different models. The experimental results show that good results have been achieved on the IMDB comment dataset, and the accuracy rate is close to 89.4%.

**Keywords:** Neural Network Algorithm · Double Attention Mechanism · Text Detection · Recognition System

## 1 Introduction

As a new computer aided learning method, neural network algorithm has been paid more and more attention. The adaptive training process based on neural network is to establish a mapping relationship from the input level to the output level and then to the target detection level to effectively process and predict the noise. It can implement both functions at the same time. Dual core intelligent text number recognition technology is based on artificial intelligence. In practical applications, it can realize automatic sorting, positioning and tracking of objects. At the same time, a large number of complex problems can be transformed into a few representative models and simplified.

There are many theories about neural network and attention mechanism text detection and recognition. For example, some proposed a neural network detection algorithm based on attention mechanism. A new training learning vector is introduced into the bidirectional neural network of the gate ring unit, which further improves the recognition and classification ability of the neural network [1, 2]. Other researchers use the double door recurrent neural network layer to obtain text information and extract functions,

The original version of this chapter was revised: The affiliation of the author “Hailin Tang” has been updated. The correction to this chapter is available at

[https://doi.org/10.1007/978-981-99-2092-1\\_131](https://doi.org/10.1007/978-981-99-2092-1_131)

and introduce attention mechanism level to enhance the ability to capture semantic and logical relationships between words [3, 4]. Other scientists assign the attention weight of the feature vector through the attention mechanism to focus on the feature vector with higher weight. The extracted attention features and prediction information of previous time are used as inputs to nested short-term and long-term memory networks [5, 6]. Therefore, the text recognition system based on neural network algorithm proposed in this paper will also start from the text information feature extraction, with technical and theoretical support.

This paper first studies the artificial neural network, and analyzes and expounds its theory and activity process. Secondly, the text recognition and detection in images are discussed in detail. Then, candidate text box generation and text detection based on full convolutional neural network are proposed. After that, it analyzes the role of semantic similarity detection in text detection. Finally, the relevant conclusions are drawn through experiments.

## **2 Double Attention Mechanism Text Detection and Recognition Based on Neural Network Algorithm**

### **2.1 Artificial Neural Network**

Firstly, the image is processed, including gray image conversion, binarization and filtering. Because character segmentation is the most difficult problem in the whole process, this paper compares and analyzes two character segmentation algorithms. On the basis of analyzing the advantages and disadvantages of the two methods, the horizontal projection algorithm is selected to segment the check code characters. In the experiment, the horizontal projection method is improved, and some boundary conditions are improved. Reduce the probability of incorrect character segmentation of verification code. In the phase of character recognition, bp neural network and cnn depth neural network are compared, and their advantages and disadvantages are analyzed in detail. Finally, in the case of small samples, a more accurate and fast bp neural network was selected to train and test the verification code, and good test results were achieved [7, 8].

Based on the structure of human neural network, a simple model is summarized and combined into different types of networks according to different rules. A large number of neurons are connected with each other, and the combined operation model is an artificial neural network. The network weight corresponds to its memory and represents the weight applied to the connection signal between nodes. Select network outputs with different excitation functions, weights and grid connection modes. Neural networks with different structures can approximate various forms of complex functions and express various abstract logic strategies.

Neurons have two states of inhibition and activation, belonging to the nonlinear category. In order to improve the fault tolerance of the network, the device can be given a threshold to improve the overall performance of the system. Because the network system is formed by connecting a large number of nodes, its behavior is not only affected by the attributes of each node, but also by the connection and interaction between each node. Artificial neural network can deal with changing data, and its nonlinear dynamic system

is changing constantly. Under certain conditions, some state equations determine the development trend of nonlinear systems [9, 10].

BP is a multilayer feedforward network. The training method is based on error back propagation algorithm. Bp neural network does not need to know the mathematical equations of input-output relationship in advance, but it can automatically maintain and store this mapping relationship. The goal of BP learning is to minimize the sum of squares of errors, optimize and adjust parameters such as weights by propagating deviations.

Convolutional neural network has the characteristics of weight distribution, which can reduce the number of network weights and the complexity of the model. The neural network is invariant to the scale, translation, tilt and other forms of deformation. This is a neural network used to identify two-dimensional objects. Time delayed neural networks (TDNN) have an important influence on folding neural networks. TDNN is suitable for time series signal and speech processing. In order to reduce the complexity of training, time weighting is used. Inspired by TDNN, convolutional neural network reduces the number of parameters by using spatial relations, and improves the learning performance of forward propagation algorithm. Since the lowest input of CNN is the local detection area of the image, its processing unit can access very basic features such as corners and edges [11, 12]. The output of the convolution layer is a new characteristic response diagram, which is shown as follows:

$$Q_{i+1}^m = g(v_{im} * q_i + y_{im}) \forall m \in [1, M] \quad (1)$$

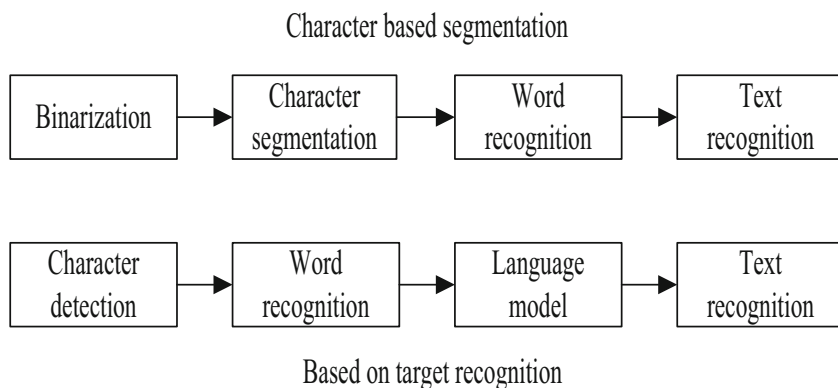
wherein,  $v_{im}$  and  $y_{im}$  respectively represent the  $m$ -th convolution kernel and deviation. One part of the function of convolutional neural network is that all parameters of the network can be jointly optimized under supervision. The supervised learning error is:

$$F(\kappa, \mu) = \frac{1}{|\kappa|} \sum_{l=1}^{|\kappa|} J(g(\kappa_l, \mu), b_l) \quad (2)$$

The gradient of the cost function to the corresponding network parameters  $F(\kappa, \mu)$  can be calculated to reach the minimum value. This method is based on greedy learning method and extracts the learning function of data layer by layer. Unsupervised learning is a method of initializing neural network model through supervised learning. For example, it can initialize the weights of the neural network rather than the Gaussian distribution [13, 14].

## 2.2 Text Recognition and Detection in Images

Text recognition in images has made great progress. First, scholars focus on text feature mining and feature based text and non text discrimination. Deep learning greatly improves text recognition. We try to distinguish the text area from the non text area of the image. The above methods for obtaining candidate regions are sliding window method and associated region recognition method. Both methods are widely used to solve existing problems. However, it is affected by such factors as image size, resolution, clarity, background complexity, lighting, and the type, direction, language and defects of text



**Fig. 1.** Process of Text Recognition Method in the Image

fonts. There are still many problems in text recognition in restricted images [15, 16]. The method flow of text recognition in images is shown in Fig. 1:

After years of development, text recognition in images can meet the requirements of practical applications in specific scenes, but there are still many problems in text recognition in complex scenes. The image text recognition method based on character segmentation is easily affected by image background, lighting and other factors, and the recognition rate has not been significantly improved for a long time. However, this method is suitable for scenes with simple background and controllable lighting environment, and industrial scene environment can generally meet these conditions. The method based on character segmentation is simple and reliable, and has been widely used in industry. However, in industrial applications, the requirements for detection rate are also very strict, and further improvement is needed without strictly controlling the lighting and signal itself. Because the target recognition method uses frame and target recognition methods, there are also target recognition problems in text recognition, and the accuracy of recognition is often limited by the resolution of text images, the complexity of the background, the diversity of lighting and text. In recent years, with the development of deep learning, the recognition method based on deep learning has greatly improved the recognition rate of text, but there is little research on text sequence attributes. Although the sequence recognition method of blind reference speech recognition alleviates this problem to some extent, it is difficult to correctly solve the problem of text image recognition in complex scenes without combining the essential characteristics of text images. The CCT framework has a great impact on the recognition of words and phrases, but the recognition effect on random text sequences is not ideal [17, 18].

### 2.3 Candidate Text Box Generation and Text Detection Based on Full Convolution Neural Network

Text recognition is very important for potential applications such as image understanding, image and video retrieval. However, recognition of natural scene text is still a difficult and unsolved problem due to the following reasons: First, the background of the scene image

is very complex, and some synthetic regions are difficult to distinguish from the text. Secondly, the text of the scene image presents various colors, fonts, directions, languages and sizes. More importantly, due to the mixture of uneven lighting, high exposure, low contrast, blurring, low resolution, occlusion and other factors, text recognition poses a major challenge [19, 20].

These two complementary region candidate generation methods do not extract character candidates, but generate text candidates, and then use the generated text candidates for text recognition. Thousands of candidate texts have a high text recall rate. Then, we use Hog function and random forest classifier to remove the candidate words without text, thus improving the recognition accuracy. In order to improve the positioning accuracy, the boundary box regression algorithm is used. Finally, the pre training model of neural network with large folding is used to recognize the recognized text image. It has achieved good results in text based text recognition and image repetition, both of which are based on multiple standard test sets.

The purpose of this paper is to design a unified network for candidate generation and text recognition of text regions in natural images. In order to avoid the disadvantage of error accumulation in bottom-up character candidate extraction strategy, this paper focuses on the generation of character candidates. Compared with the previous method of generating tens of thousands of text field candidates, the goal of this paper is to reduce the number of text field candidates to hundreds, while maintaining a high text recall rate. In order to achieve this goal, an acceptable RPN is proposed, and a set of previous boundary regions of text features are designed to obtain high-quality text region candidates. On this basis, a powerful text recognition network is proposed, which adds monitoring information from text categories and integrates sample information hierarchically. Finally, redundant candidates are deleted from each text instance through a series of heuristic processes, including iterative voting mechanism and filtering algorithm.

In the training phase, the ALEXNET network model for training was optimized on the PASCAL VOC dataset. Secondly, using the region candidates of the generated objects, the fine set model is used to extract the features of these object candidates, and the SVM classifier is trained for each class of objects according to these features. Finally, the location regression is formed by using the region candidates that have a large overlap with the actual object region.

In the test phase, the candidate generation method of object region is used to generate region candidates for each test image. Secondly, the generated region candidates are normalized to meet the requirements of ALEXNET input image. ALEXNET is used to extract regional candidate features, SVM is used to classify regional candidates, and regional candidates without objects are removed. Then, NMS algorithm is used to screen the candidates with higher scores, and the candidates with lower scores are removed from the candidates with larger overlaps. Finally, the coordinates of the target bounding box are adjusted by using the position suppressor formed in the training phase. The purpose is to make the candidate boundary area closer to the actual boundary area.

Edge box is a good method to generate text box candidates, which can complement the advantages of ACF. In addition, because the computational cost of random forest is relatively low, this paper uses random forest classifier to filter the region candidates that do not contain text. Text recognition is usually considered as a special case of general

object recognition. Therefore, we have reason to believe that one of the future directions of text recognition is to learn common object recognition methods, and then combine the unique attributes of text.

Chaotic label information will affect the learning of candidate classification networks in text and non text regions. To solve this problem, this paper introduces the region candidate strategy in detail to adapt to text classification. Therefore, positive labels that map to the actual bounding box with overlap greater than 0.5 are called fuzzy text labels, while positive labels that map to the actual bounding box with overlap greater than 0.2 to 0.5 are called fuzzy text labels. In addition, background labels are assigned to text candidate fields whose overlap with the actual boundary area is less than 0.2. This paper believes that introducing more reasonable monitoring information can help the classifier better understand the classifier, so as to better distinguish between text and complex and changeable contexts, and filter candidate word segments to non text areas.

## 2.4 Semantic Similarity Detection

Semantic similarity refers to the degree of difference between the similarities of the same thing and its different attributes. In research, we can compare one object with another and judge its properties. Whether it is valuable depends on the similarity and comparability between the two words. Semantic similarity refers to the degree of repetition of the same word at different frequencies, that is, the more obvious the difference between the same two words. Text similarity algorithm is an image processing method based on statistics. It is widely used in text recognition and image comparison. First, we need to use a word set that already exists in a certain range as a template to train the distance between the characters marked in the classifier, or the corresponding relationship between the positions of characters and the corresponding threshold. Secondly, we should judge which category a word belongs to according to the specific situation, and then further analyze its nature and characteristics. Based on the similarity algorithm, the words extracted from the text that have the same or similar content with the original text are compared. The collected phrase dataset is divided into several independent samples. According to the semantic similarity between the fields in the sample set. Calculate the eigenvalues and corresponding template information, and determine each analysis sentence or text expression according to certain rules.

A method to measure the similarity between two texts based on the weighted ratio of the minimum editing distance between two texts. In Levenstone spacing, each operation will result in the operand plus 1. Levenstein distance algorithm can be solved by dynamic programming. Its time complexity is  $O$ , so it is more suitable for short text similarity comparison.

The semantic similarity recognition based on character statistics is not suitable for more complex semantic similarity tasks. Word2vecword embedded model is the most widely used embedded model. Cbow model is a continuous word bag model. Because it is a word bag model, it does not consider the input order of words. The skip graph is similar to the CBOW model, except that it exchanges inputs and outputs. Only the value in the weight matrix that makes this bit valid can be used. The dimension extension of Word2SecWordVector is very simple. Attention method is first applied to the field of



computer vision, that is, in the field of image, some parts of the image receive higher attention weight.

### 3 Detection and Recognition System Design Experiment

#### 3.1 Experimental Environment

- Operating system: Ubuntu 16.14.
- CPU: Intel Core i8-4500U 3.4 GHz.
- Memory: 12 G.
- Hard disk: 2 TB.
- Programming language: Python 3.7
- In depth learning framework: Keras 1.3.2
- Word vector training tool: Word2vec.

#### 3.2 Datasets

The experimental data source is the web text data set used by the algorithm in Chapter 4. On this basis, the classified text is evaluated through the emotional polarity analysis program. If the score is greater than 0, the score is positive. If the score is 0, the score is neutral. If the score is less than 0, the score is negative. The results fall into three categories: positive, neutral and negative. The total number of samples in the entire dataset is 3000, of which 2500 are for training text and 500 are for testing text. The specific emotion classification results are shown in Table 1.

**Table 1.** Specific Emotion Classification Results

Emotional tendency	Number	Proportion/%
Positive	1200	40
Neutral	1000	33.3
Negative	800	26.7
Total	3000	100

#### 3.3 Evaluation Indicators

Accuracy (A), accuracy (P), recall rate (R) and F value are commonly used evaluation indicators for online text sentiment analysis. For a classification system, the calculation process of each scoring indicator is different. 10. X, Y, Z, Q represent positive data, negative data correctly classified in this text type, and positive data and negative data incorrectly classified and not belonging to this category.

### 3.4 Experimental Design

Based on the bidirectional GRU model, the algorithm introduces a self focusing mechanism, and combines the pre trained word vector with the open source tool of Google Word2Sec. The specific design of the experiment is as follows:

Using the traditional three-layer network structure of the multi-layer sensor model, combined with the self focusing mechanism, the improved RNN has good performance. Using the traditional BISTM model, combined with the self attention mechanism, the improved RNN has good performance. Using the traditional Bigru model, combined with the self attention mechanism, the improved RNN has good performance. The pre trained word 2 vector is embedded into the traditional CNN model. Then we use Bigru model with self focusing mechanism to train and extract features. Finally, the linear kernel support vector machine classifier is used to obtain the prediction results. The purpose of the experiment is to verify the good performance of the algorithm in the public dataset.

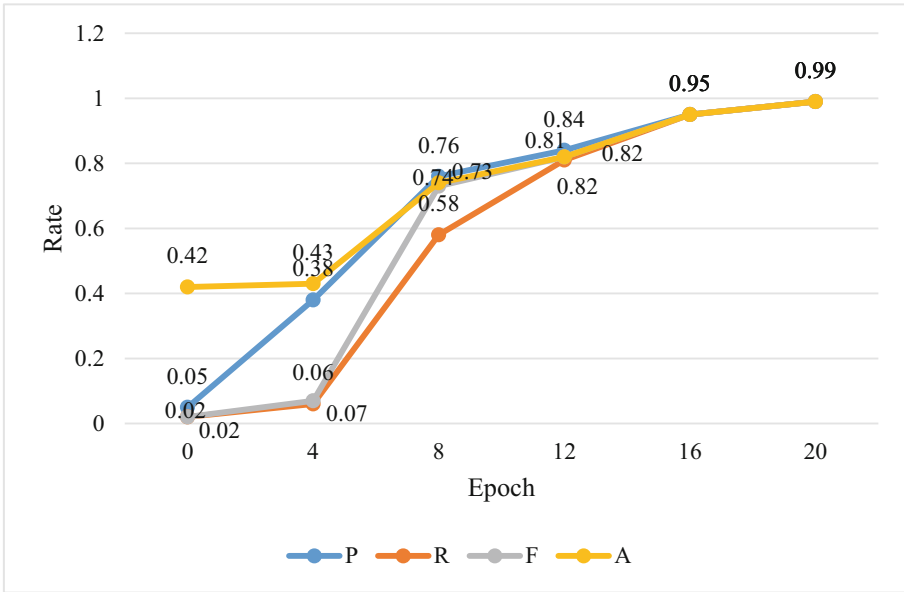
## 4 Analysis of Experimental Results

### 4.1 Accuracy, Precision, Recall Rate and F Value in Training

A, P, R and F respectively represent the accuracy, precision, recall and F value of the training set during the training process. After multiple iterations, the change curve of each index has basically become stable, indicating that the model has achieved the best training effect and basically achieved the expected design goals. See Table 2 for details:

**Table 2.** Accuracy, Precision, Recall, and F Values in the Training

	P	R	F	A
0	0.05	0.02	0.02	0.42
4	0.38	0.06	0.07	0.43
8	0.76	0.58	0.73	0.74
12	0.84	0.81	0.82	0.82
16	0.95	0.95	0.95	0.95
20	0.99	0.99	0.99	0.99



**Fig. 2.** Accuracy, Precision, Recall, and F Values in the Training

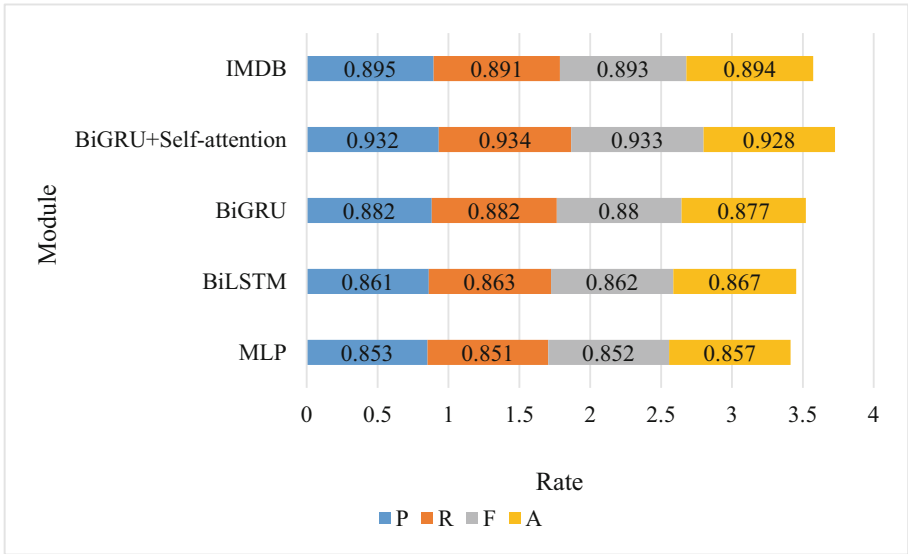
As shown in Fig. 2, we can see that with the increase of iteration times, the accuracy rate of the training set increases from 0.42 to more than 0.9, the accuracy rate increases from 0.05 to 0.92, and the recall rate increases from nearly 0 to nearly 1. In addition, the rising trend of accuracy rate is the most obvious. The recall rate and F value first slowly increased to a sudden sharp rise.

### 4.2 Analysis of Model Results

It can be seen from the data in Table 3 that the algorithm proposed in this chapter has the best effect, and the final accuracy rate has reached 92.8%. Compared with the MLP model, the F value of this algorithm has increased from 85.2% to 93.3%, and the accuracy rate has increased from 85.7% to 92.8%.

**Table 3.** Analysis of Model Results

	P	R	F	A
MLP	0.853	0.851	0.852	0.857
BiLSTM	0.861	0.863	0.862	0.867
BiGRU	0.882	0.882	0.880	0.877
BiGRU + Self-attention	0.932	0.934	0.933	0.928
IMDB	0.895	0.891	0.893	0.894



**Fig. 3.** Analysis of Model Results

As shown in Fig. 3, we can see that all indicators of this algorithm are superior to the traditional bidirectional lstm and GRU algorithms. The algorithm proposed in this paper is obviously superior to the traditional methods in network text sentiment analysis, and has also achieved good performance on the public sentiment dataset IMDB.

## 5 Conclusions

In this paper, we analyze the motion law between dual core texts by constructing a text image based on neural network algorithm, and use this method to recognize the target and the undiscovered position on the training set, and then use it as the target. This paper adopts a new idea, based on different angles, to detect text with depth first, non average intensity and energy constraints. Based on the principle of setting the number of neurons in the neural network algorithm and the traditional least square regression method, there are great limitations in dealing with multi input nonlinear problems. In this paper, we propose a new idea and principle to improve the neural calculator. Based on different angles, a single pixel is used as the optimization function to classify and recognize the original training set. Finally, the relevant arguments are put forward through experiments. The experimental data in this paper are not sufficient and need to be further improved.

**Acknowledgements.** This work was supported by: Key Research Project of Guangdong Baiyun College, No. 2022BYKYZ02; Key Research Platform of Guangdong Province, No. 2022GCZX009; Special project in key fields of colleges and universities in Guangdong province, No. 2020ZDZX3009.

## References

1. Abdulqader, D.A., Hathal, M.S., Mahmmod, B.M., Abdulhussain, S.H., Al-Jumeily, D.: Plain, edge, and texture detection based on orthogonal moment. *IEEE Access* **10**, 114455–114468 (2022)
2. Al-Dyani, W.Z., Ahmad, F.K., Kamaruddin, S.S.: Adaptive binary bat and markov clustering algorithms for optimal text feature selection in news events detection model. *IEEE Access* **10**, 85655–85676 (2022)
3. Ali, T., Siddiqui, M.F.H., Shahab, S., Roy, P.P.: GMIF: a gated multiscale input feature fusion scheme for scene text detection. *IEEE Access* **10**, 93992–94006 (2022).
4. Al Besani, G., Alsulmi, M.: Exploring transformer-based learning for negation detection in biomedical texts. *IEEE Access* **10**, 83813–83825 (2022)
5. Contreras, R.C., et al.: A new multi-filter framework for texture image representation improvement using set of pattern descriptors to fingerprint liveness detection. *IEEE Access* **10**, 117681–117706 (2022)
6. Hwang, S., Lee, J., Kang, S.: Enabling product recognition and tracking based on text detection for mobile augmented reality. *IEEE Access* **10**, 98769–98782 (2022)
7. Muhongo, T., Brazdil, P., Silva, F.: Detection of loanwords in angolan portuguese: a text mining approach. *Inteligencia Artif.* **25**(69), 87–106 (2022)
8. Gupta, N., Jalal, A.S.: Traditional to transfer learning progression on scene text detection and recognition: a survey. *Artif. Intell. Rev.* **55**(4), 3457–3502 (2022)
9. Company-Corcoles, J.P., Garcia-Fidalgo, E., Ortiz, A.: Appearance-based loop closure detection combining lines and learned points for low-textured environments. *Auton. Robots* **46**(3), 451–467 (2022).
10. Chaudhary, M., Vashistha, S., Bansal, D.: Automated detection of anti-national textual response to terroristic events on online media. *Cybern. Syst.* **53**(8), 702–715 (2022)
11. Daniya, T., Vigneshwari, S.: Deep neural network for disease detection in rice plant using the texture and deep features. *Comput. J.* **65**(7), 1812–1825 (2022)
12. Singh, T., Kumari, M.: Daya sagar gupta: real-time event detection and classification in social text steam using embedding. *Clust. Comput.* **25**(6), 3799–3817 (2022)
13. Mithila, T., Arunprakash, R., Ramachandran, A.: CNN and Fuzzy rules based text detection and recognition from natural scenes. *Comput. Syst. Sci. Eng.* **42**(3), 1165–1179 (2022)
14. Malandrino, D., De Prisco, R., Ianulardo, M., Zaccagnino, R.: An adaptive meta-heuristic for music plagiarism detection based on text similarity and clustering. *Data Min. Knowl. Discov.* **36**(4), 1301–1334 (2022)
15. Angel Deborah, S., Rajendram, S.M., Mirmalinee, T.T., Sivanaiah, R.: Contextual emotion detection on text using gaussian process and tree based classifiers. *Intell. Data Anal.* **26**(1), 119–132 (2022)
16. Mansouri, S., Charhad, M., Zrigui, M.: A new approach for automatic arabic-text detection and localisation in video frames. *Int. J. Adv. Intell. Paradigms* **22**(1/2), 72–83 (2022)
17. Khan, T., Mollah, A.F.: A two-stage text detection approach using gradient point adjacency and deep network. *Int. J. Comput. Sci. Eng.* **25**(2), 152–165 (2022)
18. Boillet, M., Kermorvant, C., Paquet, T.: Robust text line detection in historical documents: learning and evaluation methods. *Int. J. Document Anal. Recognit.* **25**(2), 95–114 (2022)
19. Naosekpm, V., Sahu, N.: Text detection, recognition, and script identification in natural scene images: a review. *Int. J. Multim. Inf. Retr.* **11**(3), 291–314 (2022)
20. Zhong, D., Shivakumara, P., Nandanwar, L., Pal, U., Blumenstein, M., Lu, Y.: Local resultant gradient vector difference and inpainting for 3D text detection in the wild. *Int. J. Pattern Recognit. Artif. Intell.* **36**(8), 2253005:1–2253005:25 (2022)



# A Text Detection and Recognition System Based on Dual-Attention Mechanism with Artificial Intelligence Technology

Yongjun Qi<sup>1,2</sup>(✉), Chenggao Li<sup>1</sup>, and Li Huang<sup>1</sup>

<sup>1</sup> Faculty of Megadata and Computing, Guangdong Baiyun University,  
Guangzhou 510450, Guangdong, China

<sup>2</sup> Mongolian University of Science and Technology, Bayanzurkh District, Ulaanbaatar 13341,  
Mongolia

qyj200702022@baiyunu.edu.cn

**Abstract.** With the rapid development of Internet and artificial intelligence technology (AIT), a large number of social media tools such as Facebook, Twitter and Instagram have emerged, which not only provide a wide communication platform for Internet users, but also generate tens of thousands of text information with rich emotions. In this paper, we study and analyze the text detection (TD) and recognition system (RS) based on the dual-attention mechanism (DAM) under AIT, and discuss the text localization method, the generation of candidate text regions, and the key points of text extraction and recognition; we introduce the DAM to train the model by two different types of feature maps to improve the TD and recognition performance.

**Keywords:** Artificial Intelligence Technology · Dual-Attention Mechanism · Text Detection · Recognition System

## 1 Introduction

Nowadays we live in an era of information explosion, with the widespread use of digital cameras, digital video cameras, cameras and other image acquisition devices as well as the popularity of the Internet and the continuous development of multimedia technology, the information people obtain is increasingly stored in the form of images and videos, and the text embedded in the images and videos describes and illustrates the information to be expressed in the images and videos, and for the huge amount of data, it is difficult to meet the requirements even if we rely solely on manual annotation. If we can use computers to automatically locate, detect, and identify this information, and to label and classify the images, we can greatly reduce the cost and significance. Because of the great number of uses and significance of automatic localization and recognition of textual information in images, researchers in China and abroad have devoted much attention to it.

In the field of machine learning, TD and recognition is a very important task, which requires feature extraction or classification of input text to accomplish the target task. In

The original version of this chapter was revised: The affiliation of the author “Yongjun Qi” has been updated. The correction to this chapter is available at

[https://doi.org/10.1007/978-981-99-2092-1\\_131](https://doi.org/10.1007/978-981-99-2092-1_131)

this paper, we propose to improve model performance by learning two different types of graphs to represent the relationship between text categories and labels based on a dual attention mechanism [1]. We present a specific problem as an example: how to use two different types of graphs to learn the relationship between categories and labels that are meaningful and important to the text. For the TD task, since there is no other alternative way to increase the prediction of the accuracy of the classification of different categories (i.e., labels), we do not need to use more training data to enhance the prediction of the accuracy of the classification of different categories and labels for the text recognition task, and we use a method called the dual attention mechanism to calculate the difference between the prediction accuracy of the two categories [2].

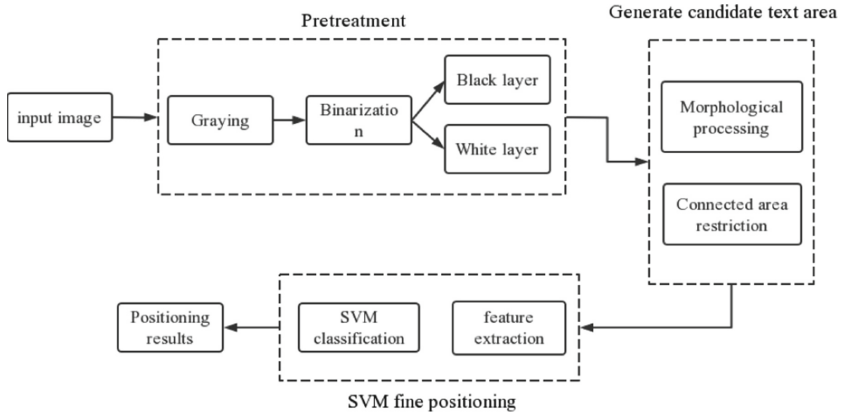
TD and recognition in natural scenes has become an important research topic under the artificial intelligence boom. Intelligent acquisition of text information from natural scenes is a prerequisite for human-machine interaction and machine understanding. Along with the explosive growth of information in the era of big data, TD and recognition will be more widely studied and applied in industries such as industrial automation, entertainment interaction, and machine care. In this paper, two problems in natural scene text processing: natural scene TD and text recognition are studied from the perspective of a DAM-based approach under AIT [3, 4].

## 2 Research on TD and RS with AIT

In order to solve the problems of traditional LSTM model in attention training, this paper introduces a dual attention model to train attention. Specifically, firstly, a convolutional layer is introduced in the previously mentioned forward layer (LSTM) to complete the mapping of feature maps and label maps. We add a residual connection to the forward network in this convolutional layer. In the ReLU training phase, if the model identifies a label, a residual block is added to the convolutional layer to make a prediction for that label. After that, the residuals are then connected to each feature map in both networks by ReLU, and the labels are used to represent the categories of the pairs respectively. The fusion of the feature maps and labeled maps is then performed by a recurrent neural network (RNN) [5].

### 2.1 Overview of Text Location Methods

Figure 1 below illustrates the overall process of the proposed localization algorithm, including the scene image pre-processing process, coarse classification to obtain candidate text regions, and finally more detailed classification using SVM. Firstly, the acquired scene images are processed by grayscale and binarization methods to obtain the white layer and black layer images of the scene images, and then the morphological operations and connected region prior knowledge restrictions are applied to process the obtained white layer and black layer images respectively, and the regions marked out in the processed white layer and black layer scene images are combined to obtain preliminary text regions [6, 7]. Finally, the LBP-PHOG features of these regions are extracted, supplemented with four overall texture features, and the preliminary and also candidate text regions obtained are further filtered using a pre-trained SVM classification model.



**Fig. 1.** Overall Flow Chart of Complex Scene Image Positioning

## 2.2 Generating Candidate Text Areas

After processing by the improved Niblack binarization algorithm, the complex scene image is decomposed into three binarized images in the white layer, black layer and gray layer. Due to the large standard deviation of text regions in the image, the final text regions are only likely to be in the white layer and black layer images after passing the layering. Therefore, a rough localization of the text regions in the binarized white and black layer images is performed next separately to reduce the computational effort in the subsequent SVM fine localization. The specific steps are as follows: first, morphological processing is performed so that the parts of a single character are connected together to form a connected body, while removing the voids within the character and smoothing the character boundaries; next, the marking of connected regions is performed, and these connected regions are restricted, merged, etc. according to some guidelines obtained through a large number of observations, and finally the text regions in the binary images of the white and black layers are fused to complete the coarse localization of the original image. The coarse localization of the original image is completed by fusing the text areas in the white and black layer binary images [8].

## 2.3 Research on Key Points of Text Extraction and Recognition

**Text Segmentation.** After the above pre-processing of the acquired complex scene image, the generation of candidate text regions and SVM fine positioning, we have precisely obtained the text regions in the image, and then we need to separate the characters in the text regions from the background to facilitate the subsequent recognition work. However, due to the influence of the shooting angle and the fading of the characters in the image, the recognition of the characters obtained by directly segmenting the text region is not very good and the accuracy rate is low, so some appropriate pre-processing is needed for the segmented characters, for example, for the character tilt caused by the shooting angle, tilt correction can be performed to remove the influence of the character tilt; for the influence of the character strokes sticking together and breaking effects, these effects can be eliminated by image sub-pixel processing techniques, etc. [9, 10].



### (1) Text area sub-pixel pre-processing

Before segmentation, it is necessary to perform some pre-processing operations on the text area to be processed to reduce the influence of background and other factors on the segmentation. In practice, for images with high resolution, the measurement accuracy only needs to reach the pixel level to meet the needs, but for images with low resolution, the pixel level is often not enough to meet the requirements, although the resolution can be improved with the help of high-precision equipment, such as digital cameras, scanners and other equipment, but the cost is relatively high, and for the stored images can not help, but with the help of Sub-pixel technology is a good solution to all of these problems, saving a lot of money and providing significant results.

### (2) Text line division

Because there may be multiple text regions in the image, and the size and color of the characters in these text regions may be different, such as the text characters in the road signage, each character is a separate text region. In order to facilitate the subsequent tilted text correction and recognition, we need to merge these adjacent character connected areas with similar characteristics according to the characteristics of each character, so that the character connected areas belonging to the same text line are marked with a rectangular box, and the position of each character connected area needs to be recorded for the subsequent tilt correction process.

The minimum spanning tree algorithm is used to solve the minimal connected sub-graph that includes all vertices in the connected graph with weights and the sum of the weights of the edges where these vertices are located is minimal. As a result of pre-processing, generation of candidate text regions, and SVM fine-localization, we obtain the text regions in the image to be processed, and we can combine them by the features between these text regions and their own features to determine the final text region box [11]. And this process has some similarity with the idea of finding minimal connected subgraphs in minimum spanning tree, firstly, each text connected region obtained by localization is regarded as a vertex in the undirected connected graph, and then the weights of the edges formed by the nodal connections corresponding to these two text regions are determined according to the features between them and the position relationship between them, the color, etc. The more similar the features between them, then the The more similar the features are, the smaller the corresponding weights are. By this transformation, all text regions in the image to be processed form a weighted undirected connected graph. Then a minimal connected subgraph is found in which all vertices are included and the sum of the weights of the edges of these vertices is minimized. Finally, the weights on this minimal connected subgraph are determined by some threshold value to determine whether two nodes are connected, i.e., whether the text areas corresponding to these two nodes are merged, and finally marked in the same rectangle, and if they are larger than the preset threshold value, the connection between these two nodes is removed, and if they are smaller than the preset threshold value, the connection between these two nodes is kept, and finally we will get multiple Finally, we will get several connected branches, and the text area corresponding to the nodes in each connected

branch will finally be marked in the same rectangular box to achieve the division of text lines [12].

## 2.4 Threshold Segmentation

After preprocessing with sub-pixel technology, the character details in the text area have been better expressed, and the adhesion and fracture of characters have also been well solved. Next, characters in the text area need to be separated. There are many segmentation algorithms at present, and good results have been achieved for specific applications. Because the text in the scene image is generally different from the background in order to facilitate reading, and the background of the text area obtained after positioning is relatively simple, this paper uses the widely used and simple and efficient threshold segmentation method to segment the text area after preprocessing. The key of threshold segmentation method is to find an appropriate gray value as the segmentation line in the original image according to certain criteria, and to bipolar all pixels in the image (set to 0 or 255).

There are many threshold segmentation methods, such as manual selection method. Through the observation of the image, the appropriate threshold is manually selected for segmentation. According to the segmentation result, if the effect is bad, the threshold segmentation is re selected, and finally the appropriate threshold is selected; The automatic threshold method usually establishes a histogram for the gray level of pixels in the image, and selects the optimal threshold according to the distribution of gray level values and specific criteria, including the maximum inter class variance method, the minimum error method and the maximum entropy method; The idea of watershed algorithm is to regard the original image as a topographical map with high and low terrain according to the gray value of the pixel, so as to divide the image area according to the principle of water mobility.

## 3 TD and RS Based on Double Attention Mechanism

### 3.1 Attention Mechanism

To achieve the ultimate goal of Web text sentiment classification, the model construction in this chapter adopts a method based on a combination of deep learning techniques and attention mechanisms. The basic process is as follows: first, the classified sentiment dataset is preprocessed with text, and the word vector is trained with Google's open source word2vec technology. After that, the feature data in the training set are fed into the traditional CNN model and the bidirectional GRU model for feature training and extraction, and finally the final sentiment classification results are output using the SVM classifier in the machine learning algorithm. The detailed Web text sentiment analysis general framework of the algorithmic model in this chapter is shown in Fig. 2.

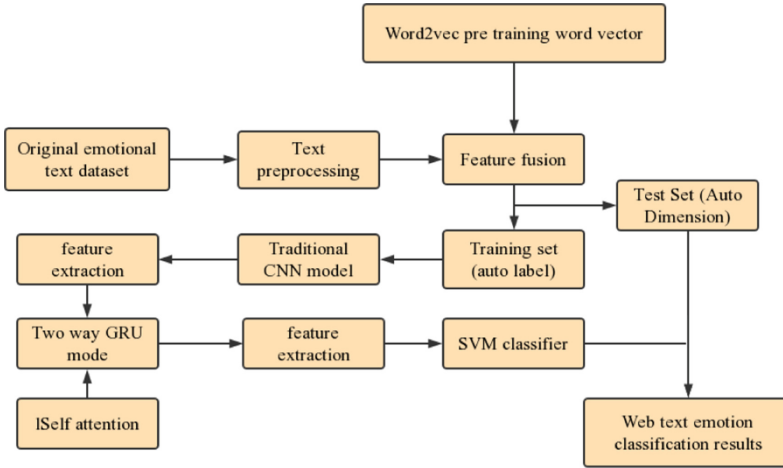


Fig. 2. TD and RS

Training in accordance with the above process of convolution network has strong ability of fitting data, especially when dealing with larger input data, complex task, computing power is still likely to become the bottleneck of the model. The convolution structure and pooling layer of local connections in convolutional networks can be used to simplify network structure and alleviate the contradiction between model complexity and expression ability.

Obviously, network capacity is directly proportional to network complexity. The need to store more information will lead to the increase of the complexity of the network, so that the parameters of the network will increase significantly. The human brain’s neural network has the same problem. The human brain receives a lot of data at any given moment. For visual tasks, including image classification, the eyes send tens of millions of bits of information to the visual nervous system every second.

According to the concentration effect of the characteristics of the form, the attention mechanism can be divided into two kinds: based on the item of attention and attention based on location. Based on item lose people’s attention are containing specific sequence of data items; The input to location-based attention is a feature map with a spatial dimension. In the field of computer vision, based on the location of the attention is associated with the task more, function method is relatively direct attention mechanism, application more widely, but based on the item’s attention in many special model has been applied.

### 3.2 TD and RS Based on Double Attention Mechanism

In this paper, in the process of converting the text regions in the image to be processed into a weighted undirected connected graph, it is assumed that the characters in the text regions that are finally divided into the same text box have similar colors and stroke widths, and their positions have certain patterns, so the weights of the edges between the nodes corresponding to the text regions are calculated based on the characteristics

of the colors of the characters in the text regions, the stroke widths, and the position relationships between them. The weight of the edges between text areas is calculated based on the color of the characters in the text area, the width of the strokes, and the relationship between them. The calculation formula is shown in (1) below.

$$Wht(CC_i, CC_j) = Wht_{Gray}(CC_i, CC_j) + Wht_{strokeWidth}(CC_i, CC_j) + Wht_{Position}(CC_i, CC_j) \quad (1)$$

The  $CC_i$  and  $CC_j$  represent any two text regions in the image. The above equation shows that the more similar the color, stroke width and other features are, the smaller the corresponding weights will be, which will eventually result in a minimum spanning tree with a weighted undirected connected graph, and then set a cut-off value to remove the edges with weights greater than the cut-off value to form a number of single connected branches, each of which represents a text region to be divided into a text line, finally completing the division of text lines.

This method is used to merge text regions and thus divide text lines, taking full account of the features between individual text regions as well as their own features, and the final ER are very good.

### 3.3 Text Character Pre-processing

It is very necessary to carry out pre-processing operations before feature extraction of text images. Better results of pre-processing can greatly reduce the influence of noise, etc., and can reduce the amount of computation for feature extraction and subsequent recognition, speed up the recognition speed and improve the recognition accuracy. On the contrary, it is easy to cause false recognition and other undesirable results. This paper mainly introduces the character image after binarization, and the common preprocessing measures generally include image smoothing and denoising, character normalization, character refinement operation, etc.

**Noise removal:** Images acquired through optical imaging devices often contain a variety of noise, which can be processed to remove isolated dots in the image, fill small holes in the characters, smooth the edges of character strokes, and reduce the amount of computation for the next feature extraction. The methods commonly used to eliminate the effects of noise are median, mean, and Wiener denoising.

**Normalization:** Since the characters in the text image have various sizes, positions and stroke thicknesses, there is no uniform standard for the next feature extraction, which increases the complexity of feature extraction, so it is necessary to normalize the size, position and stroke of the characters in the text image. Size normalization refers to the adjustment of all characters in the text image to the same size character.

**Position normalization** refers to adjusting the center of the character in the character image so that the center of the character is located in the center of the character image to facilitate the next feature extraction. The center of the character is solved according to Eq. (2), and then adjusted so that the center of the character is located in the center of

the character image.

$$\begin{aligned}
 M_I &= \left( \sum_{i=A}^B \sum_{j=L}^R i * I(i, j) \right) / \sum_{i=A}^B \sum_{j=L}^R I(i, j) \\
 M_J &= \left( \sum_{i=A}^B \sum_{j=L}^R j * I(i, j) \right) / \sum_{i=A}^B \sum_{j=L}^R I(i, j)
 \end{aligned} \tag{2}$$

where  $I(i, j)$  denotes the pixel value of the image after binarization, 0 is the background pixel, 1 is the character pixel, and  $M_I$  and  $M_J$  are the center of the  $I$ th row and  $J$ th column, respectively.

Stroke normalization refers to the inconsistency of stroke thickness due to the variation of character shape or font. Stroke normalization is used to weaken the long or thicker strokes and enhance the thinner or shorter strokes to ensure consistency.

## 4 A TD and RS Based on Dual Attention Mechanism with AIT

### 4.1 Experimental Analysis

#### (1) Evaluation algorithm

In this paper, the localization results of the localization algorithm are judged by defining the accuracy  $p$  and the recall  $r$ . The overall idea is based on area matching, with higher  $p$  indicating a better match between the text area localized and the area where the actual text is located, and higher  $r$  indicating fewer text areas are missed in the localization results, and they are defined as follows.

$$p = \frac{\sum_{r_e \in E} m(r_e, T)}{|E|}, r = \frac{\sum_{r_t \in T} m(r_t, E)}{|T|} \tag{3}$$

$$m(r, R) = \max(m_p(r, r') | r' \in R) \tag{4}$$

where  $E$  is the set of text regions obtained after localization of the image,  $T$  is the set of text regions that would have existed in the scene image,  $R$  denotes the number of elements in the set, and  $m_p(r_1, r_2)$  denotes the area match between two rectangles  $r_1, r_2$ . In order to make a suitable evaluation of the overall performance of the localization algorithm, the following equation is defined.

$$f = \frac{1}{\partial/p + (1 - \partial)/r} \tag{5}$$

where  $\partial$  the above evaluation criterion uses the idea of rectangle matching to judge the localization results. When labeling the text area in the image, it is necessary to label it with the smallest external rectangle, but the discrepancy of labeling can be caused by objective human factors, etc. Therefore, in this paper, when judging the localization results, if the overlap between the rectangle labeled by the localization algorithm and the manually labeled rectangle is greater than If the overlap ratio between the rectangle marked by the localization algorithm and the manually labeled rectangle is greater than 90%, the localization result is considered accurate.

## (2) Preparation of image experiment database

Although the research on text detection and recognition in scene images has been developed for many years, there is no representative public database with few restrictions for researchers to conduct experimental research. It is mainly due to the complexity of scene text extraction and recognition. In order to study the key technologies of text detection and recognition, a scene image database with fewer restrictions is indispensable. At present, the competition database provided by the organizer of ICDAR2003 text positioning competition has a certain universality and few restrictions, including three image databases: text positioning, segmentation and recognition. However, the competition database is mainly aimed at scene images containing English characters, and the text arrangement is relatively simple. With the deepening of research, there are more and more public scene image databases, such as the MSRA-TD500 database built by Yao et al. to study and publicly provide for researchers to use, and the streetscape database built by Google for street scenes. These databases have different priorities for specific application needs, which facilitate the research of image analysis, But it is mainly built for English or other non Chinese character scenes. For scenes containing Chinese character information, there are few public databases at present, and some of the public databases cannot meet the actual needs in terms of quantity or complexity of the situation. Therefore, after analyzing the existing image databases, this paper has built an image database. The entire database has about 800 scene images, Digital cameras and other photographic equipment.

### 4.2 Selection of the SVM Training Sample Set and the Training Process

In the experiments, a total of 1200 positive samples and 2000 negative samples were collected and mixed together, two-thirds of which were used as training sample images for the SVM classifier and the other one-third were used to test the classification model obtained from the training. The positive samples were obtained by randomly selecting 300 images from the training images in the self-built complex scene image library and manually segmenting the text characters in the images to include different sizes, colors, fonts, etc. as much as possible. By observing and analyzing the text characters in the images and after several tests, we normalize the sample size to 64\*64 pixels size, which can speed up the processing while providing the necessary texture information.

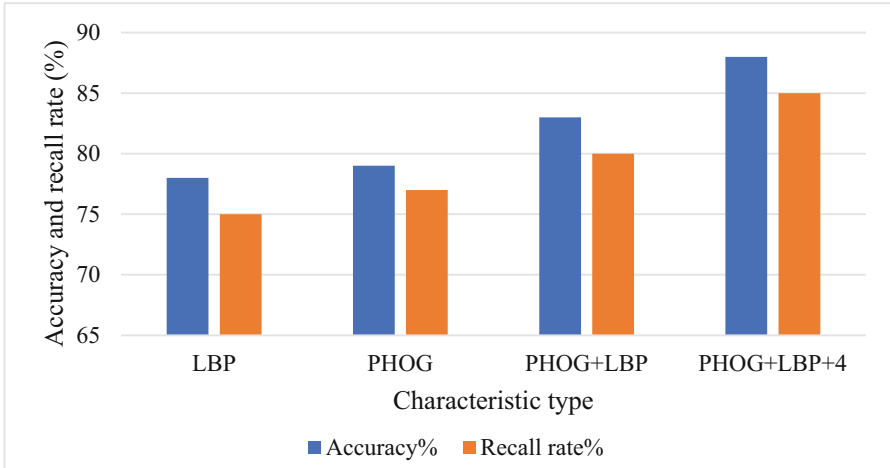
### 4.3 Algorithm Verification and Result Analysis

In the process of conducting experiments, the effectiveness of the combined features selected in this paper is tested first, then the performance of the localization algorithm in this paper is tested comprehensively, and then the DAM proposed in this paper is compared with other algorithms in detail.

Feature validity test. In order to illustrate the effectiveness of the combination features selected in this paper, the combination features of this paper were verified by experiments, and the ER show that the combination features selected in this paper can effectively express the text character features in the scene images. The ER are shown in Table 1, Fig. 3.

**Table 1.** ER

Characteristic type	Accuracy%	Recall rate%
LBP	78	75
PHOG	79	77
PHOG + LBP	83	80
PHOG + LBP + 4	88	85

**Fig. 3.** Feature Effectiveness Test Results

The ER of the above data show that the PHOG features extracted in this paper can describe the spatial distribution of the image contour, shape and other information by counting the gradient size and direction of the image pixels, and combined with the LBP features, which describe the local texture features by counting the size relationship between pixels in the image neighborhood, can effectively improve the feature detection effect of each region of the image, and then supplemented with The four overall texture features extracted from the image after wavelet transform can be used to effectively represent the features of text characters.

## 5 Conclusions

Regarding the study of TD and RS based on DAM under AIT, this paper investigates the detection method of DAM after analyzing the limitations of existing deep learning-based multi-directional TD algorithms. In the extraction of image information of natural scenes, TD and recognition are two very critical tasks. The algorithm studied in this paper accomplishes these two tasks to a certain extent, but there are still some problems that need to be improved and perfected in future research work. The algorithm mainly makes

targeted design and improvement in the feature extraction network part and detection part, however, the recall rate of the algorithm is not high. How to improve the recall rate while improving the speed and reducing the pre-set boxes is an urgent problem to be solved. Therefore, how to detect the text for large skew angle and how to correctly divide the text block content are the key contents of future TD research.

**Acknowledgements.** This work was supported by: Key Research Project of Guangdong Baiyun College, No. 2022BYKYZ02; Key Research Platform of Guangdong Province, No. 2022GCZX009; Special project in key fields of colleges and universities in Guangdong province, No. 2020ZDZX3009.

## References

1. Gupta, N., Jalal, A.S.: Tradition to transfer learning progression on scene TD and recognition: a survey. *Artif. Intell. Rev.* **55**(4), 3457–3502 (2022)
2. Dikubab, W., Liang, D., Liao, M., Bai, X.: Comprehensive benchmark datasets for Amharic scene TD and recognition. *Sci. China Inf. Sci.* **65**(6), 1–2 (2022)
3. Mithila, T., Arunprakash, R., Ramachandran, A.: CNN and fuzzy rules based TD and recognition from natural scenes. *Comput. Syst. Sci. Eng.* **42**(3), 1165–1179 (2022)
4. Naosekpan, V., Sahu, N.: TD, recognition, and script identification in natural scene images: a review. *Int. J. Multim. Inf. Retr.* **11**(3), 291–314 (2022)
5. Naiemi, F., Ghods, V., Khalesi, H.: Scene TD and recognition: a survey. *Multim. Tools Appl.* **81**(14), 20255–20290 (2022)
6. Pang, Y., et al.: Graph Decipher: a transparent dual-attention graph neural network to understand the message-passing mechanism for the node classification. *Int. J. Intell. Syst.* **37**(11), 8747–8769 (2022)
7. Long, S., He, X., Yao, C.: Scene TD and recognition: the deep learning era. *Int. J. Comput. Vis.* **129**(1), 161–184 (2021)
8. Kazmi, W., Nabney, I.T., Vogiatzis, G., Rose, P., Codd, A.: An efficient industrial system for vehicle tyre (Tire) detection and text recognition using deep learning. *IEEE Trans. Intell. Transp. Syst.* **22**(2), 1264–1275 (2021)
9. Raghunandan, K.S., Shivakumara, P., Roy, S., Hemantha Kumar, G., Pal, U., Lu, T.: Multi-script-oriented td and recognition in video/scene/born digital images. *IEEE Trans. Circuits Syst. Video Technol.* **29**(4), 1145–1162 (2019)
10. Rao, D., Huang, S., Jiang, Z., Deverajan, G.G., Patan, R.: A dual deep neural network with phrase structure and attention mechanism for sentiment analysis. *Neural Comput. Appl.* **33**(17), 11297–11308 (2021)
11. Gupta, N., Khosravy, M., Gupta, S., Dey, N., Crespo, R.G.: Lightweight AIT for health diagnosis of agriculture vehicles: parallel evolving artificial neural networks by genetic algorithm. *Int. J. Parallel Program.* **50**(1), 1–26 (2022)
12. Lee, S., Hwang, J., Cho, E.: Comparing technology convergence of artificial intelligence on the industrial sectors: two-way approaches on network analysis and clustering analysis. *Scientometrics* **127**(1), 407–452 (2021). <https://doi.org/10.1007/s11192-021-04170-z>





# Data Transmission and Isolation Method for Power Bus Communication

Yuxuan Ye<sup>(✉)</sup>, Rongtao Liao, Lei Li, Shengwei Wang, and Jian Zhang

Information and Communication Branch of State Grid, Hubei Electric Power Company,  
Wuhan 430000, Hubei, China  
ruo77yizhigan9@163.com

**Abstract.** This paper mainly studies the power grid monitoring and transmission system using CAN bus, including CAN bus, data information collection unit, data information monitoring unit and communication unit. The data information collection unit and communication unit are linked on the same CAN bus, while the data information monitoring unit is connected with the communication unit. Among them, CAN bus is used to complete the two-way transmission and multi-party transmission of data information; The data information collection unit is used to collect the data information of power grid devices in real time; The data information monitoring unit is used to collect, save and display the monitoring and management data information of all power grid equipment; The communication unit forms a communication module between the CAN bus and the data information monitoring unit to complete the two-way transmission of data information between the CAN bus and the data information monitoring unit. The communication unit also includes PCI communication module, processing module and CAN control module. This paper can also monitor the current, voltage and temperature control on the spot under the high-voltage environment, which points out the decision-making basis for the maintenance and repair of electrical equipment.

**Keywords:** Power · Bus · Communication · Data Transmission

## 1 Introduction

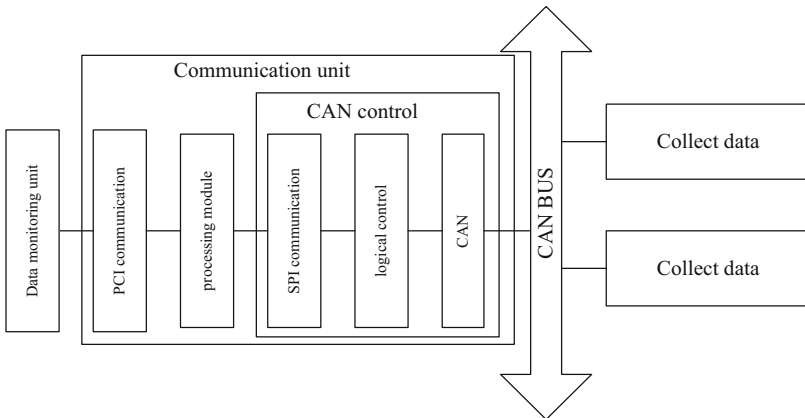
In recent years, the remote detection and control of electric power has been rapidly popularized and has attracted increasing attention from relevant personnel in operation, business management, technology and other posts [1]. The embedded technology has been applied in electric power because of its advantages of strong flexibility and tailoring operation [2]. Fieldbus detection and control is a way to apply embedded technology to power detection system [3].

Fieldbus technology turns a decentralized power grid detection and control system or device into a network node. Based on the field lines, they are connected into a network that can exchange messages and complete automatic control tasks together [4–15]. At present, the signal transmission technology of line based power grid detection and control system generally uses RS negative three or RS-485, but this technology has the

following disadvantages: the network is relatively closed, and there is no unified standard and communication protocol; The detection and control system is generally a slave architecture [16]. There is only one master node on the network, which cannot form a multi master redundancy system. If the master node has problems, the entire network will not work; The baud rate is low and the data transmission time is short, which can not meet the real-time requirements. The purpose of this paper is to solve the defects of the current power detection system, and propose a power monitoring transmission system using CAN bus. The control system conducts on-site detection and management of flow, pressure and ambient temperature in a high-voltage environment, providing decision-making basis for device maintenance and repair.

## 2 Power Monitoring Transmission System Based on CAN Bus

Figure 1 adopts the information electrical control transmission system of CAN bus, which is divided into CAN bus, information collection unit, information monitoring unit and communication unit [17]. The information collection unit and communication unit are all connected on the same CAN bus, while the information monitoring unit is connected with the communication unit. Among them, CAN bus is used for two-way transmission and multi-party transmission of information; The information collection unit is used to collect information about the pressure, flow and ambient temperature of electric equipment, and then send the information to the CAN bus; The information monitoring module is used to collect, save and provide monitoring and control information about power equipment; The communication module forms a communication mode between the CAN bus and the information monitoring module to complete the two-way transmission of information between the CAN bus and the information monitoring module.



**Fig. 1.** Structure diagram of power monitoring transmission system based on CAN bus

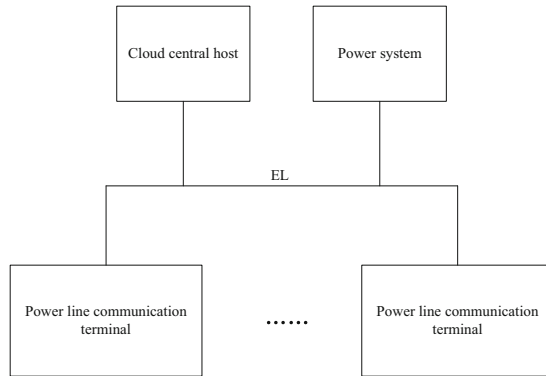
In this paper, each module in the whole control system is connected to a CAN bus, thus forming a logical layer connection. CAN is a bus type serial communication system, which can work through multi master mode, and can complete the mutual standby work

of multi master and multi system. As long as a device system node is located in the LAN, it can transfer information data with other nodes, and divide the power saving data of the device system into different levels through time [18]. The CAN transmission bus usually uses CRC detection method and has strong error correction capability, so the system can achieve good data communication. The digital monitoring device can truly show the data of the power equipment of the information collection unit to the monitoring personnel, so that the monitoring personnel can have a direct understanding of the situation of the accident site, and the management of the accident site is more convenient. The communication unit is a device inserted on the data monitoring unit to connect the CAN bus. It is mainly used for CAN communication and communication with the data monitoring unit. The communication unit basically uses the event driven method to complete its functions. This process is usually: after initialization and self inspection, add the main working cycle, wait for the event to trigger, and then use the event trigger method to call the function modules in the memory. When this paper is used, the information acquisition module first collects the current and voltage information of each line loop and between lines of the electrical equipment, and transmits the data to the communication module via the CAN bus, and then the communication module transmits the information to the operator of the data monitoring module, so as to monitor the voltage, flow and conditions of each part of the electrical equipment.

The communication unit in this paper is mainly divided into PCI communication module, processing module and CAN control module [19, 20]. The PCI communication module manages the sending and receiving of PCI bus data, while the CAN control module is responsible for the sending and receiving of data on the CAN bus, and the processing module is used to manage the digital communication and collaborative management between the PCI communication module and the CAN control module, and convert the data frame or control frame suitable for CAN communication or PCI communication. In order to meet the requirements of configuration and information exchange of power monitoring system, monitoring controller must be connected with a large number of communication modules. The data monitoring unit usually has a PCI interface, and the communication unit in this monitoring system has a matching PCI communication module to realize the communication of the data monitoring unit. Therefore, PCI 954 core chip can be selected. This core chip can meet the full functions of PCI master module and target module, transform the relatively complicated PCI bus connection into relatively simple user connection, and complete all hardware system connection signals and installation space registers required in PCI standard. The communication unit must occupy the center of the overall structure in this transmission system, and become the most critical part of the whole communication unit. The processing module must have strong performance and good stability. ARM processor can be selected. When MCU receives the data and control character recognition sent by PCI communication module, it will convert them into data frame or control frame of CAN bus after judgment and processing, and send them to CAN control module, and the CAN control module will automatically send a frame. When MCU receives the data frame or control frame sent by CAN communication interface, it will convert it into the data frame or control frame of PCI bus after judgment and processing, and send it to PCI communication module, which will automatically send it to the data monitoring unit.

### 3 Power Line Communication System

Figure 2 is the basic schematic diagram of the PLC device in this paper. Power supply line communication equipment includes cloud central host, power system PWR and communication terminal equipment of several power supply lines. In Fig. 2, the communication terminal equipment of the power supply line can be connected to the cloud central host and the power system PWR through the EL coupling of the power transmission line.



**Fig. 2.** Schematic diagram of power line communication system

In Fig. 2, the information between the data service central computer and the power line communication terminal device can be directly transmitted through the power network communication protocol using the power transmission line EL, thus simplifying the hardware structure of the power line communication network.

When using the power line communication terminal device and the power dedicated line communication system, the power or network flow used in all the home electronic devices connected with it can be recorded, so that in the actual use of smart home, users can easily monitor the changes in the power or network consumption of all the electronic devices in the smart home, so that the application scope of home electronic devices can be further standardized and adjusted; In commercial use, users can also easily measure and charge according to the electric energy of different electronic devices or the online usage, and can also use the electric energy of different electronic devices or the historical data of online usage to realize big data mining.

### 4 Power Line Communication Terminal

Figure 3 is the basic schematic diagram of the power network communication terminal device in this paper. In this paper, the power line communication terminal device can have the same basic architecture and working principle as the power network communication terminal device. The power network communication terminal device usually includes USB port, power network communication interface, universal serial bus unit, power network communication unit and processor.

Power network communication connection can be coupled by power transmission network EL and power system PWR. In this paper, the power supply line communication connector can also use the general power plug. When the distribution network communication connector enters the power socket, it can use the power transmission line EL that has been erected by the power company to couple with the power system PWR, so as to obtain the normal power supply.

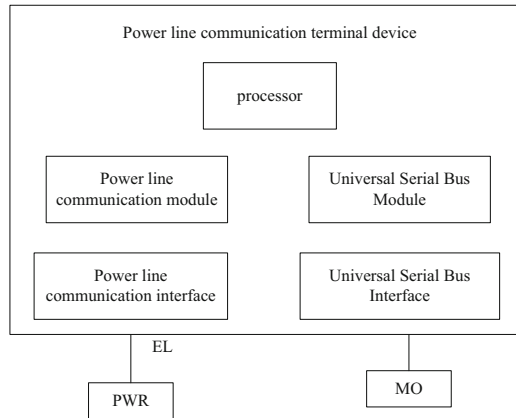
The power line communication system can package the data transmitted from the processor, and then transmit the packaged data to any equipment coupled with the power transmission line EL through the power line communication interface or power transmission line EL through the power line communication protocol to achieve data transmission and network communication. In Fig. 3, the power network communication system can send the packet data to the cloud central server coupled with the power supply transmission line EL. In this way, the cloud central server can obtain the information transmitted by the power network communication terminal device, and can conduct further data processing and calculation.

In this paper, when the common serial bus interface is coupled with the peripheral electronic device Mzero, the common serial bus module can charge the peripheral human resource electronic devices via the common serial bus interface according to the power supply requirements of the peripheral electronic device Mzero.

In terms of USBPD specification, the power port through the serial bus can coordinate with the power receiving port. In the process of cooperation, the power receiving port can tell the power terminal what its required power supply or acceptable power supply specification is, and the power terminal can also tell the power receiving port what its available power supply specification is. After the cooperation between the two is completed, the power terminal can supply power to the power receiving port according to the power requirements of the power receiving terminal. In this paper, since the universal serial bus interface, the universal serial bus module and the external electronic devices can support the input requirements of the universal serial bus, the universal serial bus module can also follow the process of the input specification requirements of the universal serial bus. In the process of communication and coordination with the external electronic devices, after obtaining the power supply requirements of the external electronic devices, Then the USB interface directly recharges the external electronic devices.

Because, in the power transmission specification of the universal serial bus, the universal serial bus interface can provide CC pin in addition to the pin previously required in the common serial bus standard, so on the premise that the common serial bus port, the common serial bus module and the peripheral electronic devices fully meet the power transmission specification of the common serial bus, The ordinary serial bus module can also directly verify the human resource electronic devices other than the configuration of the ordinary serial bus port and the channel pins, such as the identification and comparison of product categories. After the verification is completed, the ordinary serial bus module can also charge the external electronic devices directly through the ordinary serial bus port, which can more ensure the security of the application.

The data processor can also obtain and generate a message packet according to the power supply demand and/or charging time of the external electronic device, and transmit



**Fig. 3.** Schematic diagram of power line communication terminal device

the message packet to the communication module of the power supply line. While the power line communication module can transmit the information in the information packet to the cloud server or central host through the power line communication interface according to the power line communication protocol when receiving the information packet.

In this paper, the power line communication terminal device can receive the network packet information sent by the external electronic device through the universal serial bus, and then use the power supply network communication method to transmit the information about the Internet packet, that is, by using the power line communication terminal device, the external electronic device can directly realize packet exchange with the Internet, thus realizing the network function.

For example, the USB module can receive network packets from external electronic devices through the USB interface. The processor can also include a network adapter that supports RNDIS, and can implement the network layer and application layer. In this way, when external electronic devices transmit network packets through the universal serial bus port, the network adapter of RNDIS can directly convert network packets into packets that support TCP/IP, and can further analyze the content of network packets through the network layer and application layer of the processor. After analyzing the content of the network packet, the processor can re packet the content of the network packet according to its actual use needs through the application layer or network layer, and transmit it to the communication module of the power supply line. The power line network module can reorganize the internal network packets and make them into new packets more suitable for the power line connection platform, and then transmit them to the Internet through the power line communication interface.

After the network packets collected by the communication connector of the power supply line, the communication components of the power supply line can be analyzed and processed by the processor, transmitted to the universal serial bus module, and transmitted to the peripheral electronic technology equipment through the universal

serial bus connector. In this way, external electronic devices can directly exchange data with the Internet, thus meeting the online function of external electronic devices.

In addition, the processor can also obtain the network transmission data used by external electronic devices from the above process, and use the power line communication module or power line communication interface to transmit the network transmission data to the cloud server or central server. In this way, the cloud server and the central server will be able to understand the amount of data transmitted by the external electronic devices coupled to the power line communication terminal equipment, and can further control and charge according to the amount of data transmitted by the network.

Because any power line communication terminal device can, as from the power line communication terminal device, send the data of power supply time and/or charging time used by external electronic equipment and facilities to the cloud central host, so that the cloud central host can count the usage of each external electronic equipment and facilities coupled to the power line communication terminal device according to the data transmitted from the power line communication terminal device, And make further control and charges.

## 5 Conclusions

Compared with the current information transmission technology, this paper has the following important roles: because this paper uses the CAN bus network, according to the characteristics of good autonomy of the CAN bus, this system adds a device node or removes a device node in the serial bus, which does not cause any interference to the normal work of the entire system. In addition, the network has high communication frequency and good real-time data communication characteristics, The monitoring system has good independence, stability and free performance. In addition, with the help of the establishment of information collection module, information monitoring module and communication module, we will establish a decentralized, two-way, interconnected and interactive data transmission and control system. The system has unified network layout, low cost, easy connection, and is convenient for building a strong system monitoring and management network.

**Acknowledgement.** State Grid Hubei Electric Power 2022 Power Communication Information Security Interaction Technology Research and Application Technology Project under the cloud edge collaboration framework.

## References

1. Ghorbani, M., Dehmer, M., Zangi, S.: Graph operations based on using distance-based graph entropies. *Applied Mathematics and Computation* (2018)
2. Parisot, S., et al.: Disease prediction using graph convolutional networks: Application to Autism Spectrum Disorder and Alzheimer's disease. *Medical Image Analysis* (2018)
3. Singh, M.: Protection coordination in distribution systems with and without distributed energy resources- a review. *Protection Control Modern Power Syst.* **2017**(1), 1–17 (2017)

4. Saleh Al-rimy, B.A., Maarof, M.A., Mohd Shaid, S.Z.: Ransomware threat success factors, taxonomy, and countermeasures: a survey and research directions. *Computers & Security* (2018)
5. Suchita, D., Fabio Di, T., Mark, S.: Vigenère scores for malware detection. *J. Comput. Virology Hacking Tech.* **2018**(2), 157–165 (2018)
6. Coutinho, M., de Oliveira Albuquerque, R., Borges, F., Villalba, L.J.G., Kim, T.-H.: Learning Perfectly Secure Cryptography to Protect Communications with Adversarial Neural Cryptography. *Sensors* 2018(5)
7. Shah, S.A.R., Issac, B.: Performance comparison of intrusion detection systems and application of machine learning to Snort system. *Future Generation Computer Systems* (2018)
8. Vijayakumar, P., Chang, V., Jegatha Deborah, L., Balusamy, B., Shynu, P.G.: Computationally efficient privacy preserving anonymous mutual and batch authentication schemes for vehicular ad hoc networks. *Future Generation Computer Systems* (2018)
9. Andreas, B., Anupam, C., Yan, L.K.: Wireless Communication and Security Issues for Cyber-Physical Systems and the Internet-of-Things. *Proceedings of the IEEE* 2018(1) (2018)
10. Gholizadeh, N., Saadatfar, H., Hanafi, N.: K-DBSCAN: an improved DBSCAN algorithm for big data. *J. Supercomput.* **77**(6), 6214–6235 (2020). <https://doi.org/10.1007/s11227-020-03524-3>
11. Scitovski, R., Sabo, K.: A combination of k -means and DBSCAN algorithm for solving the multiple generalized circle detection problems. *Advances in Data Analysis and Classification* 2020 (prep)
12. Govender, P., Sivakumar, V.: Application of k -means and hierarchical clustering techniques for analysis of air pollution: a review (1980–2019). *Atmospheric Pollution Research* 2020(1) (2020)
13. Kang, J.S., et al.: Development of a systematic, self-consistent algorithm for the K-DEMO steady-state operation scenario. *Nuclear Fusion.* 2017(12)
14. Doostan, M., Chowdhury, B.H.: Power distribution system fault cause analysis by using association rule mining. *Electric Power Systems Research* (2017)
15. Longjun Wang, A., Xiaomin Chen, B., Gang Wang, C., Hua, D.: Non-intrusive load monitoring algorithm based on features of V–I trajectory. *Electric Power Systems Research* (2018)
16. Yasir Saleem 0001, No?l Crespi, Mubashir Husain Rehmani, Rebecca Copeland. *Internet of Things-Aided Smart Grid: Technologies, Architectures, Applications, Prototypes, and Future Research Directions.* *IEEE Access* (2019)
17. Pablo, C., Albert, B., de Gustavo, V., Xavier, C.P., Arturo, A.: Network slicing for guaranteed rate services: admission control and resource allocation games. *IEEE Trans. Wirel. Commun.* **2018**(10) (2018)
18. Yoon, S., Lee, Y.-J., Jung, H.-J.: A comprehensive framework for seismic risk assessment of urban water transmission networks. *Int. J. Disaster Risk Reduction* (2018)
19. Berrueta, A., Urtasun, A., Ursúa, A., Sanchis, P.: A comprehensive model for lithium-ion batteries: From the physical principles to an electrical model. *Energy* (2018)
20. Hooshyar, D., Pedaste, M., Saks, K., Leijen, L., Bardone, E., Wang, M.: Open learner models in supporting self-regulated learning in higher education: a systematic literature review. *Computers & Education.* 2020 (prep)





# Configuration Method of Safety Link of Electric Power Communication Network

Huanjun Hu<sup>(✉)</sup>, Shengwei Wang, Yixi Wang, Lei Li, and Hao Ning

Information and Communication Branch of State Grid Hubei Electric Power Company,  
Wuhan 430000, Hubei, China  
pushi743090gwo@163.com

**Abstract.** This paper studies a power communication link establishment method, resource allocation method and system. The power communication link establishment method includes: sending a configuration request to the configuration server until the configuration server determines that the sender meets the access conditions of the communication link to be accessed according to the security requirement list and the received configuration request, and the sender establishes a connection with the communication link to be accessed. By using this technology, the power service terminal can select the corresponding communication link according to the security preference, and send the configuration request to the configuration server; The configuration server then reversely determines whether the power business terminal meets the access conditions according to the security requirements list, and finally enables the power business terminal to access the communication link that meets its own security needs and preferences, meeting the differentiated security requirements of the power business terminal for the communication link resources.

**Keywords:** Power Communication · Configuration Link · Server

## 1 Introduction

In the field of electric power communication, due to the complexity of electric power communication business, the risk level and planning strategy of electric power communication network are different from those in other fields, and existing research results in other fields cannot be used [1–3]. It is necessary to develop targeted and personalized methods according to their unique characteristics [4–7]. The key to reduce the operation risk of power communication network is how to equalize the communication business risks of different levels in the communication network, that is, how to find a safe link that can always maintain the lowest overall risk balance of power communication network is an urgent problem to be solved at present. At present, the maximum channel gain method is generally used to deal with the problem of network resource allocation [8]. When different link nodes forward, the channel gain of each node is identified to complete the selection of the next node, so as to achieve the transmission of service data [9]. However, this method ignores the security of communication link resources and does not

take into account the differentiated security requirements of different service terminals. This paper provides a method of establishing power communication link, a method of resource allocation and a system to solve the technical problem of differentiated security requirements of service terminals for link resources.

## 2 Process of Resource Allocation Method

Figure 1 is a flow diagram of the power communication resource allocation method provided in this paper. The details are as follows:

Step (a): Receive the configuration request sent by the sender.

Step (b): Query the security requirement list according to the terminal information and link information, and determine the first security requirement level of the sender and the first access level corresponding to the communication link to be accessed.

In this paper, the security requirement list records the security requirement level of each terminal equipment and the access level required by each communication link.

Step (c): judge whether the first security requirement level is less than or equal to the first access level; If the first security requirement level is less than or equal to the first access level, perform step (d); If the first security requirement level is greater than the first access level, perform step 404;

Step (d): Confirm that the sender meets the access conditions of the communication link to be accessed; Perform step (f).

Step (e): Determine that the sender does not meet the access conditions of the communication link to be accessed.

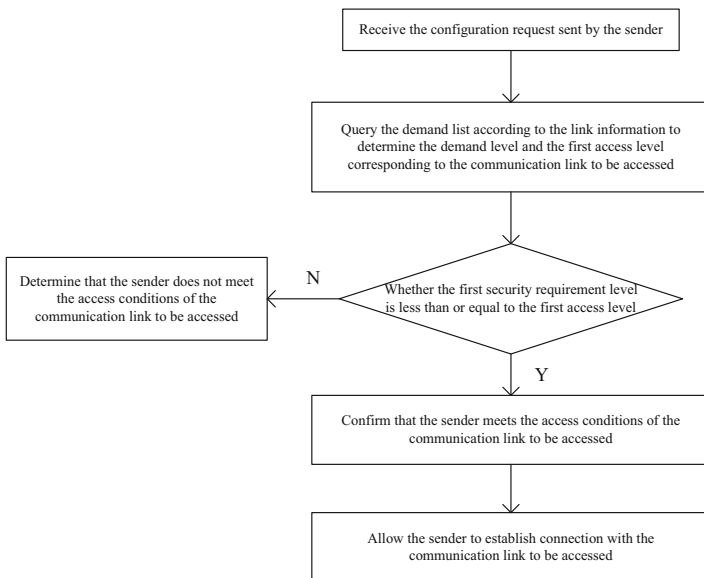


Fig. 1. Flow of power communication resource allocation method

Step (f): Allow the sender to establish a connection with the communication link to be accessed.

In this paper, after determining that the sender does not meet the access conditions of the communication link to be accessed if the first security requirement level is greater than the first access level, it also includes: when determining that the sender does not meet the access conditions of the communication link to be accessed according to the security requirement list and the received configuration request, the sender is refused to establish a connection with the communication link to be accessed [10–18]; According to the first access level, obtain a number of communication links that do not meet the access conditions, including the communication links to be accessed, and generate feedback results based on a number of communication links that do not meet the access conditions [19]; Send feedback results to the sender, so that the sender can update the current security preference list according to the feedback results.

The configuration server in this paper judges whether the power business terminal meets the access conditions according to the security requirement list containing the security requirement level of each power business terminal, and matches the communication link that meets the security requirement level for the power business terminal [20]; In addition, when the sender does not meet the conditions for establishing a connection with the communication link to be accessed, it rejects the sender’s connection establishment request and sends the feedback results in a timely manner, so that the sender can update the current security preference list and improve the matching efficiency between the power business terminal and the communication link.

### 3 Structure of Resource Allocation System

Figure 2 is a structural diagram of the power communication resource configuration system provided in this paper, which mainly includes: configuration server, power service terminal and communication link resources.

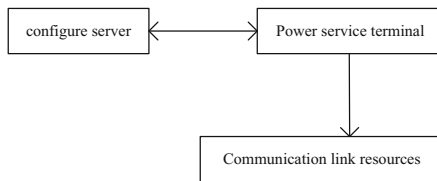


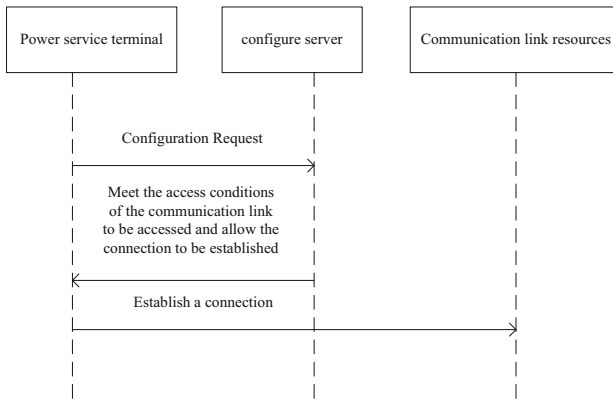
Fig. 2. Structure of power communication resource allocation system

In this article, the configuration server executes any power communication resource configuration method as described in this article. The power service terminal performs any of the power communication link establishment methods described herein. The communication link resource includes several communication links including the communication link to be accessed, which is used to establish a connection with the power service terminal when the configuration server determines that the power service terminal meets the access conditions of the communication link to be accessed.

The power service terminal in this paper establishes a security preference list through the analytic hierarchy process, selects the corresponding communication link from the communication link resources according to the security preference, and sends the configuration request to the configuration server; The configuration server then reversely determines whether the power business terminal meets the access conditions according to the security requirements list, and finally enables the power business terminal to access the communication link that meets its own security needs and preferences. Through the bilateral matching mode, the differentiated security requirements of the power business terminal for the communication link resources are met.

#### 4 Timing of Communication Link Establishment Method

In order to better illustrate the step flow of this paper, Fig. 3 is a sequence diagram of the power communication link establishment method provided in this paper. The power service terminal sends a configuration request to the configuration server to enable the power service terminal to establish a connection with the communication link to be accessed; When the configuration server determines that the power service terminal meets the access conditions of the communication link to be accessed, the power service terminal is allowed to establish a connection with the communication link to be accessed.



**Fig. 3.** Timing of power communication link establishment method

In this paper, the current security preference list is updated according to the feedback results of the last configuration request. Specifically, when the sender does not meet the access conditions of the communication link to be accessed, it receives the feedback results sent by the configuration server; The feedback result records several communication links that do not meet the access conditions, including the communication link to be accessed; The access level corresponding to several communication links that do not meet the access conditions is less than or equal to the first access level corresponding to the communication link to be accessed; According to the feedback results, the current security preference list is updated by deleting the communication links to be accessed or several communication links that do not meet the access conditions.

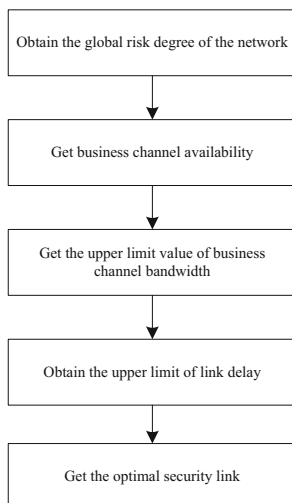
In this paper, through the security preference list, the power service terminal selects the corresponding communication link according to the security preference, and sends the configuration request to the configuration server; The configuration server then reversely determines whether the power business terminal meets the access conditions according to the security requirements list, and finally enables the power business terminal to access the communication link that meets its own security needs and preferences, meeting the differentiated security requirements of the power business terminal for the communication link resources.

This paper can also delete the communication links that do not meet the access conditions by updating the current security preference list, so that the power business terminal can match the communication link corresponding to its security requirement level, and improve the efficiency of establishing the connection between the power business terminal and the communication link.

## 5 Structure of Processor Hardware

Figure 4 shows the structure diagram of a processor provided in this paper. As shown in Fig. 4, this processor is used to implement the above power communication network security link configuration method. The processor comprises: a network global risk degree acquisition module; Business channel availability acquisition module; Business channel bandwidth upper limit value acquisition module; Link delay upper limit value acquisition module and optimal security link acquisition module. Among them, the network global risk degree acquisition module is used to obtain the network global risk degree; The business channel availability acquisition module is used to obtain the business channel availability; The upper limit value acquisition module of the service channel bandwidth is used to obtain the upper limit value of the service channel bandwidth; The upper limit value acquisition module is used to acquire the upper limit value of the link delay; The optimal security link acquisition module is used to acquire the optimal security link.

The application of the above processor can effectively reduce the overall risk balance of power communication network, and has high availability and effectiveness in reducing transmission delay and link calculation time, which is helpful to configure the optimal security link of power communication network. At the same time, quantitative calculation to determine the importance of services has important theoretical and practical significance for analyzing the impact of communication services, equipment, etc. on the reliability of communication networks.



**Fig. 4.** Processor structure

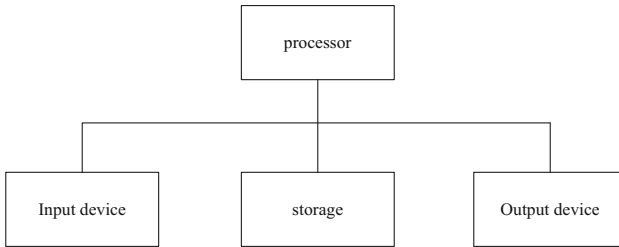
As shown in Fig. 5, the electronic device includes one or more of the above processors and memories. The processor may be a central processing unit or other form of processing unit having data processing capability and/or instruction execution capability, and may control other components in the electronic device to perform a desired function. The memory may include one or more computer program products, which may include various forms of computer-readable storage media, such as volatile memory and/or non-volatile memory. The volatile memory may include, for example, random access memory (RAM) and/or cache memory. The nonvolatile memory may include, for example, a read-only memory (ROM), a hard disk, a flash memory, and the like. One or more computer program instructions can be stored on the computer-readable storage medium, and the processor can run the program instructions to realize the power communication network security link configuration method and/or other desired functions in the above cases.

In one example, the electronic device may also include an input device and an output device, and these components are interconnected through a bus system and/or other forms of connection mechanisms. When the electronic device is a stand-alone device, the input device can be a communication network connector for receiving the collected input signals from the first device and the second device.

In addition, the input device may also include, for example, a keyboard, a mouse, and the like. The output device can output various information to the outside, including the determined distance information, direction information, etc. The output device may include, for example, a display, a speaker, a printer, a communication network, a remote output device connected thereto, and the like.

Of course, to simplify, only some of the components related to this article in the electronic device are shown in Fig. 4, and components such as buses, input/output interfaces and the like are omitted. In addition, the electronic device may include any other appropriate components, depending on the specific application.

In addition, a computer-readable storage medium is provided herein. The storage medium stores a computer program, which is used to perform the following steps: acquiring the global risk degree of the network; Obtain the service channel availability; Obtain the upper limit value of service channel bandwidth; Obtain the upper limit value of link delay; The optimal security link is obtained according to the global risk degree of the network, the availability of the service channel, the upper limit of the service channel bandwidth and the upper limit of the link delay.



**Fig. 5.** Structure of electronic equipment

## 6 Conclusions

This paper provides a security link configuration method for power communication network, including obtaining the global risk degree of the network; Obtain the service channel availability; Obtain the upper limit value of service channel bandwidth; Obtain the upper limit value of link delay; Obtain the optimal security link according to the global network risk, service channel availability, upper limit of service channel bandwidth and upper limit of link delay; Among them, when the network global risk degree of the link is minimum and the service channel availability is maximum, the link is the optimal security link. The above methods can effectively reduce the overall risk balance of power communication network, and have high availability and effectiveness in reducing transmission delay and link calculation time, which is helpful to configure the optimal security link of power communication network. At the same time, quantitative calculation to determine the importance of services has important theoretical and practical significance for analyzing the impact of communication services, equipment, etc. on the reliability of communication networks.

**Acknowledgement.** State Grid Hubei Electric Power 2022 Power Communication Information Security Interaction Technology Research and Application Technology Project under the cloud edge collaboration framework.

## References

1. Motalleb, M., Siano, P., Ghorbani, R.: Networked Stackelberg competition in a demand response market. *Appl. Energy* (2019)

2. Andoni, M., et al.: Blockchain technology in the energy sector: a systematic review of challenges and opportunities. *Renew. Sustain. Energy Rev.* (2019)
3. Ryu, U., Wang, J., Kim, T., Kwak, S., Juhyok, U.: Construction of traffic state vector using mutual information for short-term traffic flow prediction. *Transport. Res. Part C* (2018)
4. Noor, S., Yang, W., Guo, M., van Dam, K.H., Wang, X.: Energy demand side management within micro-grid networks enhanced by blockchain. *Appl. Energy* (2018)
5. Suchita, D., Di Fabio, T., Mark, S.: Vigenère scores for malware detection. *J. Comput. Virol. Hack. Tech.* (2) (2018)
6. Coutinho, M., de Oliveira Albuquerque, R., Borges, F., Villalba, L.J.G., Kim, T.-H.: Learning perfectly secure cryptography to protect communications with adversarial neural cryptography. *Sensors* (5) (2018)
7. Shah, S.A.R., Issac, B.: Performance comparison of intrusion detection systems and application of machine learning to Snort system. *Future Gener. Comput. Syst.* (2018)
8. Vijayakumar, P., Chang, V., Jegatha Deborah, L., Balusamy, B., Shynu, P.G.: Computationally efficient privacy preserving anonymous mutual and batch authentication schemes for vehicular ad hoc networks. *Future Gener. Comput. Syst.* (2018)
9. Andreas, B., Anupam, C., Yan, L.K.: Wireless communication and security issues for cyber-physical systems and the internet-of-things. In: *Proceedings of IEEE* (1) (2018)
10. Gholizadeh, N., Saadatfar, H., Hanafi, N.: K-DBSCAN: an improved DBSCAN algorithm for big data. *J. Supercomput.* **77**(6), 6214–6235 (2020). <https://doi.org/10.1007/s11227-020-03524-3>
11. Scitovski, R., Sabo, K.: A combination of k-means and DBSCAN algorithm for solving the multiple generalized circle detection problems. *Adv. Data Anal. Classif.* (2020, prep)
12. Govender, P., Sivakumar, V.: Application of k-means and hierarchical clustering techniques for analysis of air pollution: a review (1980–2019). *Atmos. Pollut. Res.* (1) (2020)
13. Kang, J.S., et al.: Development of a systematic, self-consistent algorithm for the K-DEMO steady-state operation scenario. *Nuclear Fusion* (12) (2017)
14. Doostan, M., Chowdhury, B.H.: Power distribution system fault cause analysis by using association rule mining. *Electr. Power Syst. Res.* (2017)
15. Wang, A.L., Chen, B.X., Wang, C.G., Huo, D.: Non-intrusive load monitoring algorithm based on features of V–I trajectory. *Electr. Power Syst. Res.* (2018)
16. Saleem, Y., Crespi, N., Rehmani, M.H., Copeland, R.: Internet of things-aided smart grid: technologies, architectures, applications, prototypes, and future research directions. *IEEE Access* (2019)
17. Caballero, P., Banchs, A., De Veciana, G., Costa-Pérez, X., Azcorra, A.: Network slicing for guaranteed rate services: admission control and resource allocation games. *IEEE Trans. Wirel. Commun.* (10) (2018)
18. Dorsch, N., Kurtz, F., Wietfeld, C.: On the economic benefits of software-defined networking and network slicing for smart grid communications. *NETNOMICS: Econ. Res. Electron. Network.* **19**(1–2), 1–30 (2018). <https://doi.org/10.1007/s11066-018-9124-3>
19. Ahmad, A., Rehmani, M.H., Tembine, H., Mohammed, O.A., Jamalipour, A.: IEEE access special section editorial: optimization for emerging wireless networks: IoT, 5G, and smart grid communication networks. *IEEE Access* (2017)
20. Aitzhan, N.Z., Svetinovic, D.: Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams. *IEEE Trans. Depend. Secure Comput.* (5) (2018)





# Method for Locating Communication Device Faults

Lei Li<sup>(✉)</sup>, Zheng Zhou, Yixi Wang, Shengwei Wang, and Jian Zhang

Information and Communication Branch of State Grid, Hubei Electric Power Company,  
Wuhan 430000, Hubei, China  
ziyongsha8532y@163.com

**Abstract.** This paper provides a communication network fault location method and a fault monitoring device. The technical solution first grabs data from the error log of the communication network management system, and associates the reporting time with the error category. At the same time, the peak value of such errors is taken as the time interval to obtain the network access volume at each time point, and the access volume is compared with the error occurrence rate to obtain the key time node. The network environment within the time node has a greater potential impact on the occurrence of such errors. On this basis, this paper divides the process of the same type of error into several access behaviors, and uses the terminal equipment to repeatedly test each access behavior under the above network environment to collect the access behavior and its frequency of the same type of error, so as to determine the fault location. This paper determines the communication network fault by trial and error, and fully restores the network environment, so that the fault location is more accurate.

**Keywords:** Communication · Fault Location · Data Capture

## 1 Introduction

Communication network failure refers to the condition that the network cannot provide normal service or reduce the quality of service due to hardware problems, software vulnerabilities, virus intrusion, etc. [1]. When the communication network fails, the failure phenomenon shall be reported to the network management system as an abnormal signal and recorded in the report [2–7]. The network maintenance personnel shall evaluate the cause and level of the fault and locate it according to the abnormal signals recorded in the report and their own experience [8].

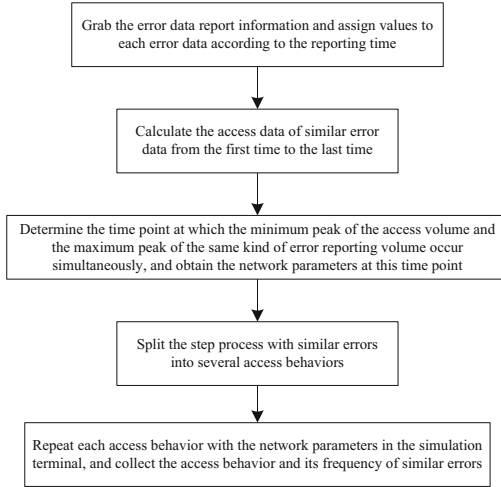
In general, the management system of large-scale communication network receives fault information more frequently, so it consumes a lot of professional human resources to analyze, which is not only inefficient, but also may have errors in manual analysis [9–12]. In addition, since the reported information is automatically captured based on software and is not differentiated according to the fault category and level, there are problems such as duplication, redundancy and concealment between them, which makes it difficult to locate the true cause of network faults [13]. In response to such problems,

researchers in the existing technology try to develop automatic positioning methods around communication network failures [14–18]. For example, they use software tools to count abnormal data, classify and summarize error data according to error category, time of occurrence, IP location and other information, screen out repeated reporting behavior caused by communication delay, and then summarize a large number of redundant error information to form key errors. Thus, repetitive work is avoided [19]. Another example is to monitor the logs of each terminal system connected to the communication network in real time, so as to monitor the frequency of access behavior within a unit time, determine the degree of abnormality according to the set threshold value, so as to provide quantifiable fault indicators to the staff and provide reference for subsequent manual analysis.

Although the above methods reduce the workload of fault analysis to a certain extent, they still require personnel to locate faults based on abnormal characteristics. In this case, if automatic fault location can be realized based on software tools, it is expected to improve the efficiency of fault troubleshooting and improve the security of communication networks. The purpose of this paper is to provide a communication network fault location method and fault monitoring device against the technical defects of the prior art, so as to solve the technical problem that the prior art is difficult to automatically locate the network fault according to the abnormal characteristics.

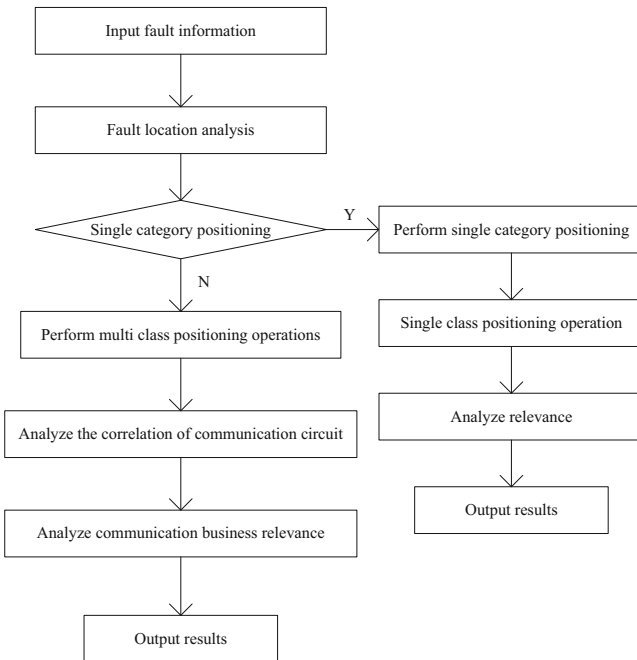
## 2 Communication Network Fault Location Process

The communication network fault location method, as shown in Fig. 1, includes: connecting with the communication network management system, capturing the error data report information, and assigning values to each error data according to the reporting time; [20] Calculate the access data of similar error data from the first time to the last time; Determine the time point at which the minimum peak of the access volume and the maximum peak of the same kind of error reporting volume occur simultaneously, and obtain the network parameters at that time point; Divide the step process with similar errors into several access behaviors; Repeat the access behavior with network parameters in the simulation terminal, and collect the access behavior and its frequency of similar errors.



**Fig. 1.** Fault location method of communication equipment

After the alarm correlation analysis, the power communication network fault management system can get the real fault information. The intelligent analysis of fault location refers to locating the real fault to the site location, physical devices, logical circuits



**Fig. 2.** Flow Chart of Fault Location Intelligent Analysis

and communication services. Through positioning, the site location, involved equipment, circuit connection and business of the fault can be accurately known after the fault occurs, and the intelligent analysis results can be displayed. The overall process is shown in Fig. 2.

It can be seen from Fig. 2 that the intelligent analysis of fault location mainly includes the following four parts.

(A) Failure of the site location

Fault location of locating site refers to locating the fault equipment to the corresponding network site through the corresponding mathematical logic relationship.

(B) Multi category location of equipment failure

The fault location of equipment is not only for the site location, but also from different angles to achieve multi category location of equipment faults. In the actual fault location of the equipment, the logical relationship between the included elements can be determined according to the function and purpose of the equipment, such as all of the optical transceiver, all of the power supply, and all of the switches. At the same time, it can also be positioned according to the model, type and line. If it is located by site, for a certain site, it contains the complete set of local devices.

However, if it is located by equipment type, such as optical transmission equipment, it will become the complete set of all optical transmission equipment, and the scope may involve each node of the whole network, while the optical transmission equipment of a certain manufacturer will constitute a subset of the complete set. In this way, double positioning is required, that is, first locate the optical transmission equipment, and then locate the equipment manufacturer. The more location classes associated with the device, the location of the fault can be obtained from all angles after the fault occurs.

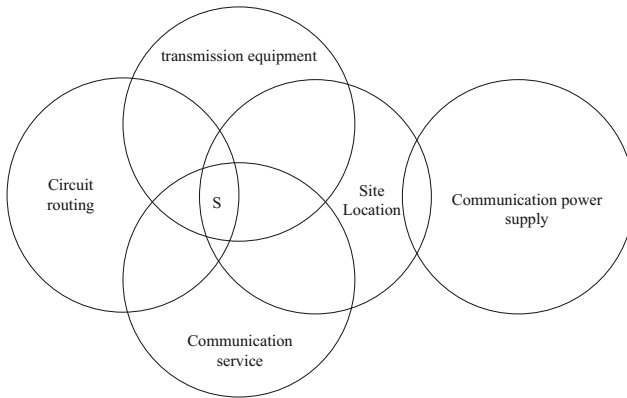
(C) Carrier failure

Carrier fault location refers to locating the fault equipment to the circuit routing and other communication carriers, that is, analyzing the correlation between fault information and communication circuits.

(D) Fault location of communication service

The fault location of communication service is the correlation analysis of fault information and communication service. The circuit route is directly related to the fault location of communication services. When all the circuit routes corresponding to a communication service fail, the communication service fault is located.

All object information associated with sites, devices and devices in the communication network is massive. How to face different location associated objects and quickly find their respective ownership is one of the keys to multiple location of equipment failures. If logical correlation is carried out one by one, it will lead to large amount of fault location calculation, slow speed and low intelligence.



**Fig. 3.** Schematic Diagram of Multi type Fault Location of Equipment

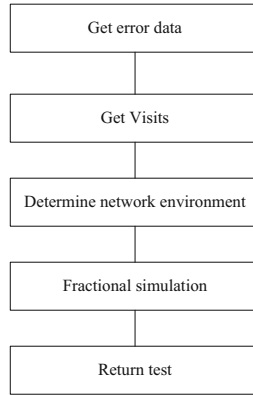
In this paper, the intersection operation is used to realize the one-time multi class location of equipment faults. When the fault of some equipment involves multiple positioning classes, it will form a non empty set intersection, and the non involved ones will not intersect and form an empty set. Based on this idea, as long as the device is associated with which positioning classes, all classes can be located through one-time intersection operation, as shown in Fig. 3. S in the figure refers to synchronous digital series optical transceiver equipment.

In Fig. 3, the transmission equipment, site location, communication location and circuit route of S and positioning class intersect. All the equipment meeting this intersection constitute the intersection of the four positioning classes. The communication power supply equipment in the figure can also be said to intersect with the four positioning classes, but its intersection with transmission equipment, communication services and circuit routes is an empty set, and it only forms another intersection with the site location, which belongs to single class positioning. Therefore, once the relationship between the device and the location class is determined, all intersecting classes will be located in one operation.

### 3 Structure of Communication Equipment Fault Location System

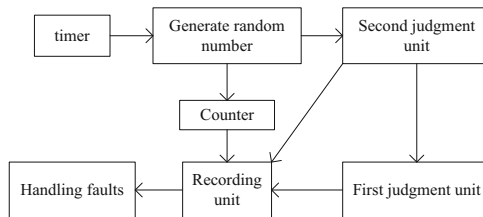
A communication network fault monitoring device, as shown in Fig. 4, comprises: an error data acquisition unit, which is used to connect with the communication network management system, capture the error data report information, and assign values to each error data according to the reporting time; Access amount acquisition unit: used to calculate the access amount data of the same type of error data from the first time to the last time; Network environment determination unit: used to determine the time point at which the minimum peak of access and the maximum peak of similar error reports occur simultaneously, and obtain the network parameters at this time point; Step by step simulation unit: used to split the step process with similar errors into several access behaviors; Loopback test unit: Repeat the access behavior with network parameters in

the simulation terminal, and collect the access behavior and its frequency of similar errors.



**Fig. 4.** System structure of fault location device

This paper relates to a fault location system in communication equipment. Its structure diagram is shown in Fig. 5, including a random number generation unit, which is used to call the CPM module in the embedded processor to generate random numbers; Timer, which is used to periodically instruct the random number generation unit to call the CPM module to generate random numbers. The timer is a periodic timer, and the period of the timer is less than or equal to 100 ms; A counter for counting the random numbers generated by the random number generation unit; The first judgment unit is used to judge whether the consecutive  $N$  random numbers obtained from the CPM module are the same. If they are the same, the CPM module is judged to be faulty, where  $N$  is greater than or equal to 1.



**Fig. 5.** System Structure of Communication Equipment Fault Location

The system can also include a second judgment unit, which is used to judge whether the random number generated by the communication processing module is obtained within a predetermined time limit after the random number generation unit calls the communication processing module to generate the random number. If the random number is not obtained after the timeout, the communication processing module is determined to be faulty.

The system can also include a recording unit, which is used to record the failure information of the CPM module when the first or second judgment unit determines the failure of the communication processing module; The fault processing unit is used for fault processing when the first or second judgment unit determines that the CPM module is faulty. The fault processing unit can include a reset sub unit for resetting the boards corresponding to the communication processing module.

The system can regularly monitor the working state of the CPM module. When the working state of the CPM module is abnormal, the system will record the log and give an alarm at the first time, and handle it accordingly. It is convenient to locate the problem of system reset caused by the abnormal working state of the CPM module, and can notify the board to reset in the shortest possible time to resume normal business.

It should be noted that in this way, in order to improve the accuracy of fault judgment, the number of random numbers involved in fault judgment is greater than 1, that is, when the random numbers generated for two or more consecutive times are the same, the CPM module fault is judged; In practical applications, CPM module failure is usually determined when the random numbers generated for three consecutive times are the same.

To sum up, in this paper, we call the random number generated by the CPM module regularly, and judge whether the CPM module is faulty according to whether the random number generated is the same for consecutive times; Since the random number is generated by the CPM module hardware, the failure of the random number generation function can be directly identified as the failure of the CPM module, and will not be confused with the failure of other software or hardware, which can facilitate and accurately locate the failure of the CPM module.

Since the random number is generated by the CPM module hardware, there is almost no additional burden on the equipment, so the detection can be completed in a few milliseconds or tens of milliseconds, and it is usually known whether the CPM module has a fault within 100 ms; However, in the prior art, a fault can only be found when a certain type of protocol has not been properly handled for a certain period of time, which usually takes several seconds or even minutes; It can be seen that compared with the prior art, this technology can locate the failure of the CPM module more quickly.

## 4 Conclusions

This paper provides a fault location method and a fault monitoring device for communication networks. The technical solution first grabs data from the error log of the communication network management system, and associates the reporting time with the error category. At the same time, the peak value of such errors is taken as the time interval to obtain the network access volume at each time point, and the access volume is compared with the error occurrence rate to obtain the key time node. The network environment within the time node has a greater potential impact on the occurrence of such errors. On this basis, this paper divides the process of the same type of error into several access behaviors, and uses the terminal equipment to repeatedly test each access behavior under the above network environment to collect the access behavior and its frequency of the same type of error, so as to determine the fault location. This paper

determines the communication network fault by trial and error, and fully restores the network environment, so that the fault location is more accurate. The application of this technology can achieve effective monitoring and positioning of communication network faults, and has a good use effect.

**Acknowledgement.** Outsourcing research project of key technologies of communication room visual control of digital twin technology of State Grid Hubei Xintong Company from 2022 to 2023.

## References

1. Mateska, A.K., Borozan, V., Krstevski, P., Taleski, R.: Controllable load operation in microgrids using control scheme based on gossip algorithm. *Appl. Energy* (2018)
2. Nasrallah, A., et al.: Ultra-Low Latency (ULL) Networks: The IEEE TSN and IETF DetNet Standards and Related 5G ULL Research. In: *IEEE Communications Surveys & Tutorials* (1) (2019)
3. Paolucci, F.: Network service chaining using segment routing in multi-layer networks. *J. Opt. Commun. Network.* (6) (2018)
4. Stojkoska, B.L.R., Trivodaliev, K.V.: A review of Internet of Things for smart home: challenges and solutions. *J. Clean. Prod.* (2017)
5. Suchita, D., Di, T.F., Mark, S.: Vigenère scores for malware detection. *J. Comput. Virol. Hacking Tech.* **14**(2), 157–165 (2018)
6. Coutinho, M., de Oliveira Albuquerque, R., Borges, F., Villalba, L.J.G., Kim, T.-H.: Learning perfectly secure cryptography to protect communications with adversarial neural cryptography. *Sensors.* (5) (2018)
7. Shah, S.A.R., Issac, B.: Performance comparison of intrusion detection systems and application of machine learning to Snort system. *Future Gen. Comput. Syst.* (2018)
8. Vijayakumar, P., Chang, V., Deborah, J., Balusamy, B., Shynu, P.G.: Computationally efficient privacy preserving anonymous mutual and batch authentication schemes for vehicular ad hoc networks. *Fut. Gen. Comput. Syst.* (2018)
9. Andreas, B., Anupam, C., Yan, L.K.: Wireless communication and security issues for cyber-physical systems and the Internet-of-Things. *Proc. IEEE.* (1) (2018)
10. Gholizadeh, N., Saadatfar, H., Hanafi, N.: K-DBSCAN: an improved DBSCAN algorithm for big data. *J. Supercomput.* **77**(6), 6214–6235 (2020). <https://doi.org/10.1007/s11227-020-03524-3>
11. Scitovski, R., Sabo, K.: A combination of k -means and DBSCAN algorithm for solving the multiple generalized circle detection problem. *Advances in Data Analysis and Classification* (prep) (2020)
12. Govender, P., Sivakumar, V.: Application of k -means and hierarchical clustering techniques for analysis of air pollution: a review (1980–2019). *Atmospheric Poll. Res.* (1) (2020)
13. Kang, J.S., et al.: Development of a systematic, self-consistent algorithm for the K-DEMO steady-state operation scenario. *Nuclear Fus.* (12) (2017)
14. Doostan, M., Chowdhury, B.H.: Power distribution system fault cause analysis by using association rule mining. *Electric Power Syst. Res.* (2017)
15. Wang, L., Chen, X., Wang, G., Hua, D.: Non-intrusive load monitoring algorithm based on features of V–I trajectory. *Electric Power Syst. Res.* (2018)
16. Saleem, Y., Crespi, N., Rehmani, M.H., Copeland, R.: Internet of Things-aided smart grid: technologies, architectures, applications, prototypes, and future research directions. *IEEE Access* (2019)



17. Aujla, G.S., Garg, S., Batra, S., Kumar, N., You, I., Sharma, V.: DROpS: a demand response optimization scheme in SDN-enabled smart energy ecosystem. *Inf. Sci.* (2018)
18. Parisot, S., et al.: Disease prediction using graph convolutional networks: Application to Autism Spectrum Disorder and Alzheimer's disease. *Med. Image Anal.* (2018)
19. Pilerood, A.E., Heydari, M., Mazdeh, M.M.: A two-stage greedy heuristic for a flowshop scheduling problem under time-of-use electricity tariffs. *South Afr. J. Ind. Eng.* (1) (2018)



# Digital Twin Simulation System for Computer Room Equipment

Lei Li, Zheng Zhou<sup>(✉)</sup>, Rongtao Liao, Yuxuan Ye, and Huanjun Hu

Information and Communication Branch of State Grid, Hubei Electric Power Company,  
Wuhan 430000, Hubei, China  
mufang19583psj@163.com

**Abstract.** This paper mainly studies a digital contracture analog device and a digital contracture analog system. The device is divided into driver installation, and the first expansion installation, the second expansion installation and the third expansion installation connected with the driver installation respectively; The driver is installed to connect the upper computer and the production device; The first expansion installation, the second expansion installation and the third expansion installation are used to connect the production devices; The driver installation is used to receive the message collection command of the upper computer, collect the relevant parameters of the production device according to the message collection command, and/or remotely control and drive the first expansion installation, the second expansion installation and the third expansion installation to collect the relevant parameters of the production device, and transmit the relevant parameters to the upper computer, and use the relevant parameters to conduct system simulation of the production device. The above devices can collect a large number of relevant parameters of production equipment, and conduct simulation modeling according to relevant technical parameters through digital contracture technology. The characteristics of production equipment can also be mastered through simulation results, and its information processing speed is fast.

**Keywords:** Power · Machine Room Equipment · Digital Twins · Simulation

## 1 Introduction

Due to the continuous development of scientific and technological innovation and the rising quality of people's life, electronic equipment (such as instruments, equipment, etc.) has become a basic part of people's daily life [1]. With the reduction of the life cycle of electronic products and the improvement of the customization depth of product design, scientific and accurate understanding of the basic characteristics of product design and the actual application situation will help customers reduce economic losses [2–7]. It is particularly important to directly reflect the actual application situation of customers on the product design side, so as to improve the product design efficiently.[8] To master the design characteristics of products, the most traditional method is to collect the relevant parameters of product design through the data collection system of SCM,

and then transfer the parameters to the upper computer and other devices to analyze the parameters (such as simulation, modeling, etc.), and master the relevant situation of product design in this way [9–12]. However, the data collected in this way is relatively simple, while the data analysis steps are very complex.

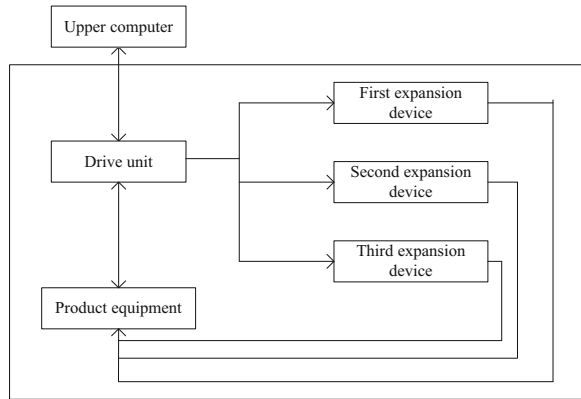
With the growing development of the data society, today's digital technology is also constantly shaping and changing the enterprise [13–17]. In the future, all companies will become fully digital enterprises, which not only requires the company to develop products with fully digital characteristics, but also refers to the use of digital technology to affect the entire company's product design, research and development, production and after-sales service processes, as well as the use of digital technology to connect the internal and external environment of the entire company [18]. The so-called information generation technology refers to making full use of physical simulation, sensor product update, product use history and other information to conduct information simulation in multiple fields, multiple physical quantities, multiple levels, and multiple probabilities, reflecting information on the entire news space, and expressing the life cycle state of the corresponding physical information. It can be seen that people can easily understand enterprise products through digital birth technology. Based on this, it is necessary to propose a digital contracture simulation system and a digital contracture simulation system to solve the technical problem that the information collected in the previous information collection, analysis and delivery system based on microcomputer and upper computer is relatively simple and very complex to process.

## 2 Digital Contracture Simulation System

Figure 1 is the schematic diagram of digital contracture simulation settings in this paper. A digital spawning simulation setting includes: a driver setting, a first extension setting, a second extension setting and a third extension setting connected to the driver setting respectively; [19] The driving device is used to connect the upper computer and the production device; The first expansion setting, the second expansion setting and the third expansion setting are used to connect the production device; The drive device is used to receive the information acquisition command of the upper computer, collect the relevant parameters of the production device according to the information acquisition command, and/or remotely control and promote the first expansion setting, the second expansion setting and the third expansion setting to collect the relevant parameters of the production device, and transmit the relevant parameters to the upper computer. The relevant parameters can be used to simulate and model the production device.

Digital contracture simulation equipment is divided into driving equipment, the first expanding equipment, the second expanding equipment and the third expanding equipment. On the one hand, the driving device is connected with the first expanding equipment, the second expanding equipment and the third expanding equipment in turn; [20] On the one hand, the driving device has the ability to start, collect and send commands at the same time; On the other hand, the device is pushed to connect with the upper computer and receive the monitoring command of the upper computer. The monitoring command includes the message collection command, the message termination collection command, etc. The message collection command includes the message collection mode,

requirements, content types, etc.; The message termination collection instruction command includes the message termination method, request type, etc.; On the one hand, the driving device can also manage the relevant parameters of the first extension device, the second extension device or the third extension device through the monitoring instruction command; On the other hand, if the driving device is connected, the driving device can collect corresponding data from the product equipment by using its own measurement capability through control commands; When the corresponding data is collected, the corresponding data will be transmitted to the enterprise's content locator.



**Fig. 1.** Digital contracture simulation system

Among them, the relevant technical parameters related to production equipment refer to all relevant technical parameters related to production equipment, including the appearance parameters and relevant technical parameters of production equipment. The appearance parameters include type, size, width, material, etc., while the relevant technical parameters include performance parameters, working conditions, etc., such as temperature, pressure, current, voltage, line grounding, etc.; Relevant parameters are defined for specific production devices, and the types of specific parameters for each production device are different.

The first extension device, the second extension device and the third extension device are all measuring and transmitting devices of product data, whose main function is to measure or control the product equipment, and collect or send relevant data, that is, the function of product data input and output. The data acquisition effect can be greatly improved by parallel work of multiple expansion units such as the first expansion unit, the second expansion unit and the third expansion unit.

The data generation simulation device given in this paper includes driver installation, the first expansion installation, the second expansion installation and the third expansion installation. The driver installation can link the first expansion installation, the second expansion installation and the second expansion installation in turn, and the driver installation can be used to link the production devices of all upper computers in turn, and the first expansion installation, the second expansion installation The third expansion installation is also used to link production devices in turn, and the driver

system directly accepts the data acquisition command of the upper computer, collects the truth data of all production devices, and/or manages and guides the corresponding data of all production devices of the first expansion installation, the second expansion installation and the third expansion installation; Then the relevant technical parameters can be transmitted to the upper computer, which can also conduct simulation for the relevant technical parameters. Through these devices, the relevant technical parameters of most production equipment can be collected (that is, the performance, status and other parameters of production equipment), and then the relevant technical parameters can be simulated through data generation technology, Finally, the characteristic technical parameters of all production equipment can be mastered through the simulation results; On the one hand, it can also collect various types of information at the same time, and the information processing speed is fast.

### 3 Structure of Expansion Device

As shown in Fig. 2, the driving device is divided into the first main control, the port circuit connected to the first console in turn, the power supply management circuit, the entrance integration circuit and the exit integration circuit; The first main control uses the port integration circuit to connect the first expansion device, the second expansion device and the third expansion device in turn; The port integration circuit is used to connect the upper computer; The inlet integrated circuit, the outlet integrated circuit, the first expansion device, the second expansion device and the third expansion device are used to connect the product equipment in turn; The power supply management circuit is sequentially connected with the first expansion device, the second expansion device and

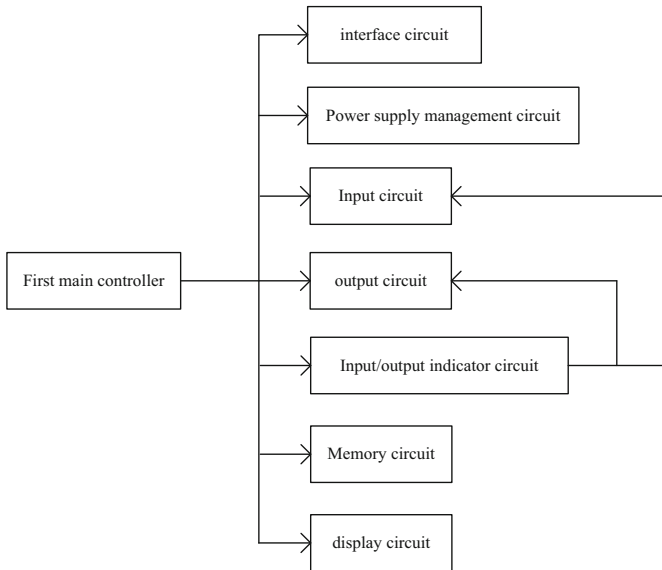
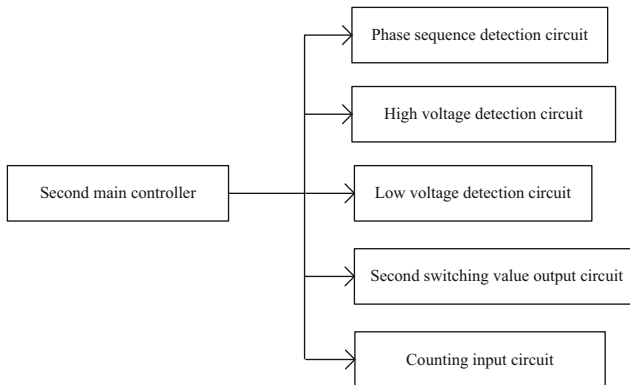


Fig. 2. Drive structure

the third expansion device; The supply power management circuit is used to convert the supply current into the corresponding target current, and supply power to the product connection integrated circuit, the inlet integrated circuit, the outlet integrated circuit, the first main control device, the first expansion device, the second expansion device, and the third expansion device.

As shown in Fig. 3, the first interface device is divided into the second main control unit, phase sequence detection circuit, high voltage inspection circuit, low temperature inspection circuit, the second switching value outlet integrated circuit and metering signal input integrated circuit; The phase sequence detection circuit, high voltage detection circuit, low temperature detection circuit, the second switching value outlet integrated circuit and the metering signal input integrated circuit are successively connected to the first main control unit through the second main control unit, while the phase sequence detection circuit, high voltage detection circuit, low temperature detection circuit, the second switching value outlet integrated circuit and the metering signal input integrated circuit are successively connected to the second main control unit interface, Access production equipment in turn; The phase sequence detection circuit is divided into multiple circuits to test the forward and reverse operation of the three-phase motor of the production device, and use it to test the high-voltage exchange information; The high-voltage inspection circuit is divided into multi-channel high-voltage test equipment to test and exchange information; The low-voltage measuring circuit includes multiple low-temperature measuring ports for measuring AC and DC signals; The measuring output circuit of the second system comprises a measuring port of a multiplexer switch for measuring all flows; The measuring input circuit of the system includes a high-speed measuring port, which is used to measure the high-speed pulse signal of the generator encoder on the production line and record the working speed of the motor at the same time.



**Fig. 3.** Structure of the first expansion device

As shown in Fig. 4, the second expansion device includes the third main controller, I/O scanning line and discharge line; The I/O scanning line uses the third master controller to link the second master control, and the I/O scanning line links the product equipment; The I/O scanning circuit contains multiple test ports. Each test port is equipped with a double disconnector, and the discharge circuit is set between the double disconnectors of the 22 test ports; The I/O scanning circuit is used to check the opening and closing status of each line connection node in the power line of the production device, and judge the wiring status of each line through the opening and closing status; The discharge circuit is used to release the storage equipment of the production device when cutting off the high-voltage current.

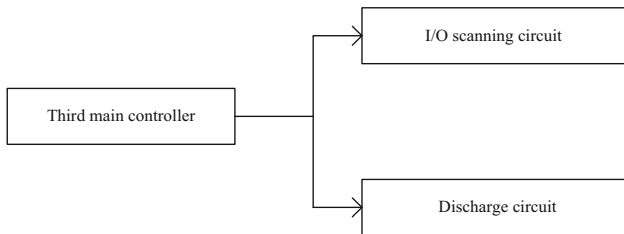
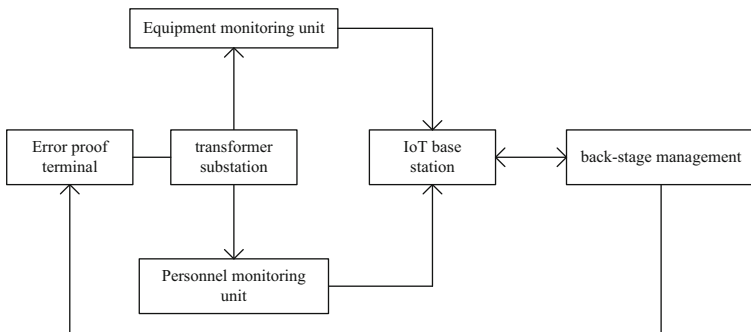


Fig. 4. Structure of the second expansion device

#### 4 Anti Misoperation System of Digital Nusheng Substation

As shown in Fig. 5, this paper provides an anti misoperation system for digital twin substation. The system includes: equipment monitoring unit, which identifies and collects data on the primary and secondary equipment in the substation and the environment where the substation is located. Multiple monitoring points are set on the primary and secondary equipment. The equipment monitoring unit is used for real-time collection of electric equipment data, environmental data, live data and image data of multiple monitoring points; The personnel monitoring unit collects the movement path, identity recognition and vital sign related data of personnel in the substation in real time; The IoT base station is used for receiving and transmitting real-time data streams, and its communication ground coupling connects the equipment monitoring unit and the personnel monitoring unit to feed back the collected data streams; The management background connects and processes the collected data flow to view, store and play back the operation status information of the substation and the status information of the personnel, judge whether the electrical equipment of the substation is abnormal, establish a 3D model based on the data flow, divide the working area and the non working area, generate the operation ticket information, and the management background interacts with the personnel monitoring unit in a two-way manner; The management background connects the IoT base station through the server two-way communication. The server can connect multiple IoT base stations, and the server is in the response mode for a long time; Anti error terminal, coupling and connecting the primary equipment and secondary equipment in the substation to perform opening and closing operations on the

primary equipment and secondary equipment, and anti error terminal communication and connection management background.



**Fig. 5.** Working principle diagram of anti misoperation system of digital Nusheng substation

## 5 Conclusions

The digital contracture simulation system uses the digital contracture simulation equipment mentioned above, thus producing a beneficial effect matched with the digital contracture simulation equipment; In this way, a large number of parameters related to the production device (i.e., parameters related to product equipment performance, operation status, etc.) can be collected through the digital contracture simulation system, and the simulation can be conducted according to relevant technical parameters through the digital contracture technology, so that the characteristics of the production device can be mastered through the simulation results; On the one hand, a large amount of data of various types can be collected at the same time, and the information processing speed is fast.

**Acknowledgement.** Outsourcing research project of key technologies of communication room visual control of digital twin technology of State Grid Hubei Xintong Company from 2022 to 2023.

## References

1. Pablo, C., Albert, B., Gustavo, de.V., Xavier, C.P., Arturo, A.: Network slicing for guaranteed rate services: admission control and resource allocation games. *IEEE Trans. Wirel. Commun.* (10) (2018)
2. Dorsch, N., Kurtz, F., Wietfeld, C.: On the economic benefits of software-defined networking and network slicing for smart grid communications. *NETNOMICS: Econ. Res. Electron. Network.* **19**(1–2), 1–30 (2018). <https://doi.org/10.1007/s11066-018-9124-3>
3. A. Ayaz, Rehmani Mubashir Husain, Tembine Hamidou, Mohammed Osama A., Jamalipour Abbas. *IEEE Access Special Section Editorial: Optimization for Emerging Wireless Networks: IoT, 5G, and Smart Grid Communication Networks.* *IEEE Access.* 2017



4. Aujla, G.S., Garg, S., Batra, S., Kumar, N., You, I., Sharma, V.: DROpS: a demand response optimization scheme in SDN-enabled smart energy ecosystem. *Inf. Sci.* (2018)
5. Suchita, D., Di, T.F., Mark, S.: Vigenère scores for malware detection. *J. Comput. Virol. Hacking Techniques* (2) (2018)
6. Coutinho, M., de Oliveira Albuquerque, R., Borges, F., Villalba, L.J.G., Kim, T.-H.: Learning perfectly secure cryptography to protect communications with adversarial neural cryptography. *Sensors* (5) (2018)
7. Shah, B.I.: Performance comparison of intrusion detection systems and application of machine learning to Snort system. *Future Gener. Comput. Syst.* (2018)
8. Vijayakumar, P., Chang, V., Deborah, L.J., Balusamy, B., Shynu, P.G.: Computationally efficient privacy preserving anonymous mutual and batch authentication schemes for vehicular ad hoc networks. *Future Gener. Comput. Syst.* 2018
9. Andreas, B., Anupam, C., Yan, L.K.: Wireless communication and security issues for cyber-physical systems and the Internet-of-Things. *Proc. IEEE.* (1) (2018)
10. Gholizadeh, N., Saadatfar, H., Hanafi, N.: K-DBSCAN: An improved DBSCAN algorithm for big data. *J. Supercomput.* 77(6), 6214–6235 (2020). <https://doi.org/10.1007/s11227-020-03524-3>
11. Scitovski, R., Sabo, K.: A combination of k -means and DBSCAN algorithm for solving the multiple generalized circle detection problems. *Advances in Data Analysis and Classification*. 2020 (prep)
12. Govender, P., Sivakumar, V.: Application of k -means and hierarchical clustering techniques for analysis of air pollution: a review (1980–2019). *Atmosph. Poll. Res.* (1) (2020)
13. Kang, J.S., et al.: Development of a systematic, self-consistent algorithm for the K-DEMO steady-state operation scenario. *Nuclear Fusion* (12) (2017)
14. Doostan, M., Chowdhury, B.H.: Power distribution system fault cause analysis by using association rule mining. *Electric Power Syst. Res.* (2017)
15. Wang, L., Chen, X., Wang, G., Hua, D.: Non-intrusive load monitoring algorithm based on features of V–I trajectory. *Electric Power Syst. Res.* (2018)
16. Saleem, Y., Crespi, N., Rehmani, M.H., Copeland, R.: Internet of Things-aided smart grid: technologies, architectures, applications, prototypes, and future research directions. *IEEE Access* (2019)
17. Domingo-Ferrer, J., Farràs, O., Ribes-González, J., Sánchez, D.: Privacy-preserving cloud computing on sensitive data: a survey of methods, products and challenges. *Comput. Commun.* (2019)
18. Koroniotis, N., Moustafa, N., Sitnikova, E.: Forensics and deep learning mechanisms for botnets in internet of things: a survey of challenges and solutions. *IEEE Access.* (2019)
19. Ylmaz, E.N., Gnen, S.: Attack detection/prevention system against cyber attack in industrial control systems. *Comput. Secur.* (2018)
20. Al-rimy, B.A.S., Maarof, M.A., Shaid, S.Z.M.: Ransomware threat success factors, taxonomy, and countermeasures: A survey and research directions. *Comput. Secur.* (2018)



# Fault Detection Methods for Communication Equipment

Hongyu Ke<sup>(✉)</sup>, Xi Zhang, Jing Dai, Chenxi Dong, and Lei Li

Information and Communication Branch of State Grid, Hubei Electric Power Company,  
Wuhan 430000, Hubei, China  
taiyan82711ykz@163.com

**Abstract.** This paper studies a communication equipment fault detection system, including at least one sensing object, at least one sensing component connected to the sensing object, and a unified fault detection device connected to the sensing component; The fault detection device includes an object information processing module and an object fault processing module connected with the sensing component and the object information processing module. This paper also studies a communication device fault detection method, including calling the attribute information of the pre stored sensing object to determine the type of the sensing object; determine the fault condition of the sensing object and generate the corresponding fault event. Through the system and method in this paper, the repeated development of fault detection device is avoided, and the cost of development and maintenance is reduced; it is helpful to improve the stability of the system; the transplantation and reuse of fault detection devices in different communication devices are realized.

**Keywords:** Communication · Fault Monitoring · Maintenance System

## 1 Introduction

As the downtime of communication equipment in the network will lead to the paralysis of local or even the entire network, the availability requirements for communication equipment are generally high; [1] Especially for the core communication equipment in the network, it is usually required to provide five nines (99.999%) availability, that is, the allowed downtime within a year cannot exceed five minutes. [2–5] To ensure such high availability, the premise is to build a perfect fault detection system and fault detection method, that is, if the equipment has a fault, it must be able to detect it as soon as possible; [6] After the fault is detected, the fault must be located as soon as possible; [7–12] After the fault is located, it must be repaired as soon as possible. Only in this way can the requirements for the availability of communication equipment be met. [13] The purpose of this paper is to overcome the defects in the existing technology, provide a communication equipment fault detection system and a concentric equipment fault detection method. [14–17] By using a unified fault detection device for different sensing objects, we can avoid repeated development, reduce the cost of development and maintenance, and improve the stability of the system.

## 2 Existing Fault Detection System

The existing communication equipment fault detection system is mainly composed of three parts: sensing object, sensing component and fault detection module, as shown in Fig. 1, including sensing object, sensing components connected to it, and their respective fault detection modules. Perceived objects mainly refer to the objects to be detected, such as temperature, voltage, fan, etc.; [18] The function of the sensing component is mainly to collect the state of these sensing objects; The fault detection module mainly coordinates the work of each part, including reading the value of the sensing component, necessary processing of the information obtained by the sensing component, and controlling the sensing component. Figure 2 is a schematic diagram of an actual fault detection system.

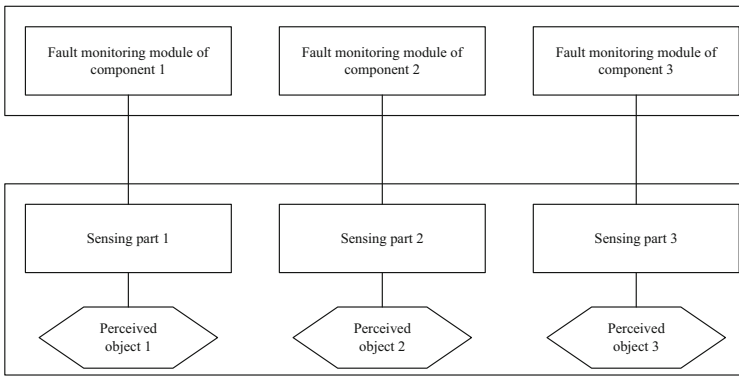
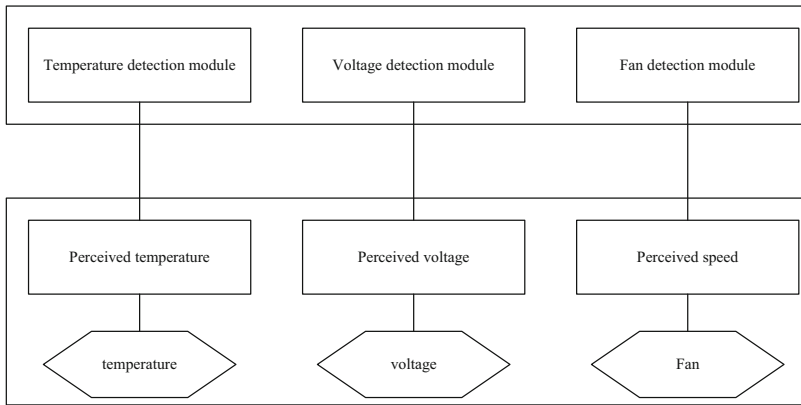


Fig. 1. Fault detection system of communication equipment in prior art

The flow of the existing communication equipment fault detection method is as follows: the fault detection module periodically (the timing time can be determined according to the specific sensing object) obtains the value of the sensing object through the sensing component, and then judges whether the sensing object is faulty by the value of the sensing object obtained. If it fails, it sends an alarm event. The first level threshold judgment method is mainly used to judge whether the sensing object is faulty. The specific steps are: acquiring the value of the sensing object through the sensing component; Compare the value of the acquired sensing object with the preset threshold value (the threshold value can be modified later, and the threshold can be divided into upper threshold and lower threshold). If the value of the sensing object is greater than the upper threshold or less than the lower threshold, it is considered that the state of the sensing object is faulty; If it is not, the state of the perception object is normal; Judge whether the last detected state is consistent with the current detected state. If not, a state change event will be generated to change the current state. Otherwise, the current state will not be changed.



**Fig. 2.** An actual fault detection system in the prior art

In general, the existing communication equipment fault detection systems and methods have the following defects:

(A). For different sensing objects, different fault detection modules are used for detection; Its disadvantage is that the fault detection function is distributed on each fault detection module, so different perceptions are required.

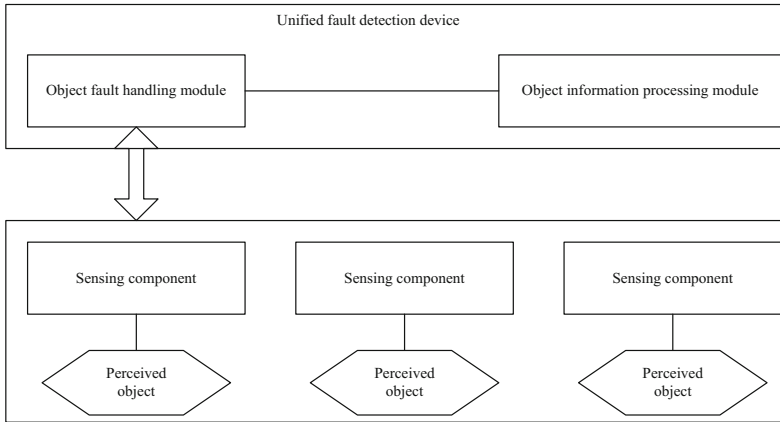
The object compiles the corresponding fault detection program, which leads to a lot of repeated work, and the cost of developing a complete set of test software remains high; It is difficult to unify the external interfaces provided by each program, which hinders the transplantation and sharing of different programs; Due to the different implementation of the fault detection program, the maintenance cost of the later software is increased; At the same time, there are system compatibility problems, which is not conducive to the stability of the fault detection system;

(B). The detection software needs to solidify the attribute description of the sensing object in the communication device; Its defects are: because the detection software is designed for communication equipment, the attributes of the sensing object are solidified in the detection software at the same time, resulting in strong limitations of the fault detection system, which can only detect the sensing object of the communication equipment at the beginning of the design, and cannot be reused between different communication devices;

(C). The first level threshold judgment method is adopted, which has the following disadvantages: the first level threshold judgment method can only judge whether the state of the sensing object is in the fault state or normal state, but can not distinguish the severity of the actual fault state (such as fatal/serious/minor), and can not handle the corresponding fault according to the different severity; The first level threshold judgment method is to immediately judge the state of the sensing object according to the obtained value. However, due to the reason of the sensing object, such as the state of some lines, there needs to be a statistical cycle. In a statistical cycle, if the number of faults detected meets the requirements, the line is considered as a state fault; Therefore, the traditional first level threshold judgment method is likely to lead to miscalculation.

To sum up, the existing communication equipment fault detection system and communication equipment detection methods cannot meet the requirements of portability, stability and accuracy, and there are also shortcomings of high development and maintenance costs and insufficient accuracy.

### 3 Two Kinds of Communication Equipment Fault Detection Systems



**Fig. 3.** Communication equipment fault detection system (a)

As shown in Fig. 3, it includes a sensing object, sensing components connected to the sensing object, and a unified fault detection device connected to the sensing component. The fault detection device is used to replace the fault detection module in the prior art, including:

The object information processing module has the function of classifying and processing the perceptual object information. Specifically, the object information processing module can extract the attribute information of the perceptual object.

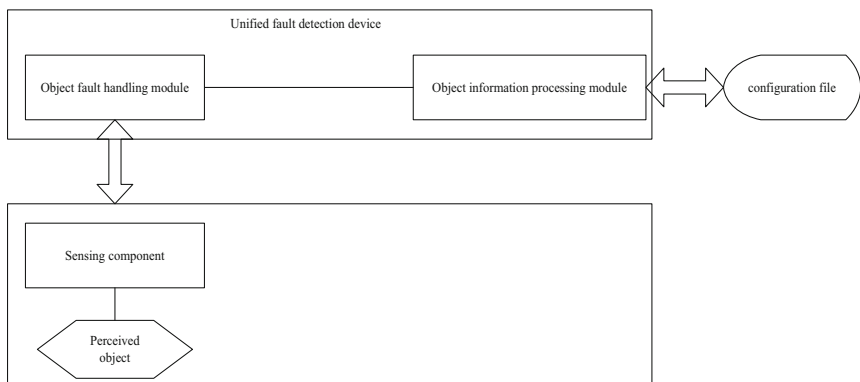
The object information processing module can obtain attribute information when the fault detection device is initialized, and also can obtain attribute information during the detection of the fault detection system. In this paper, the object information processing module is provided with a configuration file module, which is used to store the configuration files of attribute information of each sensing object in advance. That is, in this paper, the attribute information of the sensing object, including type information and parameter settings, is preset in the configuration file and stored in the object information processing module. When the fault detection device is initialized or the fault detection system is running, the object information processing module extracts the associated information from the configuration file. For example, both CPU voltage and board voltage are associated with periodic statistical sensing objects for processing.

The attribute information stored in the configuration file in this paper includes: type information, such as perceptual object name, perceptual object type, and perceptual

object type identifier; The parameter settings include the ID of the sensing object, the threshold value of the sensing object, the enabling mark of the sensing object, the physical location of the sensing object, and the data type of the sensing object. The format of the configuration file can be described in XML or in text. Set all the information of the items to be tested in the configuration file before the communication device is detected, and then it can be processed uniformly according to different modes.

The object fault processing module is connected with the sensing component to obtain the state information of the sensing object through the sensing component; Connected with the object information processing module, it is used to call the sensing object attribute information to determine the type of the sensing object; It is also used to determine the fault condition of the sensing object according to its type and state information, that is, whether its state has changed, including switching between the fault state and the normal state, switching between discrete states, etc. If the state of the sensing object changes, a fault event will be generated.

Since the object fault processing module is responsible for the generation of fault events, it is equipped with counters, timers and other modules commonly used in the field for equipment fault detection. At the same time, the sensing object is one or any combination of periodic statistical sensing object, continuous statistical sensing object, multi-level threshold sensing object, discrete sensing object, watermark sensing object and counter sensing object. For example, in this paper, it can be divided into CPU temperature, CPU voltage, board voltage, fan, etc. The sensing component is connected to the sensing object to sense the change of CPU temperature, CPU voltage, or fan. It can be seen that through the system in this paper, it is unnecessary to repeatedly edit and develop fault detection devices for different components, reducing the development and maintenance costs; At the same time, there is not too much detection software running, which is also conducive to improving the stability of the system.



**Fig. 4.** Communication equipment fault detection system (b)

As shown in Fig. 4, it includes a perception object, a perception component connected to the perception object, a unified fault detection device connected to the perception component, and a configuration file module connected to the fault detection device. The

fault detection device includes an object information processing module and an object fault processing module. The configuration file module is connected with the object status information processing module, and is used to store the configuration file of the sensing object attribute information in advance. The configuration file is used to describe the sensing object type information and parameter settings, including the sensing object name, sensing object type, sensing object type identification, sensing object ID, sensing object threshold value The enabling tag of the sensing object, the physical location of the sensing object, and the data type of the sensing object. The format of the configuration file can be described using XML.

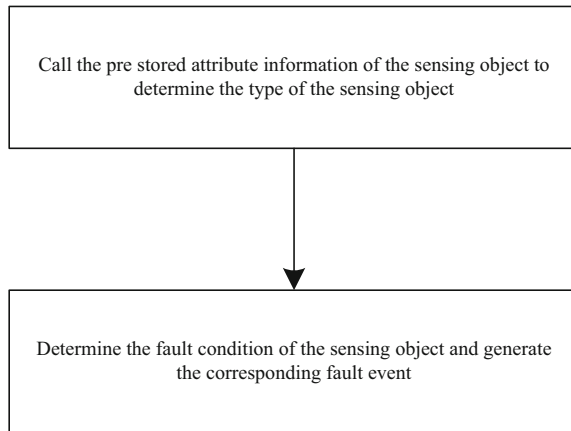
Among them, the configuration file module can be updated at any time. As long as the information of all items to be tested is updated before changing the communication equipment for detection, it can be uniformly processed according to different modes. The configuration file module can be set in the network configuration server or in the device to be tested, regardless of the specific physical address, as long as it can communicate with the object information processing module. It can be seen that the system provided in this paper can realize the transplantation and reuse of fault detection devices in different communication devices.

## 4 Detection Method

This paper also provides a communication device fault detection method, its flow chart is shown in Fig. 5, including the following steps: call the pre stored attribute information of the sensing object, and determine the type of the sensing object according to the attribute information; According to the type and state information of the called sensing object, determine the fault condition of the sensing object, and generate the corresponding fault event according to the fault condition. Before step 1, it also includes: saving the attribute information of the sensing object, and saving the attribute information of the sensing object in the configuration file. The attribute information of the sensing object can be saved in the configuration file in XML format.

After establishing the perception object attribute information and saving it in the configuration file, during the fault detection process, the configuration file is called to obtain the perception object attribute information, and according to the attribute information of the perception object, the type of the perception object is determined to be a periodic statistical perception object. As a periodic statistical sensing object, it has the following technical parameters as the basis for judging the fault state or normal state: detection cycle: the time interval for obtaining the current state of the tested sensing object; Statistical cycle: the time interval for judging the statistical status of the perceived object under test. The statistical cycle is a multiple of the detection cycle. For example, one statistical cycle is equal to 10 detection cycles.

Statistical counter: indicates how many detection cycles have been completed. Status counter: indicates the count of the status different from the statistical status of the sensing object obtained during the detection cycle. Generation threshold: In a statistical cycle, if the number of failures of the sensing object is equal to or greater than this value, the statistical status of the sensing object is considered as failure.



**Fig. 5.** Flow chart of communication equipment fault detection method

After determining that the perception object is a continuous statistical perception object, the processing flow is shown in Fig. 6:

Step (A). When the detection cycle arrives, obtain the current state of the sensing object;

Step (B). Judge the statistical state of the sensing object. If it is in the fault state, execute step (C), otherwise execute step (F);

Step (C). Judge the current state of the sensing object. If it is a fault state, execute step (I), otherwise execute step (D);

Step (D). Count the accumulated state counter;

Step (E) Judge whether the count of the state counter is greater than or equal to the set recovery threshold value. If yes, a fault recovery event will be generated. Set the statistical state of the sensing object to the normal state and execute step (I), otherwise directly execute step (J);

Step (F). Judge the current state of the sensing object. If it is normal, execute step (I), otherwise execute step (G);

Step (G). Count the accumulated state counter;

Step (H) Judge whether the count of the state counter is greater than or equal to the set recovery threshold. If yes, a fault event will be generated. Set the statistical state of the sensing object to the fault state, and execute step (I); Otherwise, directly execute step (J);

Step (I). Clear the status counter;

Step (J). End the test.



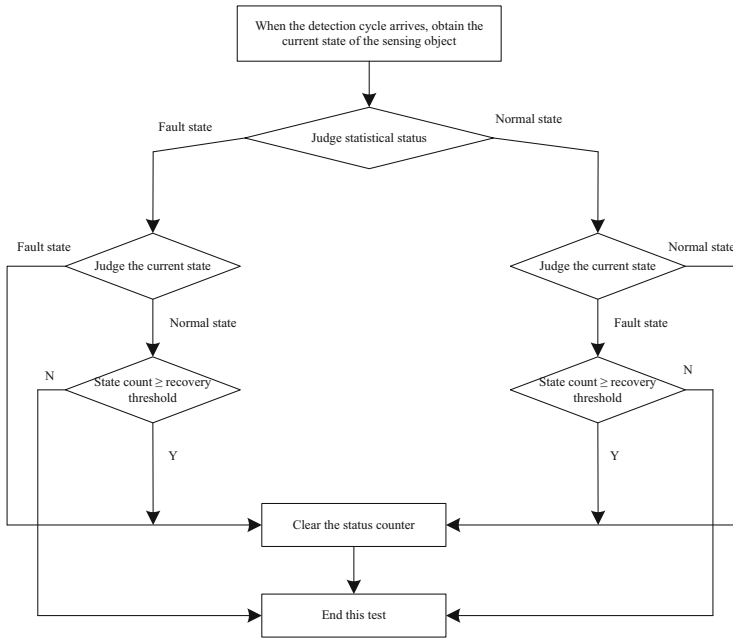


Fig. 6. Algorithm Flow Chart

### 5 Conclusions

It can be seen from the above technical scheme that this paper uses a unified fault detection device to detect different sensing objects by establishing attribute information on the sensing objects in a unified way, which has the following beneficial effects: avoid the repeated development of corresponding fault detection devices for different sensing objects, and reduce the cost of development and maintenance; Adopting a unified fault detection device is conducive to improving the stability of the system; The transplantation and reuse of fault detection devices in different communication devices are realized.

**Acknowledgement.** Outsourcing research project of key technologies of communication room visual control of digital twin technology of State Grid Hubei Xintong Company from 2022 to 2023.

### References

1. Singh, M.: Protection coordination in distribution systems with and without distributed energy resources- a review. *Protection Control Modern Power Syst.* 2017 (1)
2. Yoon, S., Lee, Y.-J., Jung, H.-J.: A comprehensive framework for seismic risk assessment of urban water transmission networks. *International J. Disaster Risk Reduct.* (2018)
3. Berrueta, A., Urtasun, A., Ursúa, A., Sanchis P.: A comprehensive model for lithium-ion batteries: From the physical principles to an electrical model. *Energy* (2018)

4. Hooshyar, D., Pedaste, M., Leijen, K.S., Bardone, E., Wang, M.: Open learner models in supporting self-regulated learning in higher education: a systematic literature review. *Comput. Educ.* (2020) (prep)
5. Motalleb, M., Siano, P., Ghorbani, R.: Networked stackelberg competition in a demand response market. *Appl. Energy* (2019)
6. Andoni, M., et al.: Blockchain technology in the energy sector: a systematic review of challenges and opportunities. *Renew. Sustain. Energy Rev.* (2019)
7. Ryu, U., Wang, J., Kim, T., Kwak, S., Juhyok, U.: Construction of traffic state vector using mutual information for short-term traffic flow prediction. *Transp. Res. Part C.* (2018)
8. Noor, S., Yang, W., Guo, M., van Dam, K.H., X. Wang. Energy Demand Side Management within micro-grid networks enhanced by blockchain. *Appl. Energy* (2018)
9. Mengelkamp, E., Grttner, J., Rock, K., Kessler, S., Orsini, L., Weinhardt, C.: Designing microgrid energy markets. *Appl. Energy* (2018)
10. Gholizadeh, N., Saadatfar, H., Hanafi, N.: K-DBSCAN: an improved DBSCAN algorithm for big data. *J. Supercomput.* 2020 (prep)
11. Scitovski, R., Sabo, K.: A combination of k -means and DBSCAN algorithm for solving the multiple generalized circle detection problem. In: *Advances in Data Analysis and Classification* (2020) (prep)
12. Govender, P., Sivakumar, V.: Application of k -means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019). *Atmospheric Poll. Res.* (1) (2020)
13. Kang, J.S., et al.: Development of a systematic, self-consistent algorithm for the K-DEMO steady-state operation scenario. *Nuclear Fusion.* (12) (2017)
14. Doostan, M., Chowdhury, B.H.: Power distribution system fault cause analysis by using association rule mining. *Electric Power Syst. Res.* (2017)
15. Wang, L., Chen, X., Wang, G., Hua, D.: Non-intrusive load monitoring algorithm based on features of V–I trajectory. *Electric Power Syst. Res.* (2018)
16. Saleem, Y., Crespi, N., Rehmani, M.H., Copeland, R.: Internet of Things-aided smart grid: technologies, architectures, applications, prototypes, and future research directions. *IEEE Access* (2019)
17. Pablo, C., Albert, B., Gustavo, deV., Xavier, C.P., Arturo, A.: Network slicing for guaranteed rate services: admission control and resource allocation games. *IEEE Trans. Wirel. Commun.* (10) (2018)
18. Dorsch, N., Kurtz, F., Wietfeld, C.: On the economic benefits of software-defined networking and network slicing for smart grid communications. *NETNOMICS: Econom. Res. Electron. Networking* **19**(1–2), 1–30 (2018). <https://doi.org/10.1007/s11066-018-9124-3>



# Design and Implementation of Teaching Assistant System Based on Intelligent Recommendation Algorithm

Jie Ding<sup>(✉)</sup>

Jilin Agricultural Science and Technology University, Jilin, Jilin, China  
1553859608@qq.com

**Abstract.** Language teaching is different from the training process of applied majors, which requires a lot of listening and speaking exercises to consolidate the knowledge learned. In order to improve the effect of online English teaching, with the help of the development of computer simulation technology and multimedia technology, based on intelligent recommendation algorithm and combined with virtual reality technology, a three-dimensional model of modular simulation teaching scene (MPS) was established. Unity3D engine was used to complete the development of eight training scenes in MPS simulation teaching platform. The test results show that the developed virtual teaching platform can meet the needs of English teaching and runs smoothly.

**Keywords:** Virtual Reality · Intelligent Recommendation Algorithm · English Teaching · Auxiliary Platform

## 1 Introduction

In the classroom teaching, teachers always play a vital role and face many challenges. The first is how to get immediate and accurate feedback from students. Teachers can rely on a variety of information, including classroom questioning, discussion, testing, students answering questions, and assessing students' learning performance through their expressions, gestures and body movements. However, if these information can not be effectively processed, it will reflect the new pattern is scattered and rough, and cannot truly reflect the learning state and effect of students. Another challenge is how to accurately document and model student learning, which is critical for assessing student status and personalizing instruction. In order to establish an accurate model of individual learning time series, it is necessary to ensure that the whole teaching process is recorded on a purely quantitative and calculable basis [1].

It is very necessary to track students' learning of each knowledge point accurately and effectively through student model. As the most commonly used student model, knowledge point tracking model models students' various learning behaviors based on their historical learning track, so as to continuously track students' mastery of knowledge points at various points in time. However, most of the existing student models are modeled

based on flashcard learning mode [2]. However, in most common teaching processes, a series of knowledge points are indirectly related to others, which is inconsistent with the flashcard situation [3].

In order to facilitate the classroom or after the classroom teaching interaction and better use of the built student model, a changes in teaching methods and teaching AIDS are very important. These auxiliary systems achieve effective support for colorful classrooms by integrating advanced technologies in fields such as computer systems, Internet, instant messaging, augmented reality, data transmission, artificial intelligence and data base mining. It improves student participation and interaction in class or after class. It also provides an easy way for teachers to understand students and help students conduct self-assessment. However, the current teaching assistance system only focuses on the specific application of a certain point, such as check-in and homework, and fails to focus on the whole process of teaching, that is, from before class to class and after class.

## 2 Knowledge Tracking Model

The essence of sequential tracking of knowledge points is to take students' learning behaviors and learning as basic data to predict students' learning behaviors at any time, so as to achieve the purpose of predicting students' learning effects and help teachers adjust teaching methods and teaching plans reasonably. The existing knowledge sequential tracking models are basically divided into three forms: knowledge tracking based on probability graph, matrix decomposition and deep learning. Knowledge Tracing models based on probability graphs include Bayesian Knowledge Tracing (BKT) and Fuzzy Cognitive Diagnosis Framework (Fuzzy CDF) [4, 5].

BKT model is a knowledge tracking model with relatively high precision. It designs the state of students' potential knowledge mastery as a binary variable, and expresses students' understanding and mastery of relevant knowledge points through different variables. Based on real-time feedback from users, BKT uses Hidden Markov Model (HMM) to update these binary variables. Fuzzy CDF uses the theory of fuzzy set to quantify the students' knowledge proficiency. It assumes the different knowledge composition of the students' needs to answer objective questions and subjective questions, and calculates the students' knowledge points by combining the results of the students' answers [6].

Knowledge Tracing models based on matrix decomposition include Knowledge Proficiency Tracing (KPT) model, etc. KPT is an interpretive probabilistic knowledge proficiency tracking model based on probabilistic matrix decomposition, which quantifies the multidimensional features of each knowledge point into a feature matrix, and calculates the complete feedback matrix of students according to the partial feedback of students [7]. At the same time, KPT combined with the human forgetting curve to take students' forgetting factors into account to update the student model over time.

Knowledge sequential tracking model based on deep learning integrates deep learning algorithms to track knowledge, and uses LSTM (Long-term-term Memory) network neural tracking to track students' understanding and mastery of knowledge points at different stages [8]. The input of LSTM neural network is the binary scalar relationship between learning effect and student input, and the output is the understanding score and

mastery value of students' related knowledge points. Due to the progressive nature of LSTM, DKT can incorporate student input feedback into the model. The k-enhancement of KT means the k-enhancement of the learning model (KSGKT) and the enhancement of KT means the knowledge structure communicator enhancement of the learning model (KSGKT). Combine these unique characteristics [9].

### 3 Classroom Intelligent Recommendation Algorithm

The ability to learn and retain large amounts of new knowledge is an essential part of human education. In 1885, Ebbinghaus innovatively proposed the forgetting curve for the emergence of human memory theory. The study is the first to identify two factors that affect recall of a knowledge or other concept: reinforcement and delay. On this basis, Reddy et al. proposed a change algorithm based on forgetting curve to describe students' learning. Settle and other schools use psychological theory and artificial intelligence technology to design the semi-decay cycle regression model, which has been widely used to simulate the memory process of students. Raffert et al. described the learning process of students as an observable Markov decision process, and used the solution method based on particle movement to approximate the optimal learning strategy. Elshani et al. regarded the information related to students, courses and grades as ordered genes and used improved genetic algorithms to describe the optimal learning path [10]. Niknam and other scholars adopt the learning path recommendation system, which designs the best learning path for learners with the help of the bionic ant colony optimization algorithm. Reddy and other scholars adopted the virtual biological evolution method, which does not need to simulate students, and introduced the former, the middle and the latter three methods to deal with the problem of interval repetition. These three methods are all based on multidimensional knowledge graph recommendation, among which Leitner system is one of the most commonly used rule-based knowledge recommendation methods [9].

Muley et al. proved the effectiveness of another rule-based Super Memo algorithm through experiments. Hoi and others define and classify online learning. Wang modeled knowledge points and courses by combining the results of learners' learning and answering questions. On the basis of establishing the optimization model, the optimization strategy of the advanced knowledge graph is designed, and the optimization topology diagram is used to guide the learner's behavior, and the personalized learning strategy is designed for the learner. In order to meet the individual needs of different learners, Shi and other scholars adopted a multi-dimensional knowledge objective framework, which can generate the most suitable learning path according to learners' learning interests and input. Reddy et al. used the deep coupled path algorithm to generate the recommended learning path and achieve the optimal evolutionary effect in the same-level memory model.

Teaching assistance systems are generally divided into three categories: one is to directly assist teachers or students in daily work or study; the second is to apply in a specific aspect, such as the use of portable devices to enhance classroom interaction. The third is to add intelligent teaching assistants to introduce artificial intelligence and make the system more intelligent. At present and for a long time to come, teachers are still the main object of classroom teaching, and the main goal of intelligent teaching

assistants is to assist teachers in teaching rather than replace teachers. Therefore, a more comprehensive teaching assistance system to assist teachers is very necessary. In the existing teaching assistance systems, many methods only enhance classroom interaction through online information exchange or using mobile or portable devices, that is, only build some special and separate applications, and fail to build an integrated information system to meet the overall challenges in classroom teaching. As far as intelligent teaching assistants are concerned, most of them take over the work of real teachers as online teachers, or focus on a specific aspect of assistance, and cannot help teachers to carry out comprehensive classroom management and auxiliary teaching.

## 4 Bayesian Knowledge Tracking Model

Bayesian knowledge tracking model was first developed in 1995 by Corbett A. T. And Anderson J. R. proposed, and it was later widely used to measure learners' mastery of Knowledge Point (KP). As a HMM, the BKT model summarizes the teaching process into a hidden Markov process, where  $S$  represents the knowledge state of students at different times, represents the results of students' participation in the test under different knowledge states, and  $a$  represents students' learning behavior under different definition modes. Markov values are also known as no after-effect values, which indicate that the latter state is only related to the current state value and not to the previous state value. For example, in game scenario design, each action is about the behavior under the previous state demarcation, rather than describing the cause and effect of the current state. The sequence of states produced under the definition of the correlation sequence description is called Markov chain. The relationship fitting of iterative function based on the law of state value change is usually called state cyclic transition function. If the state loop transition function contains the action value set in the current state, that is, the state change value will be affected by the subsequent action of the current state, which is called the Markov Decision Process (MDP). For example, in the case of solving a binary equation, the next parameter state is determined by the current solving parameter, and the solving process is affected by a state. The BKT model diagram is shown in Fig. 1.

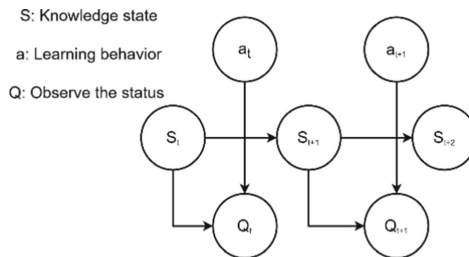


Fig. 1. The BKT model diagram

In MDP, bystanders usually know exactly what is going on. In a sports event, the audience can see a scene related to the current game. However, in the index of examining the learning process, the tester cannot grasp the accurate state of students' learning

effect through observation, but can only make indirect evaluation through examination, practical operation, interview and other ways. For example, the Markov process that examines the learning effect of students is called the stage observable Markov process, which means that the examiner can only obtain a series of observable indirect evidence. The demonstration diagram of teaching intelligence recommendation based on Markov is shown in Fig. 2.

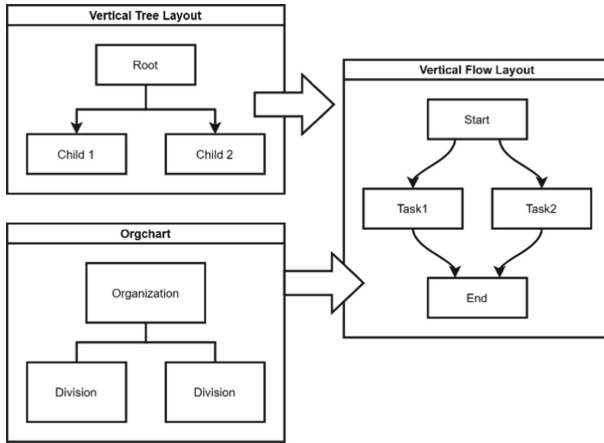


Fig. 2. Markov based instructional intelligence recommendation diagram

## 5 Platform Based on Reinforcement Learning Algorithms

According to the output type, classroom teaching learning algorithms can be divided into two types: one is the model based on learning effectiveness, the other is the independent model based on learning process. The model algorithm based on learning effectiveness uses instantaneous data to simulate the feedback given by the environment, while the model-independent feedback is treated as invalid data. The algorithms used to select classification are also divided into different types, one is value-based classification scheme, and the other is policy-based classification scheme. In the value-based classification scheme, each feedback of the algorithm represents the value of the corresponding action, and the subsequent connection of the action is determined according to the entropy ranking of each action. Commonly used algorithms include Sarsa and Q-learning. In entropy-based algorithms, the input to each algorithm is the basic probability of all actions, and the actions are classified and sampled according to the probability.

In order to achieve more complex functions, other kinds of structures are usually introduced into the neural network. For example, convolutional neural network data acquisition adds linear connections such as convolutional layer, descending hierarchy and pooling layer, and cyclic neural network designs the influence of input on feedback output at the latter moment. Combining reinforcement learning (DRL) data with deep learning mining is the main scheme to deal with the interaction between related agents and

unrelated environments in machine learning. Taking learning algorithm as an example, it uses a reward value matrix to record the rewards that can be obtained by taking various behaviors in all states. However, when the state space is too large, using neural network instead of this matrix becomes a better solution. Combined with the advantages of deep learning, reinforcement learning can accurately capture the characteristics of each dimension of the environment in many more complex and practical problems, so as to make better decisions. By maximizing the value of the objective function, students' mastery of the corresponding knowledge points can be obtained, as shown in Formula 1.

$$p(O\|K, L, T(g, s)) = \prod_{i=1}^m C(q_i) \prod_{j=1}^n W(q_j) \quad (1)$$

## 6 Design of Knowledge Tracking Model

Accurate knowledge tracking for students is the basis of intelligent recommendation system. BKT model is one of the most commonly used student knowledge tracking models, but it does not consider the influence of knowledge point correlation on learning, and can only update the model through the test results. In consideration of the shortcomings of BKT model, an improved interactive Bayesian Knowledge Tracing (IBKT) model was designed to track students' learning states. IBKT covers more elements and processes than the original BKT model. The elements include students' student number, name, age and other basic information, as well as students' personality indicators and mastery of each knowledge point, such as affinity to people, responsibility and learning habits. The process includes the learning process and the observation process. Learning process refers to the process in which the mastery of knowledge points changes after a period of learning. The observation process includes classroom observation and examination observation. Classroom observation refers to students' self-cognition of knowledge points through classroom voting, while examination observation refers to students' learning situation through examination results. Since students' learning process is usually invisible, such model updating can correct the errors caused by the evolution of one's knowledge point mastery during the learning process. With the support of more elements and processes, IBKT model can achieve far more support than BKT model can give to the teaching process. In terms of knowledge tracking, IBKT model can update and track students' learning status more timely through various interactive processes.

## 7 Conclusions

An intelligent teaching assistant system based on knowledge points is designed and implemented, which is composed of teachers' mobile phones, students' smart phones/watches, classroom computers and back-end servers. Based on the actual teaching needs, the system provides interactive and planning support for computational teaching, strengthens the interaction between teachers and students from the bottom up, and promotes the quantification and precision of classroom teaching. The main functions of the system include: before class, teachers can get the recommended teaching plan and



class grouping plan; in class, teachers can use the learning materials saved in the system, such as teaching videos, teaching plans, etc. After teaching, teachers can initiate a vote or question and answer based on knowledge points; Students can always consult the system to assess their ability, and get recommended self-study plan. Students' learning situation is captured by using knowledge-based voting and answering methods to model students. Based on the Bayesian knowledge tracking model, this model introduces the correlation degree between knowledge points to make it more suitable for the actual classroom teaching. The method of student self-evaluation to update the model, which reduces the model's dependence on the test, reduces the cost of model update and avoids the consumption of students' attention.

**Acknowledgements.** Analysis on the Types and Meanings of Common Intercultural Information in English News. Youth Fund Project of Jilin Agricultural Science and Technology College, JLASTU Contract No. 2017 [228].

## References

1. Gareche, M., Hosseini, S.M., Taheri, M.: A comprehensive literature review in competitive advantages of businesses. *Int. J. Adv. Stud. Humanit. Soc. Sci.* **8**, 223–240 (2019)
2. Injadat, M., Salo, F., Nassif, A.B.: Data mining techniques in social media: a survey. *Neurocomputing* **214**, 654–670 (2016)
3. Oluwaseun, O.O.: Employee training and development as a model for organizational success. *Int. J. Eng. Technol. Manage. Res.* **5**, 181–189 (2018)
4. Rebouillat, S., Steffenino, B., Lapray, M., Rebouillat, A.: New AI-IP-EI trilogy opens innovation to new dimensions; another chip in “the innovation wall”, what about emotional intelligence (EI)? *Intell. Inf. Manag.* **12**, 131–182 (2020)
5. Rebouillat, S.: A science & business equation for collaborative corporate innovation. business strategy, IP strategy, R&D strategy: an all-in-one business model. A review with a bio-technology & green chemistry focus. *Int. J. Innov. Appl. Stud.* **4**, 1–19 (2013)
6. Rebouillat, S.: IP & innovation, within the context of frontier technologies; a matter of accessibility with “Open” questions? *Intell. Inf. Manag.* **14**, 93–103 (2022)
7. Akinwande, M.O., Dikko, H.G., Samson, A.: Variance inflation factor: as a condition for the inclusion of suppressor variable(s) in regression analysis. *Open J. Stat.* **5**, 754–767 (2015)
8. Liang, S., Leby, K. and McCarthy, P.: wiseCIO: web-based intelligent services engaging cloud intelligence outlet. In: *SAI 2020: Intelligent Computing*, vol. 1, pp. 169–195 (2020)
9. Liang, S., Mak, L., Keele, E., McCarthy, P.: iDATA-orchestrated wiseCIO for anything-as-a-service. In: *FICC 2021: Advances in Information and Communication*, vol. 1363, pp. 401–424 (2021)
10. Liang, S., McCarthy, P., Van Stry, M.: DATA: Digital Archiving and Transformed Analytics. *Intelligent Information Management (IIM)* **13**, 70–95 (2021)



# Construction of User Profile for Teaching Archive Knowledge Service Domain

Wenyan Guo<sup>1</sup>, Liming Du<sup>2</sup>(✉), Fengying Wang<sup>2</sup>, Lei Cui<sup>1</sup>, and Jie Dong<sup>1</sup>

<sup>1</sup> School of Computer Science and Engineering, Shenyang Jianzhu University, Shenyang, Liaoning, China

<sup>2</sup> School of Information Engineering, Suqian University, Suqian, Jiangsu, China  
duliming1976@163.com

**Abstract.** With the arrival of the 5G era, the traditional archival service mode has changed. As an important support for schools to carry out education and teaching work, it is indeed necessary to promote the teaching archives of colleges and universities from information service to knowledge service. This paper firstly explains the current situation of the development of knowledge services for university teaching archives and defines the basic concept of user portraits. Then constructs a teaching archives user portrait set through K-means clustering algorithm, and builds a teaching archives knowledge service model on this basis, aiming at providing better service for the users of teaching archives and improving the awareness of related groups to the teaching archives in colleges and universities.

**Keywords:** User Profiling · Teaching Profile · K-Means Clustering · Knowledge Services

## 1 Introduction

As a product of school teaching activities, university teaching archives record the history and current situation of the school, and also undertake the important task of providing talents for society. However, with the advent of the information society, the disadvantages of the university archives management service system being out of line with the development of society have gradually come to the fore, and the mining of personalized archival information from the massive amount of data to precisely match the users has gradually become a new issue facing the university teaching archives service field under big data.

In recent years, intelligent analysis of knowledge services based on user portraits has been widely applied in the field of library intelligence. As a means of analyzing user needs and realizing accurate knowledge services in the era of big data, most scholars have adopted this method to realize personalized recommendation services in libraries. However, the results of using user portraits to study the personalized service of teaching archives are relatively few, especially in the knowledge service of teaching archives in colleges and universities, where there is still a large research gap. Therefore, in this

paper, by collecting the basic attribute information and behaviour information of users, after cleaning, classifying and fusing the basic data, the K-means clustering algorithm is used to construct a user portrait set and build a knowledge service model for teaching archives on this basis [1]. This paper provides feasible suggestions for the development of teaching archives knowledge services in universities.

## 2 User Profiling Related Technologies

### 2.1 The Basic Meaning of User Profiling

With the development of artificial intelligence, the research of user portrait rises quietly at home and abroad. User portrait was first proposed by Alan Cooper, the father of interaction. Subsequently, numerous scholars have carried out researches on user portrait successively. Korsgaard constructs personas using a design and evaluation modeling approach to qualitative data through information and communication technologies [2]. Through the potential motivations and concerns of users, Guang min Li mines representative invisible data to fully present the user group portrait [3]. Mingyu Lee used questionnaires to build profiles of millennial users that included demographic information, user behavior, motivation and personality [4].

On this basis, this paper holds that the user profile in the knowledge service environment of university teaching archives refers to the collection, analysis and processing of users' personal attribute information, behavioural attribute information and feedback attribute information, etc., to develop a highly refined feature identification that fits the characteristics of the user group, and then categorise users with similar attributes to make the user image concrete.

### 2.2 Algorithm Design for Constructing User Profiles Based on K-means Clustering

#### 1) k-value determination algorithm

Before performing K-means clustering, a predetermined k-value is needed to indicate the number of clusters to be formed, and the size of the k-value directly affects the clustering effect[5]. In practice, the contour coefficient method can be used to determine the clusters of the data to be analysed because it takes into account both cohesion and separation. Its core formula is.

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (1)$$

The specific steps to determine k value by contour coefficient method are as follows.

Step 1: Calculate the average distance  $a_i$  from sample  $i$  to other samples in the same cluster and refer to  $a_i$  as the intra-cluster dissimilarity of sample  $i$ .

Step 2: Calculate the average distance  $b_i^j$  of all samples from sample  $i$  to some other cluster  $K_j$ , and is the dissimilarity of cluster  $K_j$  of sample  $i$ .

Step 3: Define the contour coefficient of sample  $i$  based on the intra-cluster dissimilarity  $a_i$  and inter-cluster dissimilarity  $b_i$  of sample  $i$ .

## 2) K-means algorithm

The K-means algorithm is an unsupervised learning algorithm for solving clustering problems. The algorithm is simple in principle, easy to implement, and is able to remain scalable and efficient when dealing with large data sets, and it is also more effective when the clusters are close to a Gaussian distribution[6]. In addition, the algorithm is highly interpretable and often yields satisfactory clustering results when appropriate values of  $k$  are chosen.

The classical K-means algorithm proceeds as follows.

- a) Initialization: The total number of clusters  $k$  of the known dataset  $X$  and the clusters obtained by the contour coefficient method, and the  $k$  data points calculated in the dataset  $X$  are selected as the initial cluster centers, i.e. the  $k$  cluster centers are obtained.
- b) Set the iteration termination condition: set the maximum number of cycles or the change error of the cluster center as the iteration termination condition.
- c) Update the class to which the sample objects belong: calculate the distance between the data objects in dataset  $X$  and the initial centroids of the  $k$  clusters, classify each data object into the class with the closest distance according to the distance criterion, and establish the  $k$  initial clusters with the original centroids as cluster centers.
- d) Update the class centroids: the next iteration's clustering centers are the average vectors of each class.
- e) Repeat steps (c) and (d) until the iteration termination condition in step (b) is met, i.e. the clustering results have converged, then stop.
- f) Output the clustering results.

## 3 Construction of a User Profile for Teaching Profiles

### 3.1 Data Collection

User data is an implicit expression of users' needs and to a certain extent can help university archives understand users' needs preferences. Ideally, the source of data should be log data from archives' user management and service platforms, system portals, etc. However, the current Shenyang Jianzhu University archives system is immature, user information is not well recorded, and in consideration of the need to protect personal privacy information, the Shenyang Jianzhu University archives does not provide user search data to the public. Therefore, this paper uses a questionnaire to simulate the user login system and user search system of the archive system based on basic user information and user behaviour information to obtain personal information and demand preferences of users, so as to build a user profile of the teaching archives of Shenyang Jianzhu University.

### 3.2 Data Pre-processing

A total of 263 questionnaires were collected online and offline. As the original user data collected through the questionnaire had problems such as missing data and inconsistencies, it would affect the accuracy of the user profile construction. Therefore, the data needs to be pre-processed such as cleaning and transformation. The article mainly performs the following four operations: eliminating non-sample-required questionnaires, filling in missing values, eliminating and filling in incorrect values and eliminating duplicate questionnaires. Some of the statistical information is as follows, as shown in Table 1 and 2.

**Table 1.** Statistical table of user information (partial)

Gender	Identity	Type	Way	Willing
Female	student	Personal	Archives	Need
Male	student	Enrollment	APP	Need
Female	teacher	Major	Archives	Needless
Male	teacher	Enrollment	Archives	Need
Female	Social group	Personal	Website	Need
Female	teacher	Enrollment	APP	Need
Male	student	Personal	APP	Need
Female	Social group	Personal	Archives	Need
Female	student	Personal	Archives	Neutral
Male	student	Personal	Archives	Need
Female	Social group	Major	APP	Need
Male	Social group	Personal	Archives	Need

**Table 2.** Comparison table for labelling of user information

Gender	Identity	Type (Querying a File Type)	Way (File query Method)	Willing (establish knowledge service system)
Male	Student	Enrollment	Archives: Enquiry at Library	Need: necessary
	Teacher	Personal: Personal information	Website: Query official website	Needles: unnecessary
Female	Alumni	Major: Professional Courses	APP: Portal website APP	Netural: Don't care
	Social group	Course Practice		

### 3.3 Building User Profiles Based on K-means Clustering

1) Determine the value of k

In this paper, we use python software to find the best k value, as shown in Table 3. According to the contour coefficient principle, when the contour coefficient refers to the maximum, its corresponding number of clusters is the best[7], from the results, when the k value is 4, the clustering effect is the best.

**Table 3.** Table of contour coefficients corresponding to k values

K-value	.....	2	3	4	5	6	.....
Corresponding profile factor	.....	0.323	0.329	0.338	0.305	0.315	.....

2) Perform K-means clustering analysis

After determining the optimal k value, After determining the optimal k value, K-means cluster analysis was carried out on the data, and the results were shown in Table 4. The sample population was divided into four different classes of user portraits, with each column corresponding to the character at the centre of that cluster.

**Table 4.** K-means clustering results

Attribute	Full Data (232.0)	Cluster0 (90.0)	Cluster1 (71.0)	Cluster2 (63.0)	Cluster3 (8.0)
Gender	Female	Female	Female	Male	Female
Identity	student	teacher	student	student	alumni
Type	Enrollment	Course Practice	Enrollment	Personal	Personal
Way	Website	APP	Website	Archives	Website
Willing	Need	Need	Need	Need	Neutral
<b>Clustered</b>	<b>Instances</b>	39%	31%	27%	3%

The specific results are analyzed as follows.

a) Cluster 0

This portrait collection is mainly composed of teachers with a certain number of years of work, who have higher information literacy, greater demand for teaching archives, higher query methods, and more comprehensive file search content.

b) Cluster 1

From the data in the table, we can infer that the portrait collection is mainly a group of students facing admission, who need to achieve their admission goals through

the school enrollment plan. In addition, problems such as incomplete information search and difficulty in obtaining may be the reasons why this group needs to build a teaching archives knowledge service system.

c) Cluster 2

From the clustering results of the portraits, it is clear that this group is more willing to go to the university archives to find relevant archival materials because of the privacy of the archives they query, and they hope that the university archives can improve the relevant query system and establish an intelligent knowledge service recommendation system for teaching archives.

d) Cluster 3

The clustering results show that alumni do not make much use of teaching archives and are not sufficiently aware of the importance of teaching archives. It is also possible that this group is indifferent to the archival knowledge service system because of the low level of demand for archives and the fact that access is open and relatively easy to obtain.

### 4 Knowledge Services for Teaching and Learning Archives Based on User Profiling

In order to better tap into users' real archival demand preferences based on the constructed user profile set, and at the same time provide accurate teaching archival services to users. In this paper, a user-oriented knowledge service model for teaching archives is constructed from four aspects: data collection, demand analysis, data processing and recommendation service. It is shown in Fig. 1.

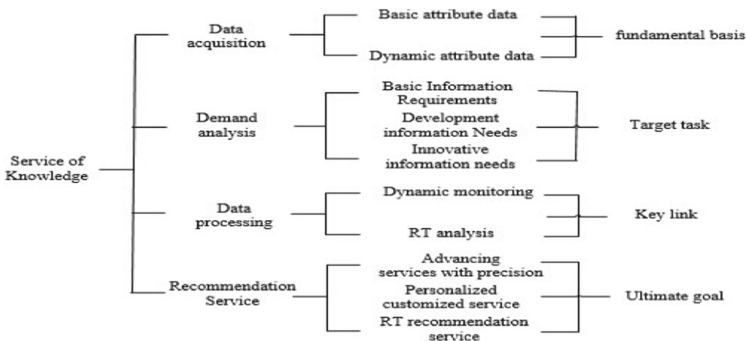


Fig. 1. Knowledge service model for teaching archives

#### 4.1 Data Collection Layer: Acquisition User Data

In realizing the knowledge service of teaching archives, collecting relevant user information is the basic basis for constructing user portraits and knowledge service models. The basic attribute data and dynamic attribute data of users can be obtained by using crawler technology or questionnaire through college portal web logs [8].

## 4.2 Needs Analysis Layer: Exploring User Needs

Exploring user's demand and solving user's problem is the target task of constructing teaching archive knowledge service model. Data processing and recommendation services are based on demand analysis, and use the collected data to explore the direct real needs and potential hidden needs of users[9]. Demand analysis shows three types of basic information demand, development information demand and innovation information demand. Teaching archive knowledge service needs to carry out corresponding construction and service according to the demand content.

## 4.3 Data Processing Layer: Building a User Profile

Building a user profile is a key part of realising the knowledge services of the teaching archive. The data processing layer mainly builds user portrait set through data, and establishes service item index for each portrait set and completes dynamic feedback and adjustment, so as to provide corresponding service content for different types of users to realize personalized recommendation[10]. In this process, considering the change of user characteristics and interest preference, the system needs to have a synchronous optimization and update mechanism, so as to establish a more suitable portrait set for the user group.

## 4.4 Recommendation Service Layer: Realizing Knowledge Services

The ultimate goal of the knowledge service for teaching archives is to achieve accurate recommendation services. While the teaching archives knowledge service model provides basic services, its highlight lies in providing scenario-based and real-time precise personalized services according to user behavior characteristics and demand characteristics. It includes precision recommendation service, personalized customized service and real-time recommendation service.

## 5 Conclusions

The knowledge service of university teaching archives based on user portraits is a user-centred process of providing diversified services, which to a certain extent alleviates the problems of information overload and semantic gap between teaching archives and users. Based on user big data, the article accurately predicts users' archival demand preferences through modelling of user portraits, and realises a seamless connection between users and resources. Through the construction of teaching archives knowledge service model, the intelligent level of college archives service can be effectively improved.

**Acknowledgements.** This work was supported by Liaoning archival science and technology project "Research on the knowledge service of university teaching archives in the context of engineering certification" (project number: 2022-R-022) and Suqian University Scientific Research Fund for Talent Introduction "Research on Analysis, Testing and Anomaly Diagnosis Methods for Complex Systems" (project number: 2022XRC091).



## References

1. Xue, H., Guo, K., He, F.: Research on personalized recommendation service of mobile library based on user portrait. *Int. J. Adv. Comput. Sci. Appl.* **13**(10), 675–683 (2022)
2. Korsgaard, D., Bjørner, T., Sørensen, P.K., Burelli, P.: Creating user stereotypes for persona development from qualitative data through semi-automatic subspace clustering. *User Model. User-Adap. Inter.* **30**(1), 81–125 (2020)
3. Guangmin, L., Wenjing, C., Xiaowei, Y., Li, W.: Modeling and analysis of group user portrait through WeChat mini program. *Wirel. Commun. Mob. Comput.* **1–8**, 2022 (2022)
4. Lee, M., et al.: Developing personas & use cases with user survey data: a study on the Millennials' media usage. *J. Retail. Consum. Serv.* **54**, 102051 (2020)
5. James, O.A., Nkele, M.D., John, I.S., Mahfouz, A.M.A.: AutoElbow: an automatic elbow detection method for estimating the number of clusters in a dataset. *Appl. Sci.* **12**(15), 7515–7515 (2022)
6. Hassan, I., Mohammad, I.: K-means-G\*: accelerating K-means clustering algorithm utilizing primitive geometric concepts. *Inf. Sci.* **618**, 298–316 (2022)
7. Kong, W., Wang, Y., Dai, H., Zhao, L., Wang, C.: Analysis of energy consumption structure based on K-means clustering algorithm. In: *E3S Web of Conferences*, vol. 267, p. 01054 (2021)
8. Tang, Y., Li, S., Song, W., Zhou, C., Niu, Z.: Service recommendation based on dynamic user portrait: an integrated approach. In: *American Society of Mechanical Engineers*, vol. 85376, p. V002T02A056 (2021)
9. Tinggui, C., Chu, Z., Jianjun, Y., Guodong, C.: Grounded theory-based user needs mining and its impact on app downloads exemplified with WeChat APP. *Front. Psychol.* **13**, 875310 (2022)
10. Li, Y., Wang, R., Nan, G., Li, D., Li, M.: A personalized paper recommendation method considering diverse user preferences. *Decis. Support Syst.* **146**, 113546 (2021)



# Artificial Intelligence Technology in Computer Network Security

Cuijin Lao and Shen Qin<sup>(✉)</sup>

Liuzhou City Vocational College, Liuzhou 545000, Guangxi, China  
shenqinsir@126.com

**Abstract.** Currently, computer networks are evolving in an extensive and highly distributed direction, Data centers based on traditional networks are gradually migrating to cloud data centers, and instead of reducing the number of network security incidents, they are causing more damage, and traditional reactive protection methods cannot reliably cope with changing security events and security hazards. The purpose of this paper is to study computer network security based on artificial intelligence technology. The application of artificial intelligence techniques in network security with intrusion detection and defense techniques is described. Because the learning efficiency of BP algorithm is low and the convergence speed is slow, the additional momentum factor is used to improve the performance of BP algorithm, and the neural network is introduced into the intrusion detection and defense system, so as to reduce the system leakage rate and false alarm rate and improve the convergence speed of neural network. At the same time, various network security products developed in combination with artificial intelligence will be deployed to the cloud data center to strengthen the network security of the cloud data center.

**Keywords:** Artificial Intelligence Techniques · Computer Networks · Network Security · BP Algorithm

## 1 Introduction

With the widespread use of computers and networks in the financial and business sectors, society as a whole is increasingly dependent on computer networks. However, the forms and means of cyber security threats have changed, it shows new characteristics. The network intrusion causes the information system to be unable to operate effectively, seriously hindering economic activities, causing huge losses, and even social turbulence. Network security has become the bottleneck of social information development, and network security research has increasingly become the focus of global attention [1, 2].

With the rapid development of new technologies such as cloud computing, in order to ensure the security of computer networks in the new situation, we must take some defensive measures and strategies. The application of data mining techniques in computer networks allows retrieval, processing and analysis of data in computer networks [3, 4]. Daniel Gerbi Duguma analyzed the policies related to information security of

computer networks and further optimized network information security by combining data mining techniques to provide a reference for evaluating the effectiveness of network information security [5]. R. Punithavathi has designed a dedicated automatic network security detection card using coretex - m 3 processor according to the requirements of the automatic network security detection function of the embedded data acquisition terminal. In addition, the RS232 interface has been extended to communicate with the data acquisition terminal and to collect the necessary cyber security events [6]. Abida Sultana Sumona started by outlining a road map of early mobile communication and artificial intelligence, the technical development from the third generation mobile communication technology (3G) to the era of artificial intelligence was reviewed. In the context of telecommunication AI, the convergence of business operations management, business support system (BSS) and operations support system (OSS) oriented intelligent applications is further detailed [7]. With the development of new technologies, it is urgent to establish a new security defense system suitable for large-scale networks, so as to actively, cooperatively and effectively deal with security threats by integrating various network security technologies.

In this paper, we introduce network security management, particularly intrusion detection and artificial intelligence. The artificial intelligence technology in network security management is analyzed. Using the artificial neural network of artificial intelligence technology, the intrusion detection system is improved, and the deployment of network security equipment based on artificial intelligence in the cloud data center is improved, in order to bring new security defense strategies to the network security field.

## **2 Research on the Application of Artificial Intelligence Technology in Computer Network Security**

### **2.1 AI in Network Security**

#### **(1) Establishing rule-generating expert systems**

The most important part of network security management is intrusion detection [8, 9]. So far, expert systems are the most popular artificial intelligence techniques. An expert system is a detection system with specialized knowledge. Administrators can collate existing intrusion patterns and use system detection to identify system security factors, which also helps to detect future intrusions [10, 11].

#### **(2) Use of Artificial Neural Network System in Network Security Management**

Artificial neural networks have a strong discriminatory ability to identify intruders with intentional noise or even malicious distortions. The system was developed by an efficient research team to simulate the learning ability of the human brain over a long period of time [12, 13].

### (3) Artificial immunity techniques

Network security is one of the application fields of artificial immune technology. Artificial immune technology is based on the human immune system. The human body can be protected in the external environment through immunity, while the network system security devices can realize the defense against intrusion through the learning of information security technology. The former focuses on preventing virus intrusion, while the latter focuses on protecting information to ensure its integrity, confidentiality and availability [14, 15].

## 2.2 Intrusion Detection and Defense Techniques

IDS is the abbreviation of intrusion detection system. There are two main types of IDS: network-based IDS and host based IDS. The data detected by network-based IDS comes from the data packets on the network, independent of the hosts on the protected network, and is suitable for intrusion detection of various network traffic. The data detected by host based IDS comes from the system log, audit log, and operating system log of the protected host. Generally, only one specific computer is protected [16, 17]. IDS can be divided into two parts, network detector and management center. The network probe monitors the inbound and outbound access to the protected network segment in real time, reports the detected violations to the console, and executes the commands returned by the console. At the same time, the network probe has client daemons that support the universal connection control protocol, which can communicate with the firewall server daemons in real time. If a violation is detected, the console will be notified of the violation message, and the firewall will also be notified. The firewall will automatically generate rules; The management center provides a graphical interface for centralized management of network detectors [18].

IPS is the abbreviation of intrusion prevention technology. IPS has more defense functions than IDS. Although IDS can prevent ordinary attacks by linking the detection results with the firewall, it has poor protection against “instant attacks”. IPS can solve this problem well, blocking the detected attacks while detecting them.

## 2.3 Improvement of BP Algorithm

The neural network detection mechanism compares the actual user behavior patterns with the learned normal behavior patterns to determine whether they are abnormal or not.

In this paper, we use the additional moments method and the adaptive learning rate method to improve the performance of the BP algorithm to handle the problem.

The correction formula for the additional moment factor weights is:

$$\Delta W_{ij}(k+1) = (1 - mc)1Err_i + mc\Delta W_{ij}(k) \quad (1)$$

$$\Delta \theta_j(k+1) = (1 - mc)1Err_i + mc\Delta \theta_j(k) \quad (2)$$

$k$  is the training number,  $mc$  is the moment factor, and  $l$  is the learning rate.

The principle of adaptive learning rate adjustment is as follows: if the weight correction can reduce the value of the error function and the selected learning rate is slightly smaller, then a smaller value can be increased to improve the learning speed and shorten the computation time. If the value of the error function cannot be reduced but is increased, the learning rate must be reduced, and the specific adjustment is as follows:

$$\eta(k+1) = \begin{cases} 1.051\eta(k) & E(k+1) < E(k) \\ 0.71\eta(k) & E(k+1) < 1.04E(k) \\ \eta(k) & \text{Other situations} \end{cases} \quad (3)$$

with adaptive learning rate, the selection of initial learning rate is more flexible. In the intrusion detection system based on neural network technology implemented in this paper, the neural network module is also the applied BP model.

### 3 A Systematic Study on the Application of AI in Computer Network Security

#### 3.1 System Framework

Considering the maturity of BP neural network used in engineering and the real-time requirement in intrusion detection and defense, the system adopts BP neural network as the artificial intelligence part and works in parallel with the traditional intrusion detection and defense part. The system consists of 3 modules: sniffer module, intrusion detection and defense module, and neural network module.

Sniffer module is responsible for getting packets from the network (for simplicity, in this paper, messages etc. are considered as packets). If you use switch in your network, you also need to configure the mirror port on the switch to listen to all packets. Considering the compatibility of LINUX and WINDOWS, the system decides to use pcap.

The neural network module and the intrusion detection and defense module work in parallel and complement each other. The preprocessing subcomponent is responsible for the numerical conversion function: converting network packets of different protocol layers into appropriate mathematical vectors. Since ICMP, TCP/UDP, HTTP, and TELNET are widely used for network communication, we perform separate numerical conversions for packets of all four protocols and use the same numerical conversions for all other types of packets. The neural network parameter settings include input vector size, excitation function type, learning efficiency, moment factor, and the number of training sessions that can be defined by the parameters. The reporting sub-module is used to determine whether it is an attack or not.

#### 3.2 Neural Network Module

In our implementation of a neural network intrusion detection and defense system, a sample of a typical network attack is first obtained as a training sample bank to train the classification mechanism of the neural network. After training the neural network stores

the characteristic patterns of these weight attacks. After the training, the neural network classification mechanism can be used to analyze and process the network data stream captured in actual operation, and determine whether it is a normal network behavior or an unknown network attack. When an unknown network attack is detected, the feedback training sample library retrains the neural network, classifies the neural network through training, and continues to expand the scope of detection.

The application of neural network in intrusion detection and defense system plays an important role in reducing the system leakage rate and false alarm rate. The improved BP algorithm further reduces the system leakage rate and false alarm rate, and improves the convergence speed of neural network.

### 4 Network Security Deployment of Cloud Data Center

Cloud data centers developed based on virtualization technology have performed well in flexible resource scheduling and full utilization of resources. Traditional data centers are gradually migrating to cloud data centers. Combined with the advantages of network security products developed by AI in intelligence and security, it is suitable for deployment to cloud data centers. The network security management and control deployment of the cloud data center is shown in Fig. 1.

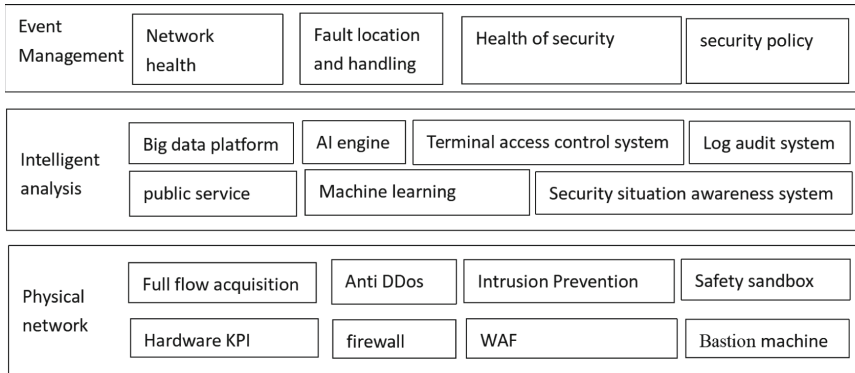


Fig. 1. Network security deployment diagram of cloud data center

The network security management and control of the cloud data center is divided into three layers: the physical network layer includes the network equipment and security equipment of the cloud data center, such as Anti DDos, firewall, intrusion prevention system, WAF, security sandbox, fortress machine, etc.; The intelligent analysis layer provides public services based on the big data platform, receives the data reported by the physical network layer, uses artificial intelligence to analyze the network security of the reported data, controls the access of terminals, audits servers, databases, network devices, network traffic, etc., and predicts the security situation awareness; The event management layer receives the audit information and security situation prediction of the intelligent analysis layer, judges the network health and security health of the cloud data

center, finds security events and faults in time, conducts fault location and processing, and dynamically generates security policys and sends them to all network security devices for joint defense.

The network security equipment deployment of the cloud data center is shown in Fig. 2. It is divided into vertical external threat security protection deployment and horizontal internal threat security protection deployment. The vertical external threat security protection deployment includes Anti DDoS, firewall, intrusion prevention system, WAF, security sandbox, etc. Horizontal internal threat security protection deployment includes firewall, intrusion prevention system, WAF, security sandbox, virtual firewall (VFW), terminal access control system, etc.

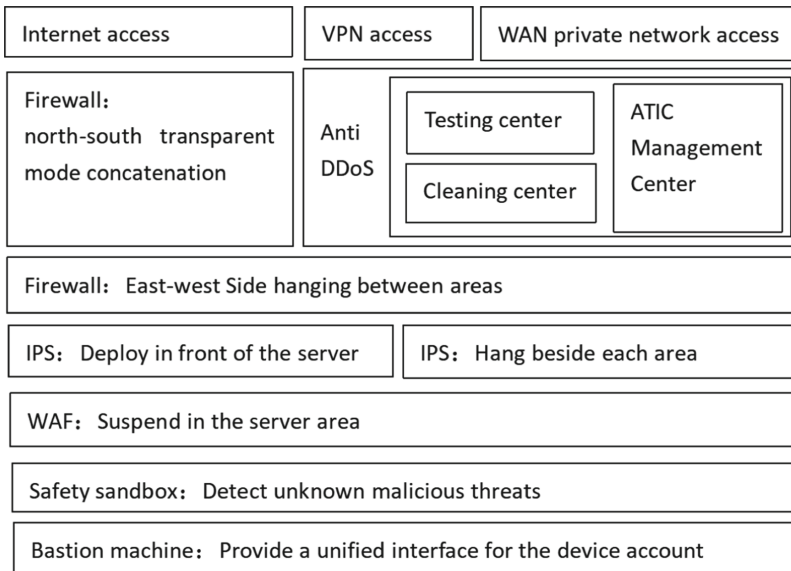


Fig. 2. Network security equipment deployment diagram of cloud data center

First, the firewall is deployed in series in the vertical transparent mode at the Internet, VPN or WAN private network access to protect the vertical security and prevent external intrusion. At the same time, a firewall is hung between the horizontal areas inside the cloud data center to protect the security between the horizontal areas.

Because the exit of the cloud data center faces the threat of distributed denial of service attacks, it is also necessary to deploy anti DDoS devices, mirror all access traffic to the detection center for detection, and then report it to the ATIC management center, which will clean the traffic according to the security policy to effectively prevent DDoS attacks and protect normal business traffic.

In order to defend against various attacks against business servers, IPS devices are directly deployed in front of the business server cluster to protect the security of the business server area; In order to monitor internal attacks, IPS devices are deployed inside the data center to detect abnormal data traffic.

In order to defend against SQL injection, cross site scripting and other attacks against the Web server, a web application firewall WAF is installed next to the server area to protect the Web server.

To prevent undiscovered and unpublished attacks, security sandboxes are deployed for detection. The security sandbox simulates the real environment through virtualization technology, allowing malicious attack processes to fully run in the virtual environment without damaging the real system, effectively identifying 0-Day vulnerabilities, targeted APT and other attacks, and intercepting them with the firewall.

In order to uniformly manage the operation and maintenance accounts of all equipment for future audit and review, the bastion machine is deployed inside the cloud data center, and the whole process of account operation is recorded and saved.

Through the reasonable deployment and virtual application of network security devices based on artificial intelligence, the cloud data center can be effectively protected, intelligent management can be realized, the utilization efficiency of network security devices can be improved, and security guarantees can be provided for the reasonable scheduling and utilization of various resources.

## 5 Conclusions

In today's world, information security has become a global issue concerning national defense security, social stability and economic development, and the international struggle to obtain, use and control information is becoming increasingly intense. In this paper, some exploratory work has been carried out in this area of computer network security based on artificial intelligence technology, and in order to further improve the network security active defense system, the following work needs to be continued in future work: Mapping the alarm types of heterogeneous IDSs such as Snort to expanded class attacks to achieve multi-detector data fusion; In future work, it is necessary to target different network worms. A feedback mechanism between the prediction results and inference rules is needed to train the optimal fuzzy inference rules.

## References

1. Silva, J.V.V., de Oliveira, N.R., Medeiros, D.S.V., Lopez, M.A., Mattos, D.M.F.: A statistical analysis of intrinsic bias of network security datasets for training machine learning mechanisms. *Ann. des Télécommunications* **77**(7–8), 555–571 (2022)
2. Senthil Kumaran, V.N., Ananthi, G.: Artificial noise aided polar code with optimal jamming position for physical layer security in mondrian loss integrated rayleigh wireless relay channel. *Ad Hoc Sens. Wirel. Netw.* **51**(1–3), 205–234 (2022)
3. Nikhade, J.R., Thakare, V.M.: BlockChain Based Security Enhancement in MANET with the Improvisation of QoS Elicited from Network Integrity and Reliance Management. *Ad Hoc Sens. Wirel. Netw.* **52**(1–2), 123–171 (2022)
4. Talpur, A., Gurusamy, M.: Machine learning for security in vehicular networks: a comprehensive survey. *IEEE Commun. Surv. Tutor.* **24**(1), 346–379 (2022)
5. Duguma, D.G., Kim, J., Lee, S., Jho, N.-S., Sharma, V., You, I.: A lightweight D2D security protocol with request-forecasting for next-generation mobile networks. *Connect. Sci.* **34**(1), 362–386 (2022). <https://doi.org/10.1080/09540091.2021.2002812>



6. Punithavathi, R., Kowsigan, M., Shanthakumari, R., Zivkovic, M., Bacanin, N., Sarac, M.: Protecting data mobility in cloud networks using metadata security. *Comput. Syst. Sci. Eng.* **42**(1), 105 (2022)
7. Sumona, A.S., Kundu, M.K., Badrudduza, A.S.M.: Security analysis in multicasting over shadowed rician and  $\alpha$ - $\mu$  fading channels: a dual-hop hybrid satellite terrestrial relaying network. *IET Commun.* **16**(1), 43–57 (2022). <https://doi.org/10.1049/cmu2.12310>
8. Nagamani, C., Chittineni, S.: Network database security with intellectual access supervision using outlier detection techniques. *Int. J. Adv. Intell. Paradigms.* **22**(3/4), 348–361 (2022)
9. Das, S., Mohanta, B.K., Jena, D.: A state-of-the-art security and attacks analysis in blockchain applications network. *Int. J. Commun. Netw. Distrib. Syst.* **28**(2), 199 (2022). <https://doi.org/10.1504/IJCND.2022.121199>
10. Mishra, S.: Blockchain-based security in smart grid network. *Int. J. Commun. Netw. Distrib. Syst.* **28**(4), 365–388 (2022)
11. Nagarajan, S.M., Anandhan, P., Muthukumaran, V., Uma, K., Kumaran, U.: Security framework for IoT and deep belief network-based healthcare system using blockchain technology. *Int. J. Electron. Bus.* **17**(3), 226–243 (2022)
12. Melo, R.V., de Macedo, D.D.J., Kreutz, D., De Benedictis, A., Fiorenza, M.M.: ISM-AC: an immune security model based on alert correlation and software-defined networking. *Int. J. Inf. Secur.* **21**(2), 191–205 (2021). <https://doi.org/10.1007/s10207-021-00550-x>
13. Eswaran, S., Rani, V., Dominic, D., Ramakrishnan, J., Selvakumar, S.: An enhanced network intrusion detection system for malicious crawler detection and security event correlations in ubiquitous banking infrastructure. *Int. J. Pervasive Comput. Commun.* **18**(1), 59–78 (2022)
14. Alshayegi, M.H., Abed, S.: Enhanced video-on-demand security in cloud computing against insider and outsider threats. *Int. J. Secur. Netw.* **17**(1), 48–55 (2022). <https://doi.org/10.1504/IJSN.2022.122550>
15. Kumar, S., Chaube, M.K., Nenavath, S.N., Gupta, S.K., Tetarave, S.K.: Privacy preservation and security challenges: a new frontier multimodal machine learning research. *Int. J. Sens. Netw.* **39**(4), 227 (2022). <https://doi.org/10.1504/IJSNET.2022.125113>
16. Ahmed, N., et al.: Security & privacy in software defined networks, issues, challenges and cost of developed solutions: a systematic literature review. *Int. J. Wirel. Inf. Netw.* **29**(3), 314–340 (2022). <https://doi.org/10.1007/s10776-022-00561-y>
17. Sarker, D.K., Sarkar, M.Z.I., Anower, M.S., Sultana, R.: Multicast network security with asymmetric cooperative relaying. *Int. J. Wirel. Inf. Netw.* **29**(3), 303–313 (2022). <https://doi.org/10.1007/s10776-022-00566-7>
18. Gharib, A., Ibnkahla, M.: Node embedding for security-aware clustering of mobile information-centric sensor networks. *IEEE Internet Things J.* **9**(18), 17249–17264 (2022)



# Dynamic Capacity Expansion System and Control Method for User-Side Power

Rongtao Liao<sup>(✉)</sup>, Huanjun Hu, Yixi Wang, Yuxuan Ye, and Jian Zhang

Information and Communication Branch of State Grid, Hubei Electric Power Company,  
Wuhan 430000, Hubei, China  
puxunhan110muw@163.com

**Abstract.** This paper provides a user side power dynamic capacity expansion system and its control method. The control method includes: uninterruptedly monitoring the power of AC input; When the power is less than the first power threshold, control the mains power to supply power to the load equipment and other devices in the system, and charge the flywheel energy storage device; When the power is greater than or equal to the first power threshold, the flywheel energy storage device is controlled to discharge, and the utility and flywheel energy storage device are controlled to jointly provide power for the load equipment and other devices in the system. The system and its control method can provide short-term high power support for the load in working state or when multiple loads are working at the same time, and achieve the effect of dynamic capacity increase without increasing the power supply capacity at the user side. Compared with the traditional power capacity expansion method, the equipment investment is low and the construction period is short. And the capacity increasing system provided in this paper can be used to support multiple sets of flywheel energy storage devices to work in parallel according to the application requirements, and the system has good scalability.

**Keywords:** Power · Dynamic Capacity Increase · Control

## 1 Introduction

In the process of production and operation, due to the expansion of production capacity and the increase of power load, the power supply capacity originally applied for cannot meet the current production and operation needs, so it is necessary to apply for increasing capacity on the original basis [1]. The process for power users to apply for power capacity increase is usually as follows: if the power user needs to increase capacity, it must submit a written application to the power supply party and issue the relevant application materials of the power user, which can be carried out only after the power supply party reviews and agrees [2–7]. When users apply for power capacity increase, it not only takes a long time to apply for and approve, but also needs to pay for equipment purchase, design, construction and other costs, which is a big expense [8].

For some power users, the normal basic load will not exceed the power supply capacity, but when their high-power electrical equipment is put into operation, it will exceed

the maximum power supply capacity for a short time [9–16]. The power consumption characteristics of these high-power equipment are intermittent use, low power in standby at ordinary times, high working power in use, and short service time each time [5]. For example, in the field of hospitals, with the development of science and technology, medical equipment such as CT, MRI, and electron accelerators are introduced into use [17]. These equipment are high-power equipment that are frequently used for a short time. The power consumed in standby mode is not large, but the power consumed in working mode increases rapidly, exceeding the maximum capacity of its power supply and distribution system design [18]. To ensure the normal operation of these high-power electrical equipment, the traditional solution is to apply to the power supply department for power capacity increase. After the approval of the power company, the survey, design and construction are carried out. It may be necessary to replace or add new transformers, old lines and corresponding power distribution equipment. It not only takes a long time, but also has a large investment.

In order to solve the problem of long time and high capital investment caused by the traditional method of power capacity increase on the user side, this paper proposes a solution of dynamic capacity increase on the user side using flywheel energy storage.

## 2 System Structure

Please refer to the structural diagram of the user side power dynamic capacity increase system shown in Fig. 1. The dynamic capacity increase system includes: mains, AC input terminal coupled with mains, load equipment, and AC output terminal coupled with load equipment; The online dual conversion uninterruptible power supply device coupled with the AC input end and the AC output end, and the flywheel energy storage device coupled with the online dual conversion uninterruptible power supply device. The flywheel energy storage device is configured when the power supply capacity of the mains cannot meet the power consumption of the load equipment and other devices (including the online dual conversion uninterruptible power supply device and flywheel energy storage device) in the dynamic capacity increase system, Together with the mains supply, it supplies power to the load equipment and other devices in the dynamic capacity increasing system. In this paper, the “first power threshold” is defined as the maximum power supply capacity of the mains under normal conditions.

The mains power supplies power to the load equipment and other devices in the system through the AC input terminal. The power required by the mains power varies according to the power of the AC input terminal, while the power required by the AC input terminal depends on the power demand of the load equipment and other devices in the system, and can only exceed the maximum supply capacity of the mains power for a short time. That is to say, when the load equipment is working, the power demand of the load equipment and other devices in the system will exceed the maximum power supply capacity of the mains, and the system will detect that the power value of the AC input terminal is greater than the first power threshold. At this time, the system will control the flywheel energy storage device to discharge, so that the power of the AC input terminal is less than or equal to the first power threshold, That is to say, the mains power will continue to supply power with the maximum power supply capacity, and the flywheel

energy storage device will provide the residual power required by the load equipment and other devices in the system.

Among them, the online dual conversion UPS device includes: AC/DC rectifier and DC/AC inverter, and the AC/DC rectifier and DC/AC inverter are connected through DC bus. The mains supplies AC power to the AC input terminal, rectifies AC power into DC power through AC/DC rectifier, and then reverses DC power into stable AC power through DC/AC inverter to supply power to the load equipment.

The online dual conversion uninterruptible power supply device also includes: a bidirectional DC/DC module set on the DC bus at the high voltage side, which is used to receive the high-voltage DC rectified by the AC/DC rectifier when there is residual power supply capacity to charge the flywheel energy storage device, and convert it into low-voltage DC suitable for the flywheel energy storage device to charge the flywheel energy storage device; When the power supply capacity of the mains power cannot meet the power demand of the load equipment and other devices in the dynamic capacity increase system, the system receives the discharge power of the flywheel energy storage device and supplies power to the load equipment and other devices in the dynamic capacity increase system together with the mains power.

The flywheel energy storage device includes a bidirectional DC/AC converter connected to the low-voltage side of the bidirectional DC/DC module and a flywheel energy storage unit connected to the AC side of the bidirectional DC/AC converter. The bidirectional DC/AC converter is used to receive the output power of the bidirectional DC/DC module to provide charging power for the flywheel energy storage unit when there is still residual power supply capacity for charging the flywheel energy storage unit; When the power supply capacity of the mains power cannot meet the power demand of the load equipment and other devices in the dynamic capacity increase system, it receives the discharge power of the flywheel energy storage unit, and then outputs it to the bidirectional DC/DC module to supply power to the load equipment and other devices in the dynamic capacity increase system together with the mains power.

In addition, the bidirectional DC/AC converter is also configured to end the discharge state of the flywheel energy storage device when the flywheel speed is lower than the first speed threshold. At this time, the flywheel energy storage device is in standby state, neither charging nor discharging.

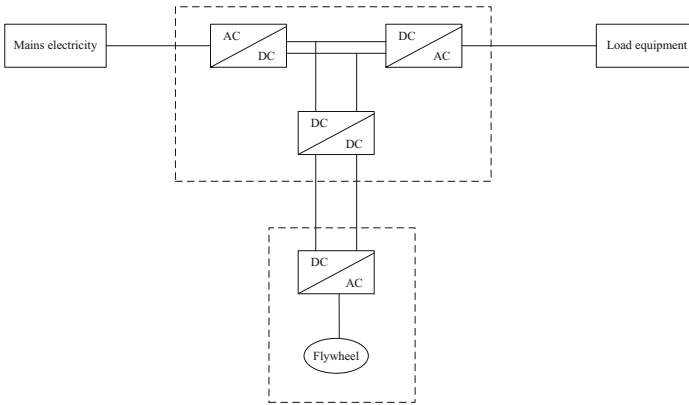


Fig. 1. Structure of power dynamic capacity increasing system on customer side

### 3 Energy Flow Direction of Dynamic Capacity Increasing System

In some feasible embodiments, the dynamic capacity increasing system also includes a control module, which is configured to perform the following steps:

Continuously monitor the power of AC input terminal; When the power of the AC input terminal is less than the first power threshold, control the mains power to supply power to the load equipment and other devices in the dynamic capacity increase system, and output the residual power supply capacity of the mains power to the flywheel energy storage device for charging. See a schematic diagram of energy flow of the user side power dynamic capacity increase system shown in Fig. 2 for details; When the power of the AC input terminal is greater than or equal to the first power threshold, control the flywheel energy storage device to discharge, and control the mains and flywheel energy storage device to jointly provide power for the load equipment and other devices in the dynamic capacity increase system. See another energy flow diagram of the user side power dynamic capacity increase system shown in Fig. 3 for details.

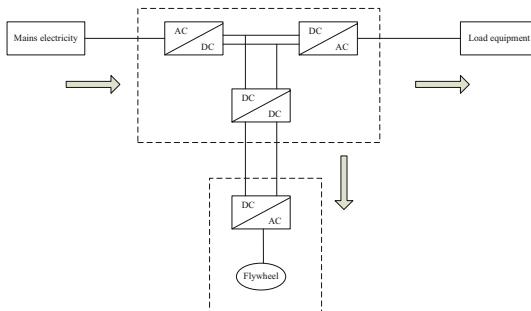
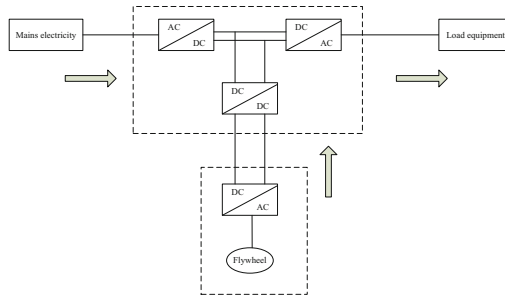


Fig. 2. A schematic diagram of energy flow direction of power dynamic capacity increase system on the user side

The control module is configured to perform the following steps: uninterruptedly monitor the power of the AC input terminal. The power monitored here corresponds to the sum of the power used by the load equipment and other devices in the system at that time, while other devices in the system include the online dual conversion uninterruptible power supply device and flywheel energy storage device.

When the power at the AC input end is less than the first power threshold, control the bidirectional DC/DC module to output power to the flywheel energy storage device, and control the sum of its output power and the power of the load equipment in the system and the power of other devices to be less than or equal to the first power threshold; When the power of AC input terminal is greater than or equal to the first power threshold, control the power of AC input terminal as the first power threshold, and control the bidirectional DC/DC module to output power to DC/AC inverter, and control the sum of its output power and the first power threshold to be equal to the sum of the electrical power of the load equipment in the system and the electrical power of other devices.

After the flywheel energy storage device discharges for a period of time, when the flywheel speed is lower than the first speed, control the bidirectional DC/AC converter to stop power output or input, and the flywheel energy storage device is in standby state, neither charging nor discharging.



**Fig. 3.** Another energy flow diagram of the user side power dynamic capacity expansion system

The first speed can be set as required, for example, 30%–40% of the highest flywheel speed. The control module comprises a programmable processor in which a software program is written, which can perform the functions described herein. Specifically, the software program can be downloaded to the processor in electronic format through the network; Alternatively, the software program may be stored on a storage medium such as a magnetic memory, an optical memory, or an electronic memory. In some cases, the control module may also include additional or embedded hardware modules for accelerating its operation. The aforementioned hardware modules may include discrete components, at least one FPGA and/or at least one ASIC.

In some feasible embodiments of this paper, the online dual conversion uninterruptible power supply device also includes a static bypass, so that when any part of the AC/DC rectifier or DC/AC inverter fails, the mains can supply power to the load equipment through the static bypass.

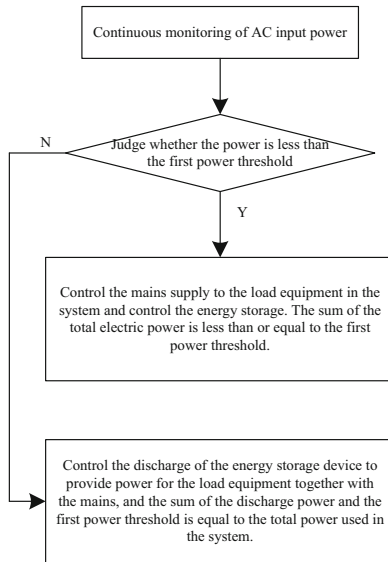
The user side dynamic capacity increase system introduced above makes full use of the high power frequent charging and discharging capability of the flywheel energy storage device and the features of the online dual conversion UPS device that can monitor the power of the AC input terminal in real time and conduct dynamic power distribution, so as to provide short-term high power support for the load equipment in the working state or when the multi load equipment is working. It achieves the effect of dynamic capacity increase without increasing the power supply capacity at the user side. Compared with the traditional power capacity increase method, it saves the process of reporting to the power supply department and waiting for approval. It has low equipment investment and short construction period.

#### 4 Control Method of Dynamic Capacity Increasing System

Figure 4 is the control method of the user side dynamic capacity increasing system as exemplified herein. A flow chart is shown in Fig. 4, which can support one or more of the following two control modes included in the described control method:

- (a) Charging mode: when the power of the AC input terminal is less than the first power threshold, control the mains power to supply power to the load equipment and other devices in the system, and control the mains power to charge the flywheel energy storage device. The sum of the charging power and the power consumption of the load equipment and other devices in the system is less than or equal to the first power threshold;
- (b) Discharge mode: when the power of the AC input terminal is greater than or equal to the first power threshold, control the flywheel energy storage device to discharge, and provide power for the load equipment and other devices in the system together with the mains, and the sum of the discharge power of the flywheel energy storage device and the first power threshold is equal to the sum of the power consumption of the load equipment and other devices in the system.

The method shown in the flow chart starts from the monitoring step, in which the power of the AC input is continuously monitored. In monitoring this power, it is constantly ensured that the appropriate power is supplied to the load equipment and other devices in the system. Then enter the judgment step to judge whether the power is less than the first power threshold. When the judgment result is yes, it is inferred that the power supply capacity of the mains can meet the power required by the load equipment and other devices in the system, and there is excess power to charge the flywheel energy storage device. The dynamic capacity increase system enters the charging mode. When the judgment result is no, it is inferred that the power supply capacity of the mains power cannot meet the power consumption required by the load equipment and other devices in the system, and the flywheel energy storage device is required to discharge. Together with the mains power, it supplies power to the load equipment and other devices in the system, and the dynamic capacity increase system enters the discharge mode. After the dynamic capacity increasing system enters the discharge mode, it also includes: judging whether the speed of the flywheel is lower than the first speed threshold.

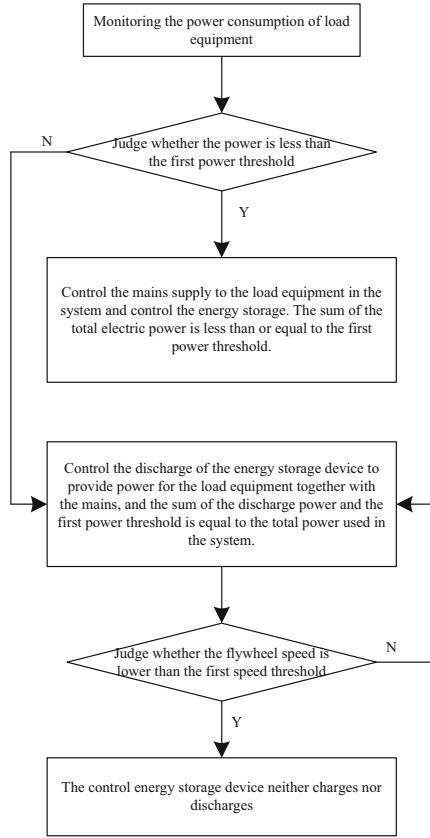


**Fig. 4.** Flow chart of control method for user side power dynamic capacity increase system

As shown in Fig. 5, on the basis of the flow chart shown in Fig. 4, enter the judgment step to judge whether the speed of the flywheel is lower than the first speed threshold. When the judgment step result is no, it is inferred that the flywheel energy storage device can continue to provide power, and the dynamic capacity increasing system continues to discharge; When the result of the judgment step is yes, the flywheel energy storage device is controlled to enter the standby state, neither charging nor discharging, and the dynamic capacity increasing system ends the discharging mode.

The control method provided in this paper compares the power of the AC input with the set first power threshold by monitoring the power of the AC input in real time, and then controls the working state of the flywheel energy storage device. It achieves the effect of dynamic capacity increase without increasing the power supply capacity on the user side. Compared with the traditional power capacity increase method, it saves the process of reporting to the power supply department and waiting for approval, and the equipment investment is low, The construction period is short. Based on the above adjustment mode, multiple sets of flywheel energy storage devices can be used to work in parallel as required to improve the output power of the flywheel energy storage system or extend the discharge time of the flywheel energy storage system. The system has good scalability.





**Fig. 5.** Flow of the optimal control method of the power dynamic capacity increase system on the user side

### 5 Conclusions

Compared with the existing technology, the user side power dynamic capacity expansion system and its control method provided in this paper have the following advantages and effects: the technical scheme provided in this paper can solve the problem that power users cannot use high-power electrical equipment due to the limitation of power supply capacity. Compared with the traditional power capacity increase method, the technical scheme provided in this paper does not need to be reported to the power supply department, which saves the application and approval process, low equipment investment and short construction period; The technical scheme provided in this paper uses online dual conversion uninterruptible power supply device to supply power to the load equipment, which can provide continuous, stable and uninterrupted power supply for the load equipment. The output waveform is good, the output voltage quality is high, and the load characteristics are good no matter when the load equipment is sudden or stable; The flywheel energy storage device used in this technical scheme can conduct frequent charging and discharging, provide high-power power supply support for load

equipment, and has the advantages of large charging and discharging power and long service life. According to the application requirements, multiple sets of flywheel energy storage devices can be used to work in parallel to improve the output power of the flywheel energy storage system or extend the discharge time of the flywheel energy storage system.

**Acknowledgement.** State Grid Hubei Electric Power 2022 Electric Power Capacity Increase Environmental Prediction and Multi scale Dynamic Capacity Increase Algorithm Research and Application Science and Technology Project.

## References

1. Keshavarzian, M., Priebe, C.H.: Sag and tension calculations for overhead transmission lines at high temperatures-modified ruling span method. *IEEE Trans. Power Delivery* **15**(2), 777–783 (2000)
2. Madadi, S., Mohammadi Ivatloo, B., Tohidi, S.: Dynamic line rating forecasting based on integrated factorized ornstein–uhlenbeck processes. *IEEE Trans. Power Delivery* **35**(2), 851–860 (2020)
3. Morozovska, K., Hilber, P.: Study of the monitoring systems for dynamic line rating. *Energy Procedia* **105**, 2557–2562 (2017)
4. Lecuna, R., et al.: Non-contact temperature measurement method for dynamic rating of overhead power lines. *Electr. Power Syst. Res.* **185**, 106392 (2020)
5. Dupin, R., Kariniotakis, G., Michiorri, A.: Overhead lines dynamic line rating based on probabilistic day-ahead forecasting and risk assessment. *Int. J. Electr. Power Energy Syst.* **110**, 565–578 (2019)
6. Kosec, G., Maksić, M., Djurica, V.: Dynamic thermal rating of power lines – model and measurements in rainy conditions. *Int. J. Electr. Power Energy syst.* **91**, 222–229 (2017)
7. Phillips, T., DeLeon, R., Senocak, I.: Dynamic rating of overhead transmission lines over complex terrain using a large-eddy simulation paradigm. *Renewable Energy* **108**, 380–389 (2017)
8. Divya shree, M., Sangeetha, A., Krishnan, P.: Analysis and optimization of uniform FBG structure for sensing and communication applications. *Photon. Netw. Commun.* **39**(3), 223–231 (2020). <https://doi.org/10.1007/s11107-020-00880-1>
9. Talpur, S., Lie, T.T., Zamora, R.: Application of dynamic thermal rating: overhead line critical spans identification under weather dependent optimized sensor placement. *Electr. Power Syst. Res.* **180**, 106125 (2020)
10. Meka, S., Fonseca, B.: Improving route selections in zigbee wireless sensor networks. *Sensors* **20**(1), 164 (2019)
11. Azumaya, C.M., et al.: Screening for AMPA receptor auxiliary subunit specific modulators. *PLoS ONE* **12**(3), e0174742 (2017)
12. Enescu, D., Colella, P., Russo, A.: Thermal assessment of power cables and impacts on cable current rating: an overview. *Energies* **13**(20), 5319 (2020)
13. Kottonau, D., Shabagin, E., Noe, M., Grohmann, S.: Opportunities for high-voltage AC superconducting cables as part of new long-distance transmission lines. *IEEE Trans. Appl. Supercond.* **27**(4), 1–5 (2017)
14. Jiang, J.A., Liang, Y.T., Chen, C.P., Zheng, X.Y., Chuang, C.L., Wang, C.H.: On dispatching line ampacities of power grids using weather-based conductor temperature forecasts. *IEEE Trans. Smart Grid* **9**(1), 406–415 (2018)

15. Theodosoglou, I., Chatziathanasiou, V., Papagiannakis, A., Wicek, B., De Mey, G.: Electrothermal analysis and temperature fluctuations' prediction of overhead power lines. *Int. J. Electr. Power Energy Syst.* **87**, 198–210 (2017)
16. Aznarte, J.L., Siebert, N.: Dynamic line rating using numerical weather predictions and machine learning: a case study. *IEEE Trans. Power Delivery* **32**(1), 335–343 (2017)
17. Tronconi, A.B., Bretas, A.S., Gazzana, D.S.: Two-dimensional transmission line modeling method: an algorithm considering non-homogeneous media and ionization. *Electr. Power Syst. Res.* **173**, 220–229 (2019)
18. Arroyo, A., Castro, P., Manana, M., Domingo, R., Laso, A.: CO2 footprint reduction and efficiency increase using the dynamic rate in overhead power lines connected to wind farms. *Appl. Therm. Eng.* **130**, 1156–1162 (2018)



# Secure and Non-secure Data Processing of Power Bus Communication

Guowei Zhu<sup>(✉)</sup>, Dangdang Dai, Jing Li, Jie Huang, and Shengwei Wang

State Grid Hubei Electric Power Company, Wuhan 430000, Hubei, China

minang15537fmd@163.com

**Abstract.** This paper studies the methods and devices of secure and non secure data transmission and isolation in power bus communication, which belongs to the field of power communication. In this paper, the power communication equipment is interconnected through the bus, and the safe and non safe data are transmitted on the bus; Identify safety and non safety data through bus arbitration unit, and conduct data diagnosis and data verification for complete and non safety data; Determine the safe activity area for operating safety data and the non safe activity area for operating non safety data through regional division; The access of non security tasks to security data is realized through the data sharing unit. This paper makes the integration of security controller and terminal more convenient, provides the possibility for non security functions to call security data, improves the utilization rate of security controller resources, improves the flexibility and scalability of monitoring system integration, reduces the cost of monitoring system, and has great application and promotion.

**Keywords:** Power · Communication · Data Processing

## 1 Introduction

In the power industry, the safety instrument system, as the main equipment of safety protection, has been widely used. Safety instrumented system is a control system used to realize one or more safety functions [1]. It can take emergency measures against the possible dangers of devices or equipment to minimize the dangers and losses and ensure the safety of production equipment, environment and personnel [2–7]. The general safety instrumented system consists of measuring instruments, safety controllers and actuators. The equipment is connected through the field bus [8]. The composition of the safety instrumented system conforms to IEC61511 standard and needs to meet different SIL requirements for different applications [9–12]. The general non safety system consists of measuring instruments, non safety controllers and actuators [13]. The equipment is also connected through the field bus. The composition of the non safety system has no SIL requirements [14].

In the power station, the plant level monitoring system includes the safety instrument system and the non safety system [15–18]. The safety instrument system performs the safety protection function, while the non safety system performs the basic process control

function [1]. The safety instrument system has SIL requirements, which is the essential difference between the two. The safety controller and the non safety controller perform safety protection and basic process control. Due to the limited processing capacity of the safety controller and the limited isolation capacity of the safety and non safety communication data in the early stage, the safety controller and the non safety controller are independent of each other and operate independently, without data interaction and associated logic control [2]. Therefore, they are configured independently, It has caused a great waste of resources.

With the continuous development of technology, the computing and processing capacity of the safety controller has been greatly increased. In this way, a fusion control method is proposed, in which both the safety protection and process control functions are placed in the safety controller. The safety controller has the functions of both the safety controller and the non safety controller. This method is different from the previous structure, in which the safety controller and the non safety controller remain independent, The fusion technology makes the controller application more flexible, and the integrated control system has a higher SIL level, while providing more customization advantages, improving resource utilization and reducing costs.

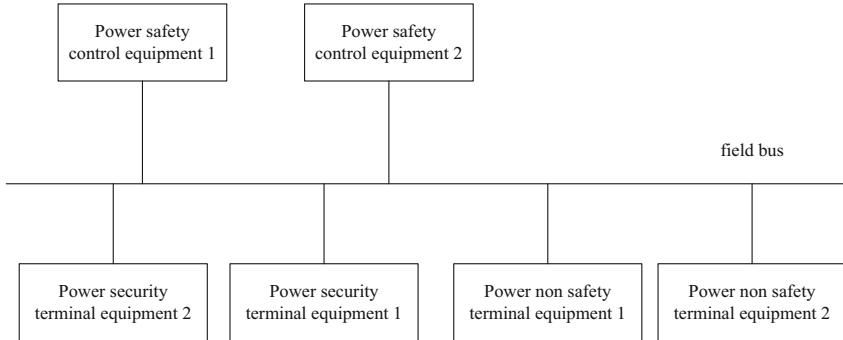
For the above fusion method, the security controller communicates with the terminal device in the form of bus; However, it should be noted that the communication between the security controller and the security terminal equipment conforms to the IEC-61784–3 standard, needs to meet the SIL requirements, and has certain information security measures against malicious attacks to ensure the integrity, confidentiality and timeliness of the security data; However, the communication between the security controller and the non security terminal equipment does not require the above indicators. The advantages of the safety controller are obvious, but there are also some problems. Because the safety and non safety measuring instruments and actuators are connected together through the field bus, the transmission process may lead to the impact of non safety data on the safety data. Therefore, how to transmit the safety and non safety data through the public field bus, and how to isolate the safety and non safety data during the transmission process, It is an important key to integrate safety and non safety functions into the safety controller.

The solution to the above problem is to find a full process data transmission and isolation method for power equipment from terminal equipment (security and non security) to security controller control tasks, to ensure that security data is not affected by non security data, and to meet the integrity, confidentiality and timeliness requirements for security data of power communication equipment. In view of the above shortcomings in the existing technology, the technical problem to be solved in this paper is to provide a method and device for safe and non safe data transmission and isolation of power bus communication, which is mainly used for safe and non safe data transmission and isolation of communication equipment on the public field bus.

## **2 Bus Structure for Safe and Non Safe Data Transmission**

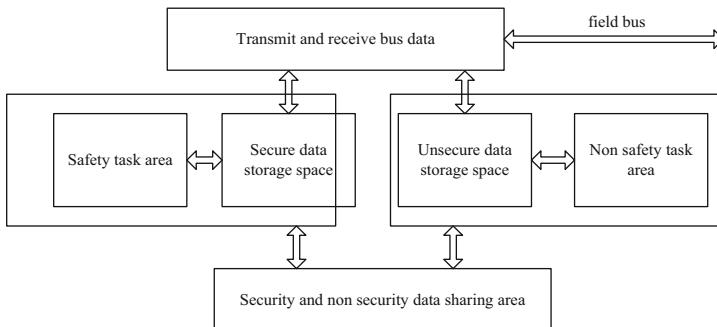
This paper implements a method and device for safe and non safe data transmission and isolation of power bus communication, as shown in Fig. 1. Security control equipment, security terminal equipment and non security terminal equipment are interconnected

through the bus; The security terminal equipment can process and transmit security data, the non security terminal equipment can process and transmit non security data, and the security control equipment can process and transmit two types of data; Safety data and non safety data are transmitted on the same field bus.



**Fig. 1.** Bus structure of safe and non safe data transmission

Figure 2 is the overall framework diagram of safe and non safe data transmission and isolation. Logically, safe work areas such as safe task execution and safe data storage are defined as safe activity areas, and safe work areas such as non safe task execution and non safe data storage are defined as non safe activity areas; The program running area, storage space, stack space, etc. in the security activity area and non security activity area are independent of each other to achieve strict physical space isolation.



**Fig. 2.** Overall framework of secure and non secure data transmission and isolation

Security activity area, non security activity area, security and non security data sharing unit, bus transceiver and arbitration unit constitute the basic framework of data transmission and isolation in this paper. Bus data can be processed and transmitted through bus data transceiver and arbitration unit, and security activity area, non security activity area, security and non security data sharing unit realize the physical isolation of security and non security data.

The bus data receiving/transmitting and arbitration unit can process the bus data and diagnose the received safety data. When the diagnosis is correct, the data will be stored as the safety data storage space in the safe activity area; Verify the received safety data. When the safety data verification result is correct, the received data will be written into the non safety data space in the non safety activity area; It is able to process the security policy of the sent security data, add the security identification and send it to the bus; It is able to add verification data to the non safety data sent, and send it to the bus after adding the safety identification.

Security and non security data can interact with each other through the data sharing unit, and non security data can access security data through padlocking, and cannot be modified. Security tasks are not allowed to access and modify security data.

### 3 Overall Framework of Data Transmission and Isolation

Figure 3 is the structure block diagram of the bus data transceiver and arbitration unit. When the bus communication device receives the bus data, it judges whether it is local data by ID or IP. When it is determined to be local data, it will respond to the security data processing preferentially through the priority response mechanism after being identified as security data through the identifier. The security data will be analyzed, and the data will be diagnosed after analysis. The diagnosis strategy complies with IEC61784-3 and meets the requirements of SIL. When the security data has requirements for information security, Information security policies such as secret key and connection authentication will be added; When the diagnosis result is successful according to the security policy, the received data will be written into the pre prepared security data storage area; When the safety data diagnosis result is failure, an alarm signal will be generated and the received data will be discarded;

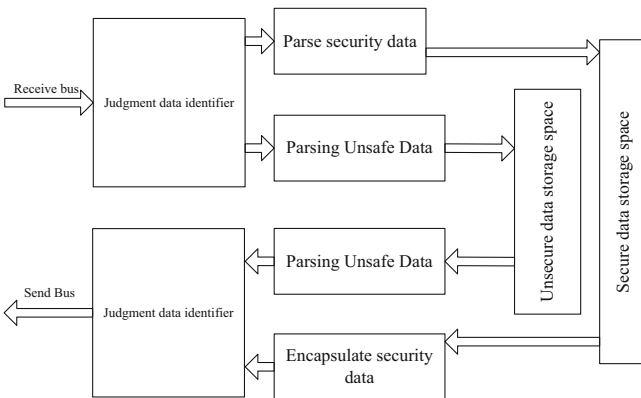


Fig. 3. Bus data transceiver and arbitration unit structure

When the bus communication device receives the bus data, it judges whether it is local data by ID or IP. When it is local data, it will check the integrity of the data after

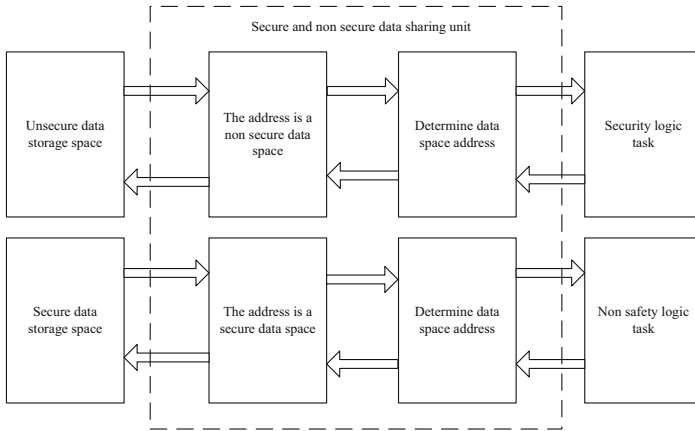
it is identified as unsafe data by identifier; When the security data verification result is correct, the received data will be written into the non security data storage area prepared in advance; When the security data verification result is error, give up the received data; When the bus communication device sends the security data to the bus, it first reads the security data, and processes it through the security policy. After the security policy processing, it encapsulates the data. Through the priority response mechanism of the security data, it adds the security identification and terminal address to send to the bus; When the bus communication equipment sends non safety data to the bus, it shall first generate the integrity check code, add the integrity check data, and then package it, and then send it to the bus after adding the safety identification.

## 4 Data Sharing and Data Transmission

Figure 4 is the structure block diagram of security and non security data sharing unit. By adding a security lock, the access control of security logical tasks and non security logical tasks to the memory area is realized. When the security logic task needs to access the non security storage space, first judge whether the access data space address is a non security data space. When it is judged to be a non security address space, padlock the read data and write data functions and do not allow access to the non security data storage space; When a non security logical task needs to access the security storage space, first judge whether the access data space address is a security data space. When it is determined to be a security address space, padlock the write data function. It is not allowed to write data to the non security data storage space. Open the read data lock and run the non security logical task to read the security data storage space data.

When the data arrives at the bus communication device through the bus, the bus communication judges whether it is local data through ID or IP. When it is determined to be local data, enter the data processing program, or discard the data; When the safety data is identified as safety data through the identifier, the priority response mechanism will give priority to respond to the safety data processing, unpack and analyze the safety data, and conduct the safety strategy diagnosis on the data after the analysis. When the safety diagnosis is passed, the received data will be written into the safety data storage area. When the safety diagnosis is wrong, an alarm signal will be issued; When identified as unsafe data by the identifier, the unsafe data is unpacked and parsed. After parsing, the data is verified. When the verification is correct, the received data will be written into the unsafe data storage area. When the security verification is wrong, the frame data will be discarded.



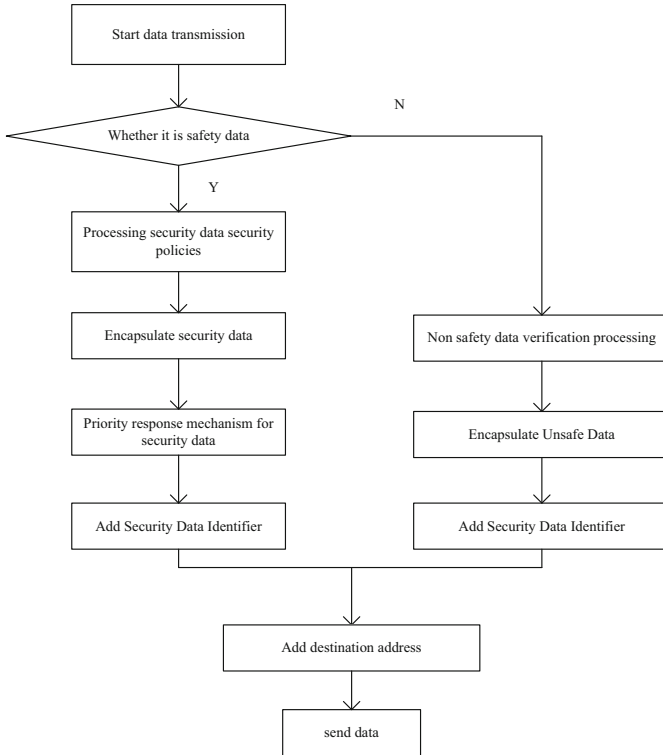


**Fig. 4.** Structure of security and non security data sharing unit

Figure 5 is the flow chart of bus safe and non safe sending and receiving. After receiving the data sending command, the bus communication device will judge whether the data to be sent is security data. When the data is security data, it will conduct security policy processing on the data. After the processing is completed, the data will be packaged according to the data format, and then the preferred processing mechanism of security data will be implemented. When there is non security data to be sent, it will process the security data in a limited way, add the security data identifier, and add the destination device address, Sending security data; When judging the non security data to be sent, the data will be processed for data verification. After the processing is completed, the data will be encapsulated according to the data format, and then the non security data identifier will be added, the destination device address will be added, and the non security data will be sent.

With the increasing demand for energy, in order to improve power generation efficiency and reduce emissions, the capacity and parameters of the main engine are getting higher and higher, but also increase its potential risk. For the above reasons, by adding the corresponding safety protection system, the occurrence of hazards can be fully reduced and the safety of thermal power plant operation can be improved.

The early power plant safety system required to collect data through hard wire. However, with the generation of functional safety bus, a large number of bus type sensors were also used in the safety system, providing great convenience for system integration, as well as the possibility of bus transmission of safety and non safety data. Through this paper, it can ensure the transmission of safe data and non safe data on the power transmission bus, and ensure the isolation of safe data and non safe data, which has a wide application value.



**Fig. 5.** Bus security and non security data transmission process

## 5 Conclusions

This paper is used to realize a method and device for safe and non safe data transmission and isolation of power bus communication, which guarantees the application of security controller. It has the following advantages: it is applicable to the integrated security control system and has high resource utilization. This paper breaks the traditional way of independent transmission of security and non security data, making the integration of security controller and terminal more convenient, providing the possibility for non security functions to call security data, and system resource utilization; High data security and good isolation performance. The communication equipment in this paper diagnoses the safety data, which improves the safety of data transmission and data itself; The communication equipment provides mutually independent areas for the storage and logical tasks of safe and non safe data to achieve regional isolation of safe and non safe data with good isolation performance; No special hardware is required. Based on the mature bus control method, this paper realizes the transmission and isolation of safe and non safe data through software, without additional hardware equipment participating in the control, which reduces the complexity of system implementation, reduces the cost of system construction, and has good usability and economy; Strong universality. With the increasing demand for integrated application of security and non security control

functions, without changing the original system bus structure, this technology can well solve the problems of coexistence, transmission and isolation of security and non security data, with strong applicability and broad application prospects.

**Acknowledgement.** State Grid Hubei Electric Power 2022 Power Communication Information Security Interaction Technology Research and Application Technology Project under the cloud edge collaboration framework.

## References

1. Motalleb, M., Siano, P., Ghorbani, R.: Networked stackelberg competition in a demand response market. *Appl. Energy* **239**, 680–691 (2019)
2. Andoni, M., et al.: Blockchain technology in the energy sector: a systematic review of challenges and opportunities. *Renew. Sustain. Energy Rev.* **100**, 143–174 (2019)
3. Ryu, U., Wang, J., Kim, T., Kwak, S., Juhyok, U.: Construction of traffic state vector using mutual information for short-term traffic flow prediction. *Transp. Res. Part C* **96**, 55–71 (2018)
4. Noor, S., Yang, W., Guo, M., van Dam, K.H., Wang, X.: Energy demand side management within micro-grid networks enhanced by blockchain. *Appl. Energy* **228**, 1385–1398 (2018)
5. Mengelkamp, E., Gärtner, J., Rock, K., Kessler, S., Orsini, L., Weinhardt, C.: Designing microgrid energy markets. *Appl. Energy* **210**, 870–880 (2018)
6. Coutinho, M., de Oliveira, A.R., Borges, F., Garcia Villalba, L.J., Kim, T.H.: Learning perfectly secure cryptography to protect communications with adversarial neural cryptography. *Sensors* **18**(5), 1306 (2018)
7. Shah, S.A., Issac, B.: Performance comparison of intrusion detection systems and application of machine learning to snort system. *Futur. Gener. Comput. Syst.* **80**, 157–170 (2018)
8. Vijayakumar, P., Chang, V., Deborah, L.J., Balusamy, B., Shynu, P.G.: Computationally efficient privacy preserving anonymous mutual and batch authentication schemes for vehicular ad hoc networks. *Futur. Gener. Comput. Syst.* **78**, 943–955 (2018)
9. Burg, A., Chattopadhyay, A., Lam, K.Y.: Wireless communication and security issues for cyber-physical systems and the Internet-of-Things. *Proc. IEEE* **106**(1), 38–60 (2018)
10. Gholizadeh, N., Saadatfar, H., Hanafi, N.: K-DBSCAN: An improved DBSCAN algorithm for big data. *J. Supercomput.* **77**, 6214–6235 (prep) (2020)
11. Scitovski, R., Sabo, K.: A combination of k -means and DBSCAN algorithm for solving the multiple generalized circle detection problem. *Adv. Data Anal. Classif.* **15**(1), 83–98 (prep) (2020)
12. Govender, P., Sivakumar, V.: Application of k -means and hierarchical clustering techniques for analysis of air pollution: a review (1980–2019). *Atmos. Pollut. Res.* **11**(1), 40–56 (2020)
13. Ghorbani, M., Dehmer, M., Zangi, S.: Graph operations based on using distance-based graph entropies. *Appl. Math. Comput.* **333**, 547–555 (2018)
14. Parisot, S., et al.: Disease prediction using graph convolutional networks: application to autism spectrum disorder and Alzheimer’s disease. *Med. Image Anal.* **48**, 117–130 (2018)
15. Singh, M.: Protection coordination in distribution systems with and without distributed energy resources- a review. *Protect. Control Mod. Power Syst.* **2**(1), 1–17 (2017). <https://doi.org/10.1186/s41601-017-0061-1>
16. Yoon, S., Lee, Y.J., Jung, H.J.: A comprehensive framework for seismic risk assessment of urban water transmission networks. *Int. J. Disaster Risk Reduction* **31**, 983–994 (2018)
17. Berrueta, A., Urtasun, A., Ursúa, A., Sanchis, P.: A comprehensive model for lithium-ion batteries: From the physical principles to an electrical model. *Energy* **144**, 286–300 (2018)

18. Hooshyar, D., Pedaste, M., Saks, K., Leijen, Ä., Bardone, E., Wang, M.: Open learner models in supporting self-regulated learning in higher education: a systematic literature review. *Comput. Educ.* **154**, 103878 (prep) (2020)



# Dynamic Monitoring Method for Power Big Data Quality

Ting Huang<sup>(✉)</sup>, Li Yuan, Yujia Zhang, Fei Xie, Zhengqin Zuo, Yuefeng Peng, Shengjie Yang, and Haige Ren

Tongren Power Supply Bureau of Guizhou Power Grid Co., Ltd., Tongren 554300, Guizhou, China

mukan396246106@163.com

**Abstract.** This paper is applicable to the computer field, and proposes a set of quality control technologies, methods and big data analysis and calculation tools for power calculation. The method is that when the actual results of the existing regulatory rule equations are invalid, the actual results are not eliminated, but the reasonable “invalid” real data results are selectively left when re evaluating. And keep these data results in the temporary database system. When the amount of “invalid” information in the temporary database system accumulates to a certain amount, these data results will be used as the initial resource to retrain the regulatory rule equation, so that the regulatory rule equation can continue to grow with the change of the actual number, improving the real-time and effectiveness of the regulatory rule equation. Thus, it can overcome the problem that the preset regulatory system without growth ability will lead to insufficient data of electric energy measurement in the database system, and the electric energy measurement information in the database system will become more “old”, which can not keep up with the reality and the statistical development after development.

**Keywords:** Power · Big Data · Monitoring

## 1 Introduction

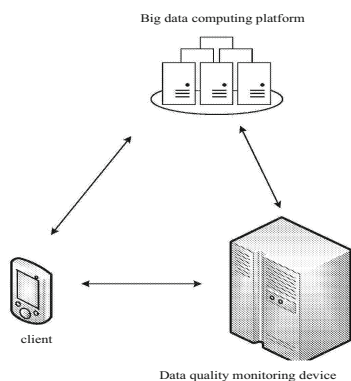
The key to big data analysis is to solve the problem of data analysis product quality, avoid errors in big data analysis, and ensure data quality, which can effectively enable SMEs to maximize the benefits in the big data analysis use process, while ensuring data quality is an important prerequisite for big data analysis to create value for SMEs. And big data analysis cannot be separated from data analysis quality and information management [1]. Through high-quality big data analysis services and efficient information management, whether in academic research or business use, the authenticity and practical value of big data analysis results can be ensured [2–7]. In recent years, great progress has also been made in the construction of Internet informatization in the power generation field in China, from the initial automation of power production to the establishment of corporate management informatization represented by the computerization of financial accounting in the early 1980s, to the construction of Internet informatization in a

large number of small and medium-sized enterprises in recent years, especially with the comprehensive establishment of the next batch of smart national power grids in China, The new IT information technology, represented by the Internet of Things and cloud computing capabilities, has been widely used in China's power generation industry [8]. Power generation data resources have also started to increase rapidly and have produced considerable scale [9].

With the continuous establishment and in-depth application of e-government in the power industry, various services of power companies have been initially combined with e-commerce [10–16]. The scale and type of service data in the information system are gradually expanding, and the requirements for resource sharing are outstanding. However, the quality of information service and the degree of resource sharing and utilization are not high [17–19]. First, the information is not enough to support management decisions. The same information has multiple numbers and multiple sources, and the data caliber is not unified; Second, the support of data enterprises for operation and management needs to be improved. The data quality is uneven. Some data enterprises have no operation system support, and there is no unified management standard, specification and strict accountability for data; Third, the data quality supervision system is backward and the management and control is one-sided. The failure to establish a unified and perfect data quality supervision system and a complete and reasonable data quality assurance mechanism has seriously affected the deep exploration of data resources.

The electric energy measurement information quality monitoring technology, methods and big data analysis and statistics tools proposed in this paper try to overcome the problem that the lack of growth of preset monitoring methods makes the electric energy measurement data sources in the information system insufficient, and the electric energy measurement information in the database will become more “old”, which can not keep up with the reality and the development trend of information after development.

## 2 Implementation Environment of Data Quality Monitoring Methods



**Fig. 1.** Schematic diagram of implementation environment of power data quality monitoring method

Figure 1 shows the implementation environment diagram of a power metering data quality monitoring method provided in this paper. For the convenience of illustration, only the parts related to this article are shown, such as servers, power metering data quality monitoring equipment, and big data analysis and statistics tools. Wireless network or cable TV connection is adopted between the client and the big data processing and measurement system of the electric energy measurement data quality monitoring system and the electric energy measurement data quality monitoring system.

After collecting the real information used to measure the electric energy, the power calculation data quality monitoring equipment detects the real number and sends a warning message to the server in case of abnormal real numbers; The server can select the processing process for the actual data corresponding to the warning information, and send the processing command to the power calculation data quality monitoring equipment; When the power metering data quality monitoring equipment runs the data processing command, it deletes and temporarily stores all the actual data, and changes the power metering data quality monitoring rule equation through the temporary actual information, thus causing the quality monitoring rule equation to be modified with the change of the actual data, and all processes and some column operations running in the power metering data quality monitoring equipment can be completed in the big data computing platform. In addition, all the actual information temporarily retained and a large amount of historical data saved can also be uploaded to the big data computing platform.

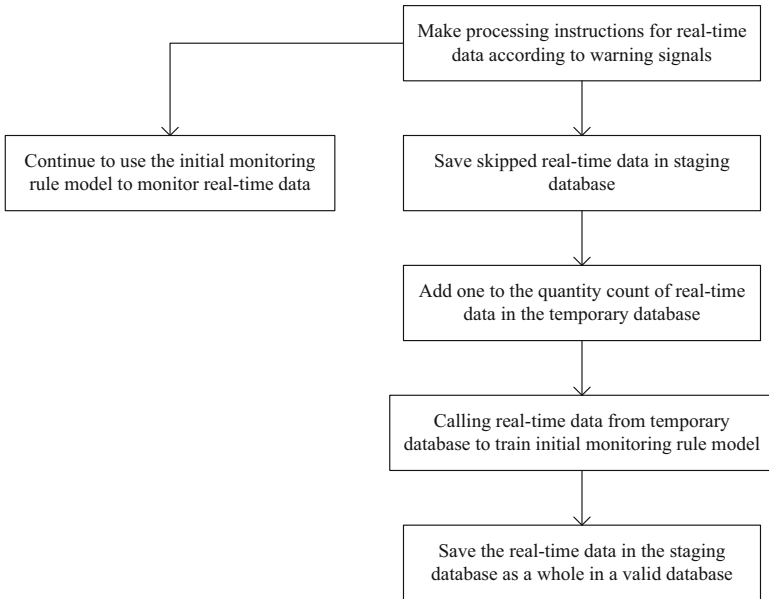
Its client allows the use of smart phones, tablet computers, portable computers and other personal electronic devices. Users can also set up and install all applications using multimedia technology. In addition, the server can also have the user registration function, and can complete various functions such as data sharing, business flow and collaborative work. In addition, the big data computing system can provide background services for applications running in the server. Users can have one or more, and any one should realize information interaction and data transmission with the electric energy measurement data quality monitoring system.

Among them, the electric energy metering data quality monitoring equipment is also connected with external internal data acquisition equipment while collecting field data, so as to obtain more effective and real-time data on the site. Depending on the application environment, the external data acquisition equipment may also be infrared detection equipment, intelligent video recognition equipment or sound acquisition equipment.

The big data analysis and calculation service platform can be either a local service platform built by the enterprise itself, or a shared cloud platform or a private cloud platform leased to other service enterprises. There is no specific provision here, but the service system used can support the background business of power metering data quality monitoring equipment and clients. In addition, security issues need to be considered.

### 3 Process of Instruction Classification Execution

Figure 2 shows the classified execution flow chart of a power metering data quality monitoring processing instruction in this paper. After receiving the processing instruction made by the client for real-time data according to the warning signal, the processing instruction is classified and discussed:



**Fig. 2.** Power metering data quality monitoring processing instruction classification execution flow chart

Method (A): After executing the command to delete real-time data, continue to use the initial monitoring rule equation to monitor the real-time information. When the command to delete the real-time information is selected, it indicates that the real-time information confirmation is unqualified information. The use of the environment and other key elements have not changed, the instant information has not changed significantly, and the alarm information generated is basically unqualified information. Therefore, the warning message can also be ignored, and the processing command of deducting real data can be directly executed. When the surrounding environment, actual use occasions and other relevant conditions change, because of the need for continuous alarm information in a short period of time, and the automatic evaluation of the reliability of the real data, when the real data becomes qualified data, you can directly execute or skip the processing command of the real data without deducting the real data.

When the real-time information is skipped during command execution, the initial control rule equation is modified to obtain the growth control rule differential equation. In the above example, “when the processing command is to skip the real-time data, modify the initial control criteria differential equation and obtain the growth control criteria differential equation” is expanded as follows:

Step (B): When the processing instruction is to skip the actual data, it means that the actual data at this time has been the valid training data, and because the actual data has been the real data after growth, the skipped real data must be retained in the temporary database and used as the valid training data.

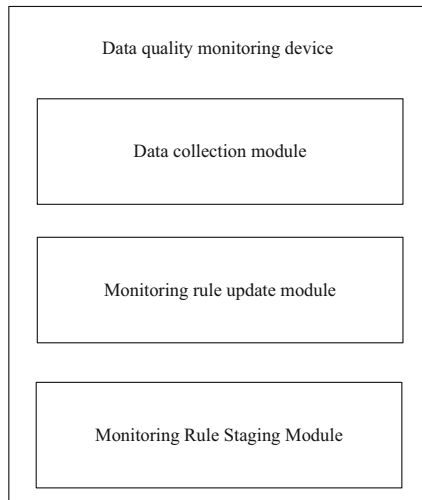
Step (C): add one to the measurement value of real-time data in the temporary database. Since the training monitoring rules have corresponding total amount provisions



for the required effective data, the total amount of all effective real-time data in the temporary database can be more accurately mastered when implementing this process, which is more convenient for the implementation of the next process.

Step (D): When the amount of information in the temporary database system reaches the threshold, call out the instant information in the temporary database system (the instant information is valid information after growing up) and train the initial control rule equation to obtain the control rule equation after growing up. The quality control of electric energy metering mainly includes:

Step (E): When all real-time information in the persistence database is effectively used, first store all real-time information in the persistence database in the effective database, and then remove all real-time information in the persistence database. Keep sufficient temporary storage system memory, which can adapt to the generation of the next actual result, and can also prevent the real-time results after multiple growth from being mixed with each other, and the call results are also easy to cause confusion.



**Fig. 3.** Structure of power metering data quality monitoring device

Figure 3 shows the structure diagram of a power metering data quality monitoring device in this paper, which is detailed as follows:

Information collection function: It is responsible for collecting instant information, evaluating it through the initial monitoring rule equation, and sending a warning message to the user when it is found that the instant information is incorrect.

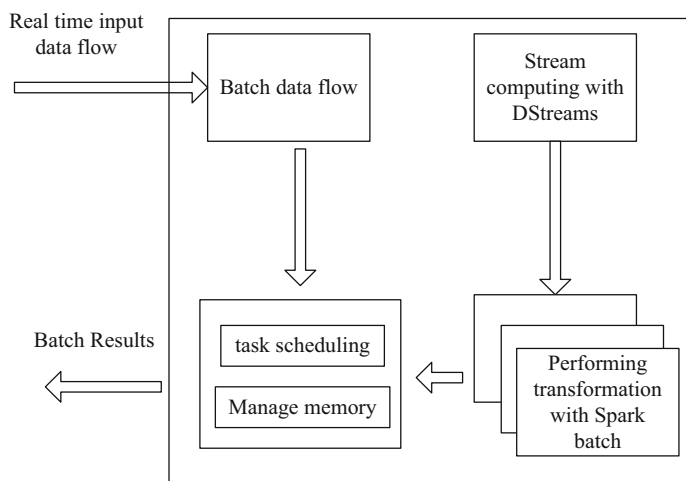
Monitoring criteria update module: used to receive data processing commands from the server to execute real-time data information based on warning information, and modify early monitoring criteria equations according to data processing commands to obtain mature monitoring criteria equations.

Temporary storage system of control methods: responsible for the control criteria process covered by the growth control criteria process, and monitoring the real-time results.

A big data analysis and calculation system includes one or more server devices, one or more data processors and one or more storage devices, and at least one computer program code is stored in each or more memories, Computer system programming is an algorithm that is actually completed by one or more information processors according to the load size and execution time limit.

#### 4 Real time Monitoring Method of Power Big Data Quality

As shown in Fig. 4, the processing process of a monitoring method for power big data processing quality in this paper includes the input of data streams in the business, the batching of data streams, the streaming operation through DStreams, the conversion through Spark batch processing, the task scheduling and memory allocation of Spark, and the output of batch processing results.



**Fig. 4.** Schematic diagram of real-time monitoring method for power big data quality

The key technical problem to be solved in this paper is to propose a real-time monitoring scheme for the quality of power big data to achieve the quality control of the whole process of power enterprise data, improve the service level of big data analysis, ensure the accuracy, timeliness, efficiency and credibility of big data mining, and provide a strong guarantee for the integration and mining application of big data mining.

To overcome these technical problems, a real-time monitoring method of power big data quality in this paper is divided into the following steps:

- (A) For the data flow system, first a topology is determined, then an interaction method between computing and frameworks is determined, and finally a computing entry architecture and a computing exit architecture are determined. Then the topology can combine various computing methods to provide the system with an overall function;

- (B) Batch the data flow, and load and start the topology. For a node, the topology must be loaded before starting. The node also needs other data, including upstream information sources and downstream information data; The topology data output by downstream data is saved in Tuple, while the topology itself is stateless; Through online update of topology, automatic update can be realized without stopping the service;
- (C) The streaming design of DStreams is used, and the data stream can be converted into a DAG diagram through a series of RDDs set by the user;
- (D) Use Spark batch execution transformation technology to transform DAG tasks into a TaskSet task set, and these TaskSets can also directly submit computing resources to the community, and then the community will deploy these TaskSets in the Worker for computing; First define RDDs, and then make corresponding conversion actions on RDDs. Finally, the system places this series of RDDs in Spark's work cluster for execution;
- (E) Spark framework task scheduling and memory management, receiving steps 2) and 4) to complete the calculation;
- (F) Output batch results.

## 5 Conclusions

The main beneficial effects of this paper are: the meta data model and data quality monitoring specification for power big financial data, using the real-time monitoring mode of power big data quality, using streaming data quality technology based on Streaming, to achieve the management mode and application strategy of real-time data quality monitoring without landing in the data process, and using the integrated memory database system in the big data process, And use the application analysis of rule base in each node to meet the function of real-time quality monitoring; Use the real-time big data information capture and synchronization skills to design a bypass data information channel in the main channel at the same time to support the need for real-time monitoring of big financial data quality. The real-time monitoring method of power big data quality based on streaming computing technology is used to complete the whole process quality control of power enterprise financial data around the data information life cycle of power enterprises, effectively improve the quality of data mining services, ensure the correctness, timeliness, efficiency and credibility of data mining, and provide a strong guarantee for the integration and mining application of big data mining.

## References

1. Zhang, M., Marculescu, B., Arcuri, A.: Resource and dependency based test case generation for RESTful web services. *Empir. Softw. Eng.* **26**(4), 1–61 (2021). <https://doi.org/10.1007/s10664-020-09937-1>
2. Abraham, R., Schneider, J., Vom Brocke, J.: Data governance: A conceptual framework, structured review, and research agenda. *Int. J. Inf. Manag.* **49**, 424–438 (C) (2019)
3. Huang, L., Zhao, Y., Mei, L., Wu, P., Zhao, Z., Mao, Y.: Structural holes in the multi-sided market: a market allocation structure analysis of china's car-hailing platform in the context of open innovation. *Sustainability* **11**(20), 5813 (2019)

4. Leonelli, S.: Data - from objects to assets. *Nature* **574**(7778), 317–320 (2019)
5. Ylijoki, O., Porras, J.: A recipe for big data value creation. *Bus. Process. Manag. J.* **25**(5), 1085–1100 (2019)
6. Yebenes, J., Zorrilla, M.: Towards a data governance framework for third generation platforms. *Procedia Comput. Sci.* **151**, 614–621 (2019)
7. Foster, J., McLeod, J., Nolin, J., Greifeneder, E.: Data work in context: value, risks, and governance. *J. Am. Soc. Inf. Sci.* **69**(12), 1414–1427 (2018)
8. Alexopoulos, K., Sipsas, K., Xanthakis, E., Makris, S., Mourtzis, D.: An industrial Internet of things based platform for context-aware information services in manufacturing. *Int. J. Comput. Integrat. Manu.* **31**(11), 1111–1123 (2018)
9. Benedikt, M., Grau, B.C., Kostylev, E.V.: Logical foundations of information disclosure in ontology-based data integration. *Artif. Intell.* (23) (2018)
10. Farkas, T.J.: Data created by the internet of things: the new gold without ownership. *La Propiedad In material* **23**, 5 (2017)
11. Uslnder, T.: Agile service-oriented analysis and design of industrial internet applications. *Procedia CIRP* **57**, 219–223 (2016)
12. Rowley, J.: The wisdom hierarchy: representations of the DIKW hierarchy. *J. Inf. Sci.* **33**(2), 163–180 (2017)
13. Halevy, A.Y.: Answering queries using views: a survey. *VLDB J.* 4 (2021)
14. Fagin, R., Kimelfeld, B., Kolaitis, P.G.: Probabilistic data exchange. *J. ACM (JACM)*. **58**(4), 1–55 (2021)
15. Bertino, E., Lobo, J., Brodie, C., Karat, C.M., Karat, J., Trombeta, A.: Privacy-aware role-based access control. *ACM Trans. Inf. Syst. Secur. (TISSEC)* (3) (2020)
16. Armbrust, M., et al.: A view of cloud computing. *Commun. ACM* **53**(4), 50–58 (2020)
17. Khatri, V., Brown, C.V.: Designing data governance. *Commun. ACM* (1) (2020)
18. Batini, C., Cappiello, C., Francalanci, C., Maurino, A.: Methodologies for data quality assessment and improvement. *ACM Comput. Surv. (CSUR)*. **41**(3), 1–52 (2019)
19. Bizer, C., Heath, T., Berners-Lee, T.: Linked data - the story so far. *Int. J. Seman. Web Inf. Syst. (IJSWIS)* (3), 205–227 (2019)



# Detection of Load Loss of Large Capacity Transformer

Feifeng Wang<sup>(✉)</sup>, Shengchao Jiang, Haoze Zhuo, Tailin Li, Bin Wang,  
and Yunqing Pei

Electric Power Research Institute of Guangxi Power Grid Co., Ltd., Nanning 530023, Guangxi,  
China

wang\_ff.sy@gx.csg.cn

**Abstract.** This paper mainly studies a live monitoring system and scheme for the current loss of power transformer. In the system, the inlet end of the primary side voltage transformer is partially connected with the high temperature side and low temperature side of the power transformer, and its output end is connected with the isolation transformer; The inlet end of the primary side voltage transformer is connected with the terminals of the high temperature side and low temperature side of the power transformer through the jaw, and the output end is connected with the voltage transformers of the two secondary sides; The output end of the isolation transformer is connected with the inlet end of the voltage transformer at the secondary side; The output end of the voltage transformer at the secondary side is connected with the input end of the optocoupler separation module; The output end of the optocoupler module and the output end of the secondary side voltage transformer are connected with the inlet end of the signal acquisition control system, and the output end of the signal acquisition control system is connected with the upper computer control system; The upper computer system is mainly used to extract and analyze the collected pressure and current information, and finally obtains the no-load loss and load loss of power transformer. Therefore, the control system in this paper can accurately evaluate the loss of power transformer, and has the advantages of on-site measurement and live measurement, with high measurement accuracy.

**Keywords:** Transformer · Load Loss · Detection · Power System

## 1 Introduction

Power transformers are electrical appliances with variable exchange voltage input forms.[1] By using the basic principle of electrical sensing, a certain level of AC voltage is changed into a certain level of exchange voltage input to meet the different requirements of different power supply devices [2]. Therefore, power transformers play an important role in power supply business [3]. With the development of power supply, the total amount of electrical equipment in transformer system is also increasing, from low voltage and low capacitor to the current high pressure and high power capacity [4]. Now, the voltage level and volume of transformer exceed the ten thousand volt level and ten

thousand kilovolt ampere level [5]. It has become an indispensable power transmission and distribution device in the work of the power supply system, which undertakes the work of power transmission and pressure and current transformation of the entire power control system [6]. Its function will directly affect the stability, safety, economy and effective work of the power supply system [7].

The energy consumption of power transformer can be divided into no-load energy consumption and load bearing energy consumption [8]. No load energy consumption refers to the damage to the core material of power transformer under the no-load condition of power transformer, which is mainly caused by hysteresis loss and eddy current energy consumption, and is related to voltage, time, core material and other conditions [9]. Load bearing energy consumption refers to the damage caused by the internal power transformer under the condition of bearing, which is mainly related to the weight form [10]. The internal energy consumption of power transformer is also the main reason to determine the operating temperature of power transformer [11]. When the internal energy consumption of power transformer is too high, its ambient temperature will gradually increase, which will weaken the security of power transformer, thus affecting the normal use and operation life of power transformer. However, when the ambient temperature rises, the no-load consumption and load cost of power transformers will increase or continue to increase, thus generating feedback to promote the continuous increase of ambient temperature, thus speeding up the aging of the insulation structure of power transformers, interfering with the normal operation and management of power transformers, thus greatly shortening the service life of power transformers [12]. At present, due to the increasing shortage of funds in our country, the problem of power saving has become an important issue concerned and discussed by the whole society. Therefore, it is of great value to study the low-noise driving effects of power transformers for saving power, optimizing environmental protection, and assisting in reducing consumption and energy saving of power supply systems.

At present, the measurement method for power transformer consumption usually adopts offline testing, that is, first disconnect the power transformer from the equipment side, then apply a certain weight voltage to its low-temperature side, open the pressure side to detect the no-load consumption of the power transformer, apply a certain voltage to its pressure side and short-circuit the low-temperature side to detect the maximum capacity consumption of the power transformer. Due to the tedious operation of the offline test mode, and the need to cut off the power transformer when it is put into operation, the test can only be carried out when it is light and the self provided power transformer can carry out the corresponding flow. The results have certain limitations. Therefore, the live test of no-load loss and load loss of power transformer is put forward, and the no-load loss and load loss of power transformer can be deduced by separately detecting the flow and pressure of high temperature side and low temperature side of power transformer. However, due to the influence of the detection accuracy of the detection instruments at one side and the secondary side, and the system error between the photoelectric elements in the detection process, the calculation of the no-load loss and equipment cost of the power transformer has some influence.

The main purpose of this paper is to propose a live measurement system of power transformer loss and solve these technical problems. This paper is mainly applied to the

live test of no-load loss and high load loss of high voltage power transformers, which can accurately evaluate the consumption of power transformers. It has the characteristics of real-time measurement and live test, and has high measurement accuracy.

## 2 Principle of Live Detection System

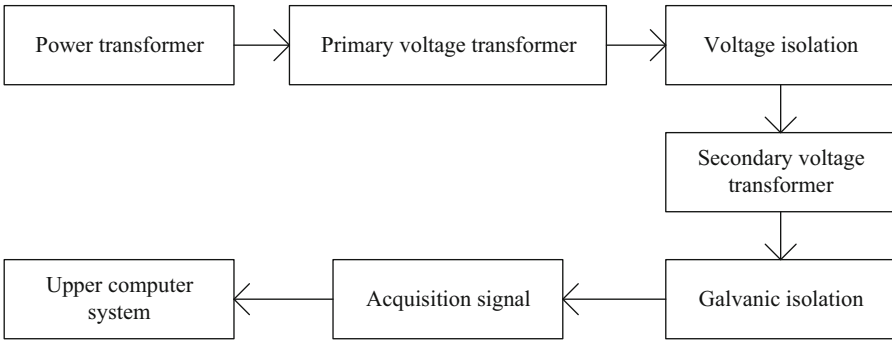


Fig. 1. Principle block diagram of live detection system for transformer loss

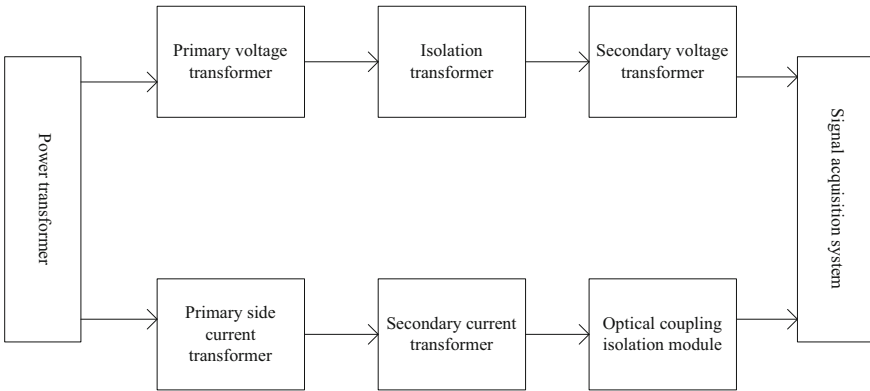


Fig. 2. Connection diagram of live detection system for power transformer loss

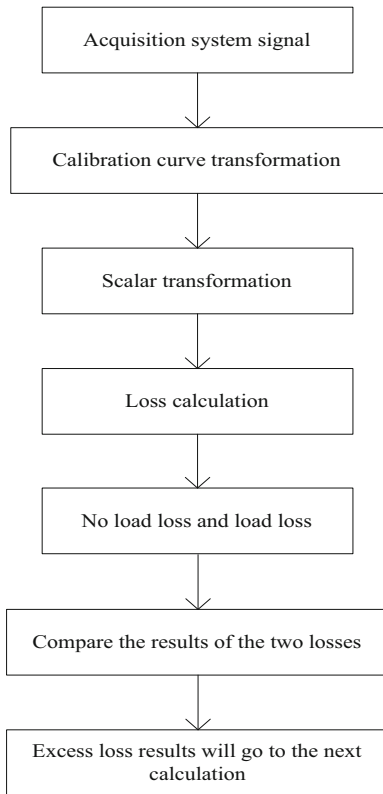
According to Fig. 1 and Fig. 2, a loss and live metering controller for power transformer includes a side voltage transformer, a side current transformer, an isolation transformer, a secondary current transformer, two side pressure mutual sensors, an optocoupler decentralized controller, a data acquisition controller and an upper computer controller.

After the inlet end of the voltage transformer at each side is connected with the high and low heat sides of the power transformer, the input and output ends are connected with the isolation transformer equipment; After the inlet end of the voltage transformer at each side is connected with the equipment at the high or low heat side of the power

transformer through the jaw, its input and output ends are connected with the voltage transformer at the secondary side; The input and output ends of the isolation transformer system are connected with the inlet end of the current transformer at the secondary side; The outlet end of the secondary current transformer is connected with the inlet end of the optocoupler separation module; The input and output terminals of the optocoupler unit and the current transformer at the secondary side are connected with the input and output terminals of the data acquisition system. The input and output terminals of the data acquisition system send the test data to the upper computer for the system, and calculate according to the pressure and flow values obtained from the test. From this, the no-load loss and load loss of the power transformer are obtained.

### 3 Detection Process

As shown in Fig. 3, during the test, the information measuring equipment successively switched the voltage and flow sequence through a voltage transformer at one side, a current transformer at the primary side, a voltage transformer at the secondary side and a current transformer at the secondary side, so the air pressure value and current value of



**Fig. 3.** Detection method flow chart



the last information measuring equipment had a corresponding deviation. Therefore, we can reduce the economic loss caused by this parameter through the current calibration curve. That is, when using this inspection equipment to test the no-load loss and load loss of power transformer, the benchmark instrument can be used to judge the water pressure and output test results of the test device, and draw the pressure calibration curve and current calibration curve.

#### 4 Detection of Abnormal Loss

Generally speaking, the load loss of power transformer consists of DC resistance loss, eddy current loss and stray loss. Among them, stray loss component accounts for about 10% of the total load loss. Although the proportion is small, with the increase of transformer capacity, its absolute value may be large. Take the power transformer with a three-phase capacity of 780 MVA as an example, when no effective measures are taken, its stray loss can be as high as about 200 kW, which not only causes a great waste of energy, but also may cause local overheating of internal structural parts of the transformer (such as iron core clamps).

The essence of transformer stray loss is the linkage between magnetic leakage and metal structure, which is mainly distributed in transformer oil tank and iron core clamp. The traditional method to reduce the stray loss of transformer is to lay magnetic shield or copper shield inside the transformer tank to reduce the stray loss of the tank structure. As for iron core clamps, because the magnetic leakage of transformer is still connected with most of them, the stray loss caused by it can not be ignored, which is an important factor leading to the large load loss of transformer. In order to effectively reduce the stray loss on the core clamp of the transformer, the flow path of the transformer magnetic leakage should be changed so that the magnetic leakage and the clamp are only partially interlocked.

The problem to be solved is that the stray loss of existing large capacity transformers is relatively high, resulting in energy waste. In order to solve the above problems, an oil immersed power transformer body magnetic shielding device can be used, which is characterized in that it includes an insulating shell composed of a matched insulating base and an insulating cover plate; Two ends of the insulating base are respectively provided with a silicon steel sheet, and the outer edge of the silicon steel sheet is provided with a shielding ring.

The insulating base is made of laminated cardboard or laminated wood; Silicon steel sheet is of sector structure. The insulating cover plate is bonded to the silicon steel sheet.

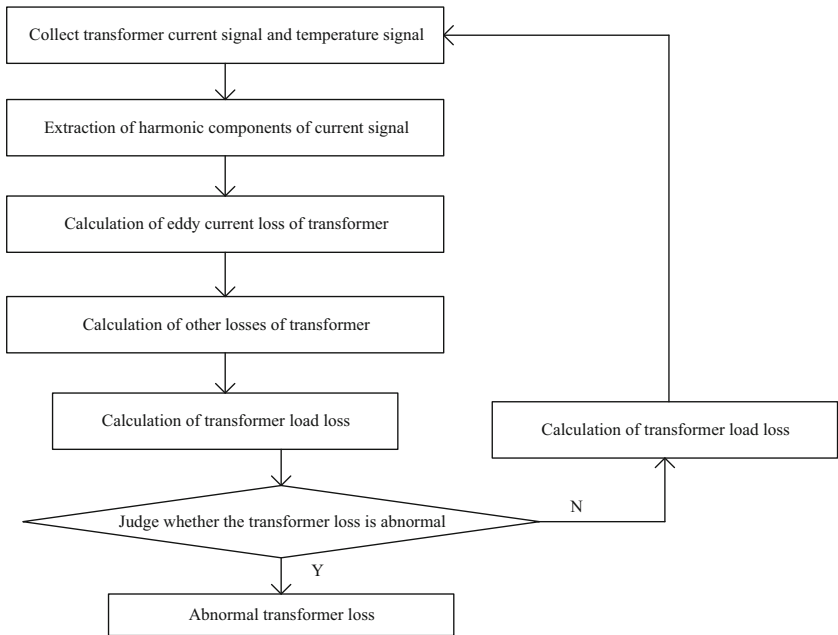
This paper also provides a method to effectively reduce the load loss of large capacity transformers, which is characterized by the use of oil immersed power transformer body magnetic shielding device, which is installed on the upper secondary pressing plate and lower supporting plate of the transformer winding to change the direction of magnetic flux leakage flow path. The magnetic circuit makes most of the concentrated distribution and a small part of the divergent distribution through high permeability components, thus effectively reducing the magnetic flux leakage, Solve the problem of local heating of iron core and minimize stray loss.

Since the magnetic shield of the transformer body is mainly composed of silicon steel sheets with high magnetic permeability, its magnetic conductivity is about 10000

times that of transformer oil. Therefore, most of the magnetic leakage at the end of the coil will be introduced into the silicon steel sheets of the transformer core through the silicon steel sheets in the magnetic shield of the transformer body, which fundamentally reduces the linkage between the magnetic leakage and the clamps, thus reducing the stray losses on the clamps.

The beneficial effect of this paper is that the stray loss of the iron core clamp of the transformer is greatly reduced by setting the magnetic shield of the transformer body in the auxiliary pressing plate and supporting plate at both ends of the transformer coil, which not only avoids the local overheating of the clamp, improves the reliability of the transformer, but also effectively reduces the load loss of the transformer.

Affected by various factors, the working loss of transformer includes other losses besides load loss. If the load loss is simply calculated, there will be error judgment when judging whether the transformer loss is abnormal. In this paper, a method of transformer loss anomaly detection is proposed. First, the harmonic component of current signal is extracted, and then the actual load loss is calculated according to each harmonic component, so as to judge whether the transformer loss is abnormal and improve the accuracy of detection results. This paper provides a transformer loss anomaly detection method, which can more accurately judge whether the transformer loss is abnormal.



**Fig. 4.** Workflow of Transformer Loss Anomaly Detection Method

The work flow of the transformer loss anomaly detection method is shown in Fig. 4. The transformer loss anomaly detection method includes the following steps:

Step (a): Collect the current signal and temperature signal of the transformer;  
Step (b): extract the harmonic component of the current signal: iteratively convert the current signal into equal interval sampling points;  
Step (c): Calculate the eddy current loss of the transformer  
Step (d): calculate other losses of transformer  
Step (e): calculate the load loss of the transformer  
Step (f): judge whether the load loss of the transformer is greater than the reference value of the load loss, if so, the transformer loss is abnormal; If not, the loss is within the normal range, return to step (a).

## 5 Conclusions

Compared with the existing technology, the technology in this paper has the following advantages: (a) The measurement is real-time. This system can monitor the no-load consumption and load loss of power transformer in real time, which plays the purpose of monitoring the consumption of power transformer in real time, and provides the corresponding basis for judging the internal temperature, working condition and life prediction of power transformer. (b). The range of measuring objects is very wide. The system can also be changed into primary side and secondary side current transformers and current transformers for power transformers of various current levels. It can measure the loss of power transformers of various voltage levels, types and capacities only after completing the standard design. Therefore, it has a wide range of applications. (c). The detection system has high security. The primary side and secondary side current transformers of the control system are shielded through isolation transformers, while the secondary side current transformers are shielded with the information acquisition system through optocoupler isolation modules, thus ensuring that the information acquisition system, upper computer and operators are always isolated from the primary side voltage and current, thus improving the safety performance of the data measurement system. (d). No load loss and heavy load loss measured by per unit value. The no-load consumption and load loss of power transformer are calculated by the per unit value, which is convenient for comparison of various types of power transformers and judgment of the operation of power transformers, and enhances the practicability of this device.

## References

1. Doostan, M., Chowdhury, B.H.: Power distribution system fault cause analysis by using association rule mining. *Electr. Power Syst. Res.* **152**(nov), 140–147 (2017)
2. Wang, A.L., Chen, B.X., Wang, C.G., Hua, D.: Non-intrusive load monitoring algorithm based on features of V–I trajectory. *Electr. Power Syst. Res.* **157**(APR), 134–144 (2018)
3. Kang, J.S., et al.: Development of a systematic, self-consistent algorithm for the K-DEMO steady-state operation scenario. *Nucl. Fusion* **57**(12), 408–414 (2017)
4. Al-rimy, B.A.S., Maarof, M.A., Shaid, S.Z.M.: Ransomware threat success factors, taxonomy, and countermeasures: a survey and research directions. *Comput. Secur.* **74** (2018)
5. Deshmukh, S., Troia, F.D., Stamp, M.: Vigenère scores for malware detection. *J. Comput. Virol. Hack. Tech.* **14**(2), 157–165 (2017). <https://doi.org/10.1007/s11416-017-0300-z>

6. Coutinho, M., de Oliveira Albuquerque, R., Borges, F., García Villalba, L.J., Kim, T.H.: Learning perfectly secure cryptography to protect communications with adversarial neural cryptography. *Sensors* **18**(5), 1306 (2018)
7. Shah, S.A.R., Issac, B.: Performance comparison of intrusion detection systems and application of machine learning to Snort system. *Futur. Gener. Comput. Syst.* **80**, 157–170 (2018)
8. Vijayakumar, P., Chang, V., Deborah, L.J., Balusamy, B., Shynu, P.G.: Computationally efficient privacy preserving anonymous mutual and batch authentication schemes for vehicular ad hoc networks. *Future Gen. Comput. Syst.* **78**(Pt.3), 943–955 2018
9. Burg, A., Chattopadhyay, A., Lam, K.Y.: Wireless communication and security issues for cyber-physical systems and the Internet-of-Things. *Proc. IEEE* **106**(1), 38–60 (2018)
10. Gholizadeh, N., Saadatfar, H., Hanafi, N.: K-DBSCAN: an improved DBSCAN algorithm for big data. *J. Supercomput.* **77**, 6214–6235 (2020) (prep)
11. Scitovski, R., Sabo, K.: A combination of k -means and DBSCAN algorithm for solving the multiple generalized circle detection problems. *Adv. Data Anal. Classif.* (2020) (prep)
12. Govender, P., Sivakumar, V.: Application of k-means and hierarchical clustering techniques for analysis of air pollution: a review (1980–2019). *Atmos. Pollut. Res.* **11**(1), 40–56 (2020)



# Intelligent Power Measurement System

Chengwu Zou<sup>1</sup>(✉), Dan Ye<sup>2</sup>, Qing Shen<sup>2</sup>, Naiqing Wang<sup>2</sup>, Jiaqi Lu<sup>2</sup>,  
Mingquan Zeng<sup>1</sup>, and Hangkang Li<sup>1</sup>

<sup>1</sup> State Grid Zhejiang Electric Power Co., Ltd., Research Institute, Hangzhou 310000, Zhejiang, China

songpantu99763@163.com

<sup>2</sup> E.Energy Technology Co., Ltd., Hangzhou 310000, Zhejiang, China

**Abstract.** This paper belongs to the field of power detection technology. It mainly studies an intelligent electrical engineering measurement system, which mainly includes: solar power module, command input and output module, parameter setting module, single-chip measurement and control module, on-site numerical simulation module, optical detection module, fault detection module, display module. The invention can obtain cleaner and more effective sunlight through the solar power module, thereby saving energy, being more economical and environmentally friendly, being able to maintain power, and effectively preventing interruption in the electrical engineering measurement process; And it is more comprehensive than the measurement technology through the traditional fault detection module, which improves the effect of troubleshooting.

**Keywords:** Intelligent · Power · System · Single Chip Microcomputer · Circuit

## 1 Introduction

Electric power is a general term for the manufacturing, transmission, distribution and reuse of energy, as well as the production of electrical equipment and other professional fields and engineering technology [1]. It is a kind of science and technology to provide, protect and change limited space and natural environment by means of energy theory, equipment and electrical science and technology, mainly including energy conversion, use and research, and also involving theory, application technology, equipment and device production [2–9]. However, the current electrical engineering measurement system uses the traditional power supply method to consume energy, and cannot continue to work if the power is cut off; And the fault measurement signal is monotonous, which makes it difficult to eliminate the fault.

At present, time division multiplexing technology has the following disadvantages: since time slots pre arrange various resource information, even if there is a time slot, the source information, rather than the transmission of information, must also be arranged to it, which inevitably leads to the loss of spectrum space [10–17]. Because each time slot is assigned to a fixed input circuit, when all input or output circuits do not fully enter the working state, an idle time slot will be generated, which makes the time transfer efficiency of traditional time division or multiplexing very low.

The frequency division reuse technology has the following defects: in order to ensure that the data transmission signals between the sub channels do not affect each other in the signal transmission process, the time division multiplexing technology usually requires that a corresponding protection frequency interval be set between the sub channels, while the frequency division multiplexing technology requires that a corresponding protection frequency band be equipped, thus greatly reducing the utilization of the frequency band [18].

The code division multiplexing method has the following disadvantages: The code division multiplexing method usually receives the influence of the code word length and the available code word amount. Therefore, no matter what code set is used, the multiple access problem will inevitably occur. At present, the traditional wireless communication mode in China is facing the problems of inefficient utilization of the volume and frequency resources of wireless network channels, inability to reduce multiple access interference, and poor effectiveness of signal propagation.

When the current FPGA is powered on, it can use special ports or peripheral EPCS elements to read out the FPGA setting signal, thus realizing the FPGA setting. Therefore, when you want to modify the logic structure of FPGA, you must burn the external EPCS again, and the configuration process can be realized. In addition, an EPCS chip with sufficient capacity only retains the internal FPGA setting information once.

At present, in the numerical simulation, because the effect of the computer network is not obvious enough, there is no large-scale quantitative data, and it is also unable to fully and truly depict various interaction characteristics, expressing uncertainty.

To sum up, the existing technology has the following problems:

The existing electrical engineering detection system uses the traditional power supply method to consume energy, so if the power is cut off, it will be unable to continue to work; At the same time, it is difficult to troubleshoot because of the simple fault detection information.

In traditional wireless communication technology, there are large volume of wireless channel, low utilization efficiency of signal resources, inability to prevent multiple access interference, and low efficiency of data transmission.

Nowadays, FPGAs have little ability to store information. It requires several chips to store configuration signals of several FPGAs at the same time; At the same time, the information configuration must also be managed by power failure, which is quite cumbersome to use.

## 2 Configuration of Intelligent System

The intelligent electrical engineering instrument control system includes: a parameter setting module, which is connected with the master control system of the single chip computer to initialize and select all the parameters of the single chip computer; The parameter setting module constitutes the setting loop; The configuration circuit includes serial FLASH memory; FCLK.FNCS of serial FLASH memory FMOSKFMIS port links to external data processor; Use these ports to write the configured data to the serial FLASH memory; Interface CLK SDA link external data processing; Boot data information using other connectors; NConfig, DI, DCLK, nCS, DO companies use wires

to connect FPGAs; The information input mode of serial FLASH memory is: the external processor uses FCLK, FNCS, FMOSKFMIS to input configuration data.

The drive steps of the configuration circuit are as follows: When the nConfig information is valid, the FPGA will reconfigure; Due to the input and output of FPGA and the combination of DLDCLK and nCS, when the sequence detector monitors the specified sequence data, the boot data buffer system generated by selecting a module through the output of the sequence detector; When no special sequence signal is detected by the sequence detector, select a module to output DI directly from FPGA; When the external processor's CLK After SDA realizes the use of guidance information, it will generate nConfig information and write the guidance information to the guidance information cache system. When the nConfig information is correct, it will generate FPGA configuration information; If the external processor fails to pass CLK The operation mode of SDA, that is, the content stored in the boot data cache module, is the same as the boot content output by DI by default.

The microcontroller controller module is connected with the parameter configuration module, the real-time numerical simulation module and the fault detection module to ensure the normal operation of the dispatching parameter setting module, the real-time numerical simulation module and the fault detection module; In the real-time numerical simulation module, it is connected with the microcomputer controller module to build the AC/DC transmission system RTDS simulation module through the RTDS real-time numerical simulator; In the AC/DC transmission system RTDS simulation module constructed by RTDS real-time numerical simulator, the AC/DC transmission system RTDS simulation module constructed by RTDS real-time numerical simulator and network computing technology; The contents include: from the perspective of cloud simulation, describing and studying the actual behavior and uncertainty in computing space of network design individuals and network design individuals; From the perspective of statistical space, the activities of network computing individuals and network computing individuals and their roles in the structure of statistical space are described and studied.

The dynamic mechanism of computer network and the dynamic mechanism of network are analyzed by simulation modeling and visualization; The fault detection module is connected with the microcomputer detection system to detect power line faults; The optical measurement module, connected with the microcomputer detection system, is used to connect the fault recorder vertically with the help of the optical cable, and monitor the corresponding electrical variable ratio with the fault recorder; When the optical measurement module is connected to the fault recorder, the optical information transmission mode is divided into: sending n-channel source signal, receiving the information of aliasing phenomenon by the receiver integrated with the optical measurement module, and separating the multi-channel aliasing information by the separation system integrated with the optical measurement module; Sending n-channel source information refers to combining the n-channel source signal with the channel information, taking the hybrid system as A, and sending it from the air by the sending terminal through n antenna systems; The transmitting end integrated with the optical measurement system receives the aliasing information, which means that the transmitting end receives the aliasing information through m reflective antennas, and the received information is

called detection information. The transmitting end needs to complete the pre-processing of detection information. The two parts of pre-processing include centralized processing and spherical processing; The optical measurement module is used to separate from the integrated discrete system. The multi-channel aliasing phenomenon information means that the discrete system  $W$  will separate the multi-channel aliasing phenomenon information from the entropy domain according to the difference of the information beam value in each source signal. The criterion of the information beam value is negative; Establish personal interaction field of network computing; The concept of field in physics is introduced into the design of number field space, and the overall number field space design forms a number of fields; Taking the individual role in the network system operation as the research goal of the overall number field space design, the data field is used to describe the local interaction between all objects, and the digital field is used to reflect the network system operation space structure as the topological potential value space structure of the arrival theory, describing and representing the network system operation and the synchronous interaction phenomenon of the network system; Synchronous driving and visual simulation of network system operation; The synchronization problem is regarded as the field evolution, and the visualization effect can be expressed by equipotential line method and equipotential surface method; The process of or synchronization is regarded as the evolution of image color, and the visualization effect is represented by the directed plane image method and the stereo image method.

The synchronous dynamics efficiency optimization analysis method of network operation; The synchronization problem of network computing is regarded as a plane or two-dimensional gray image, and the image beam is used to measure the computing intensity of network system computing individuals. The image beam is a data form representing a characteristic, reflecting the size of the average amount of information in the graph; Input and preprocess the individual characteristics and behavior characteristics of network system operation. The generation method of input data includes the generation method of typical representative data matrix of electrical engineering information collected, the generation method of typical representative data matrix customized by professionals, and the generation method of data matrix of cloud model representative of expert experience.

The software of the intelligent electrical engineering measurement system also includes: the solar power supply module, which is connected with the microcontroller control module to supply power to all working modules through the solar panel; The command input and output module is connected with the microcontroller control module to input or export test commands; The display module is connected with the measurement and control module of the SCM to represent the test data signal. The measurement method of the fault measurement module is as follows:

First of all, the real data simulation system is used to simulate the metal connection problem of DC circuit, the connection problem of high-voltage DC bus, the valve short circuit problem, the valve connection problem, the single grounding problem of converter bus, the interleaving problem of converter bus, and the three-phase fault of converter bus.



Then, by analyzing the offline wave recording at the wave recording workstation, it can be determined whether the same DC voltage and DC current passing through the optical quantity module and DC measuring equipment, and the same electrical quantity directly entering the wave recording equipment through the analog quantity output board card, are the signal amplitude consistent or the amplitude deviation lower than zero. The phase of the moon is consistent. If not all conditions are met, it is considered that the test system does not meet the requirements. If the relevant parameters are met, the DC measurement system can meet the detection conditions of the DC project site.

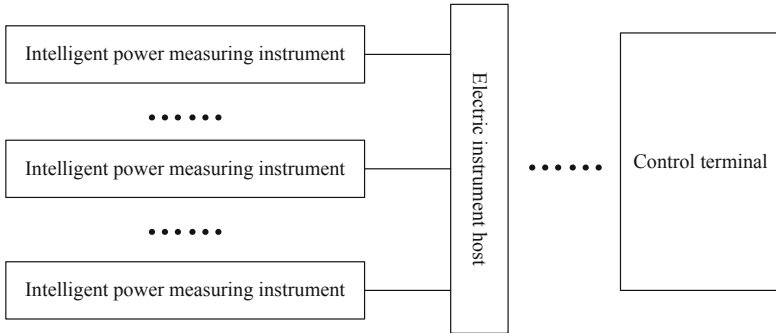
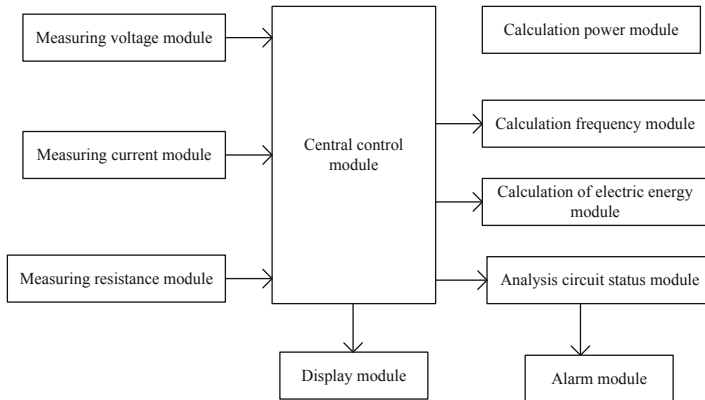


Fig. 1. Schematic block diagram of intelligent power measurement system

As shown in Fig. 1, in this paper, the intelligent power measuring instrument is used for the intelligent power measuring system. The intelligent power measuring system includes the power instrument host and control terminal. Multiple intelligent power measuring instruments are electrically connected to the same power instrument host, and the power instrument host is connected to the control terminal through communication. In other words, multiple intelligent power measuring instruments correspond to one power instrument host, and multiple power instrument hosts correspond to the same control terminal, which can facilitate efficient management of power measuring instruments.

### 3 System Structure

As shown in Fig. 2, the intelligent electrical engineering test system proposed in this paper includes: voltage measurement module, flow measurement module, resistance measurement module, central control module, power metering module, frequency metering module, power metering module, circuit status analysis module, and display module.



**Fig. 2.** Block Diagram of Intelligent Electrical Engineering Measurement System

The current detection system is connected with the central measurement and control system to measure the electrical pressure value through the current detection meter. The flow detection module, connected with the central monitoring module, is used to measure the electrical current data with a current detector. The resistance measurement module, connected with the central measurement module, is used to calculate the electrical resistance parameters through the detection instrument according to the voltage. The central control module is connected with the pressure measurement module, voltage measurement module, input and output power calculation module, frequency measurement module, input and output electricity statistics module, circuit status analysis module, display module, etc., and is used for the normal operation of all units through the microcontroller.

Power calculation module, connected with central control module, is used to calculate electrical power data through computer equipment. Frequency calculation module, connected with central control module, is used to calculate electrical frequency data through computer equipment. The electric energy calculation module, connected with the central control module, is used to calculate electric energy data through computer equipment. The circuit status analysis module is connected with the central control module to analyze the electrical circuit status according to the measured data through the analysis program. The state and abnormal data of the electrical circuit are transmitted to the alarm system to realize the alarm. The display system, connected with the central control, is used to display the measured electrical voltage, current, resistance and calculated electrical power, frequency and electric energy data information through the display.

## 4 Circuit Principle

As shown in Fig. 3, the parameter setting function given therein is connected with the master operating system of the microcontroller to provide initialization settings for all technical parameters realized by the microcontroller; The parameter configuration

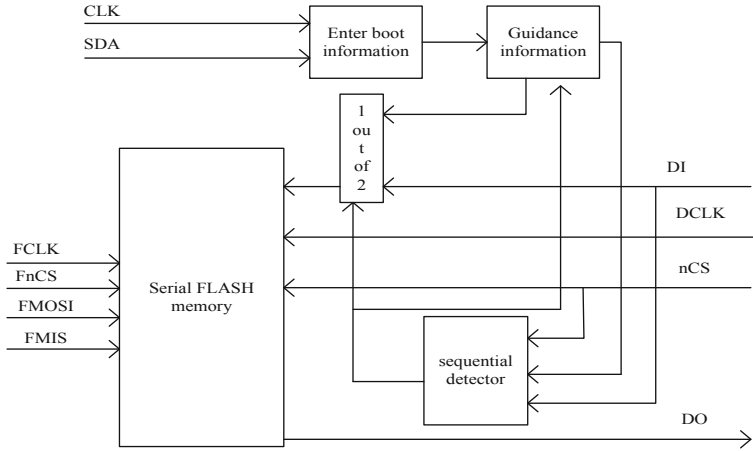


Fig. 3. Schematic diagram of configuration circuit

function constitutes the distribution integrated circuit; The distribution integrated circuit includes serial FLASH memory; Use FCLK, FNCS, FMOSKFMIS ports of serial FLASH memory to access peripheral data processors; Input the setting data to the serial FLASH memory using the above connection; The interface uses CLK and SDA lines to access peripheral data processing; Use the above connection to allocate the boot data information; Use wires to access FPGA in nConfig, DI, DCLK, nCS and DO companies; The data information input mode for serial FLASH memory is: the external data processor uses FCLK, FNCS, FMOSKFMIS lines to input the setting data information.

The drive steps of the configuration circuit are as follows: When the nConfig message is valid, the FPGA will reconfigure; Due to the input and output of FPGA and the combination of DLCLK and nCS, when the sequence detector detects the specified sequence information, the output through the sequence detector will make the boot message buffer module generated by the selected module; When no special sequence signal is detected by the sequence detector, select a module to output DI directly from FPGA; When the external processor's CLK After SDA finishes the input of guidance information, it will generate nConfig information and read and write the guidance information to the cache module of the guidance information. When the nConfig information takes effect, it will trigger the online FPGA installation process; If the external processor fails to pass CLK SDA executes actions, and the data stored in its pilot signal buffer system is the same as the pilot signal output by DI by default.

The function of microcontroller controller is connected with parameter configuration function, real-time digital simulation function and fault detection module to monitor the work of dispatching parameter setting function, real-time digital simulation function and fault detection module; The real-time digital simulation function is connected with the microcontroller controller function to establish the RTDS simulation mode of the AC/DC transmission control system through the RTDS real-time numerical simulator; In the AC/DC transmission system RTDS simulation module constructed by RTDS real-time digital simulator, the AC/DC transmission system RTDS simulation module is

constructed by RTDS real-time digital simulator using network computing technology; The content includes: describe and analyze the uncertainty of network computing individuals and network computing personal computing activities in the data space from the perspective of cloud model.

From the perspective of computing space, the behaviors of network computing individuals and network computing individuals and their roles in computing space are described and studied;

Simulation modeling and visualization analysis of the interaction mechanism of network computing and the internal dynamic mechanism of the network; The accident detection module is connected with the microcontroller control module to detect mechanical or line accidents; The optical measurement module is connected with the MCU control module, which is used to vertically access the fault recorder via the optical cable, and adjust the corresponding electrical variable ratio in the fault recorder; The optical detection module is connected to the fault recorder, and the optical data transmission mode is divided into: sending n-channel source signal, receiving the data of aliasing phenomenon by the receiving end integrated with the optical detection module, and isolating multi-channel aliasing data by the isolation system integrated with the optical detection module; Sending n-channel source signals means that after mixing the n-channel source signals through the channel, the mixed system is A, which is sent from space by n antenna systems in the sending terminal; The transmitting end of the integrated control system of the optical measurement module receives the aliasing signal, which means that the transmitting end first uses m dual polarized antennas to receive the aliasing signal, and the received message is called observation information. Then the transmitting end first completes the pre-processing data of the observation information. The pre-processing data generally involves two stages, mainly centralized processing and spherical processing; The discrete control system integrated with the optical measurement module separates multiple aliased signals, which means that the separation control system W will first separate the multiple aliased signals from the entropy according to the difference of the message beam value of each source signal, and its criterion for the message beam value is negative.

## 5 Measurement Method

Step (A). The circuit status analysis module is used to analyze the electrical circuit status according to the measured data through the analysis program, including:

- (a) For the operation state information I and n operation states of the electrical circuit to be divided with the input size of MXN;
- (b) The evolution curve can be automatically initialized and encoded as 0 with level set function;
- (c) Align the running state of n with a priori in a moment based manner;
- (d) Encode the aligned running state with a priori level set function to get 0;
- (e) The priori level set functions of all motion states are expanded into column vectors by column to form a priori matrix of motion states.

- (f) According to independent component analysis, reduce the dimension of matrix  $D$  required for motion state: preprocess the matrix required for motion state rationale; ICA dimension reduction can be carried out for the operation state or prior matrix after pretreatment;
- (g) The current level set function is expanded into vectors in columns, and then the projection matrix is obtained to map it to a low dimensional space;
- (h) Statistic the prior probability distribution of the motion state in the low dimensional space, establish the driving energy term of the operation state, and then combine it with the data driven energy term based on the region to form the total kinetic energy value;
- (i) Minimize the energy function to drive the curve evolution and obtain the operating state results of the electrical circuit;

Step (B). For abnormal data of electrical circuit status, the alarm data information is sent through the information alarm chip of the alarm module;

The exception analysis module of the alarm module receives the alarm data information sent by the information alarm chip, processes the real-time abnormal data, and captures the abnormal conditions of the abnormal data in a certain period of time; Send the captured information to the display module.

The data dimension reduction curve fitting is performed on the selected detection data to obtain the detection curve; By judging the similarity of the two curves, judge the abnormal value of the electrical circuit operation state. That is to say, there is no abnormal value at a node, otherwise, it is deemed to have abnormal value. The anomaly analysis module determines the current or voltage statistics of each circuit state; Determine each maximum manageable load through current or voltage statistics of circuit state; Through the power management method of the integrated platform, many first job functions are integrated into a single first continuous job function; Determine whether a load of the first continuous operation function is equal to the maximum manageable load; When the load of the first continuous work target is equal to the maximum machinable load, move the first operation target of one of the overloaded components of the first continuous work target out of the first continuous work target; When receiving the first continuous working target, transfer the central processing chip of the exception analysis module from a sleep mode to an operation module to carry out the first continuous working target; After the first continuous working target is processed, the central processing chip is reset to the sleep module; The central processing chip determines the next first action frequency according to the critical value of the load of the first continuous working target.

## 6 Conclusions

In this paper, the solar power module can get cleaner and more effective sunlight, thus saving investment, economic and environmental protection, and can supply power for a long time, effectively preventing interruption in the electrical engineering measurement process; At the same time, the measurement results of the fault detection module are more comprehensive, which improves the effect of troubleshooting. In this paper, we

use the Bingfen reuse algorithm to separate the multi-channel information in the Bingyu domain, so as to achieve the multi-channel multiplexing of channels. With greater channel capacity and higher spectrum signal utilization efficiency, we can obtain higher speed and high-quality broadband wireless communication services; It improves the disadvantage that the channel resources can only be allocated fixedly in the original orthogonal frequency division reuse method, alleviates the cross influence and multiple access interference of each channel information in the time domain and frequency domain during information transmission, and completes the multiplexing of time-frequency aliasing multiplexing and aliasing information in the information beam domain. When the utilization efficiency of signal resources is further improved, the content of the wireless channel is also further improved. By using negative Bing as the criterion for data independence, the division process of multi-channel aliasing information is first completed in the information Bing domain, and then the Bing division multiplexing algorithm is used to further improve the content of the wireless channel and the utilization effect of frequency resources.

The configuration circuit in this paper assumes that the capacity of FLASH memory is sufficient, that is, three or more FPGA configuration information can be stored in one FLASH chip at the same time, which can reduce the number of chips by 60% or more; The external processor can realize repeated allocation of FPGA through CLK.SDA, and can also realize "hot conversion" with FPGA interface Through the numerical simulation in this paper, the computer network has achieved remarkable results, with rich quantitative results on a larger scale, which can completely and truly describe various interaction characteristics, and the description is accurate. The simulation in this paper provides reference methods for the development of big data analysis. The information sources for big data analysis are broader, the information size is smaller, the information is more fragmented, and the structure is more diverse.

## References

1. Domingo-Ferrer, J., Farràs, O., Ribes-González, J., Sánchez, D.: Privacy-preserving cloud computing on sensitive data: a survey of methods, products and challenges. *Comput. Commun.* (2019)
2. Koroniotis, N., Moustafa, N., Sitnikova, E.: Forensics and deep learning mechanisms for botnets in internet of things: a survey of challenges and solutions. *IEEE Access* (2019)
3. Ylmaz, E.N., Gnen, S.: Attack detection/prevention system against cyber attack in industrial control systems. *Comput. Secur.* (2018)
4. Al-rimy, B.A.S., Maarof, M.A., Shaid, S.Z.M.: Ransomware threat success factors, taxonomy, and countermeasures: a survey and research directions. *Comput. Secur.* (2018)
5. Deshmukh, S., Di Troia, F., Stamp, M.: Vigenère scores for malware detection. *J. Comput. Virol. Hacking Tech.* (2) (2018)
6. Coutinho, M., de Oliveira Albuquerque, R., Borges, F., Villalba, L.J.G., Kim, T.-H.: Learning perfectly secure cryptography to protect communications with adversarial neural cryptography. *Sensors* (5) (2018)
7. Shah, S.A.R., Issac, B.: Performance comparison of intrusion detection systems and application of machine learning to Snort system. *Future Gener. Comput. Syst.* (2018)
8. Vijayakumar, P., Chang, V., Jegatha Deborah, L., Balusamy, B., Shynu, P.G.: Computationally efficient privacy preserving anonymous mutual and batch authentication schemes for vehicular ad hoc networks. *Future Gener. Comput. Syst.* (2018)

9. Burg, A., Chattopadhyay, A., Lam, K.Y.: Wireless communication and security issues for cyber–physical systems and the Internet-of-Things. *Proc. IEEE* (1) (2018)
10. Gholizadeh, N., Saadatfar, H., Hanafi, N.: K-DBSCAN: an improved DBSCAN algorithm for big data. *J. Supercomput.* (2020). (prep)
11. Scitovski, R., Sabo, K.: A combination of k -means and DBSCAN algorithm for solving the multiple generalized circle detection problem. *Adv. Data Anal. Classif.* (2020). (prep)
12. Govender, P., Sivakumar, V.: Application of k -means and hierarchical clustering techniques for analysis of air pollution: a review (1980–2019). *Atmos. Pollut. Res.* (2020) (1)
13. Kang, J.S., et al.: Development of a systematic, self-consistent algorithm for the K-DEMO steady-state operation scenario. *Nucl. Fus.* (12) (2017)
14. Doostan, M., Chowdhury, B.H.: Power distribution system fault cause analysis by using association rule mining. *Electr. Power Syst. Res.* (2017)
15. Wang, A.L., Chen, B.X., Wang, C.G., Hua, D.: Non-intrusive load monitoring algorithm based on features of V–I trajectory. *Electr. Power Syst. Res.* (2018)
16. Saleem, Y., Crespi, N., Rehmani, M.H., Copeland, R.: Internet of Things-aided smart grid: technologies, architectures, applications, prototypes, and future research directions. *IEEE Access* (2019)
17. Caballero, P., Banchs, A., de Veciana, G., Costa Perez, X., Azcorra, A.: Network slicing for guaranteed rate services: admission control and resource allocation games. *IEEE Trans. Wirel. Commun.* (10) (2018)
18. Nils, D., Fabian, K., Christian, W.: On the economic benefits of software-defined networking and network slicing for smart grid communications. *NETNOMICS Econ. Res. Electron. Netw.* (1–2) (2018)



# Routing Algorithm for Connected Vehicles Based on Bus Route Information

Ning Tong<sup>(✉)</sup>, Man Jin, and Kunrui Wang

School of Software, Dalian Jiaotong University, Dalian, Liaoning, China  
toni\_tong@163.com

**Abstract.** Aiming at the problems of the Internet of Vehicles routing algorithm in specific mobile communication scenarios of cars, buses and pedestrians, this paper presents a routing algorithm based on bus behavior (BBR). The movement patterns and routes of public transport vehicles have certain regularities. Using these regularities, this paper constructs an abstract network model, and based on the model, proposes a packet routing algorithm that keeps approaching public transport vehicles. A series of simulation experiments are carried out on the Bus Behavior algorithm and the classical routing algorithm by the ONE simulation tool. The simulation results show that the BBR performs well in the four performance indicators of message transmission success rate, load rate, average delay, and average number of hops. It shows that the algorithm has a high transmission success rate, low network load and delay, and good network stability, which is suitable for the urban traffic communication scenario studied in this paper.

**Keywords:** Internet of Vehicles · Bus Route Information · Routing Algorithm · Bus Behavior

## 1 Introduction

With the development of the Internet of Things, VANET has increasingly become a research and development hotspot. VANET is a special application of traditional MANET in the field of transportation [1]. VANET takes moving vehicles and transportation facilities as nodes and uses wireless communication technology to form a mobile network. In VANET, mobile nodes are smart cars that interact using three types of communications. The three types of communication include vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I) and hybrid communication [2]. In VANET, each vehicle is considered as a node, and its movement can be predictable because the vehicle is limited to the existing road. In addition, each vehicle will become part of the network, which manages and controls the communication on the network to conform to specified requirements [3]. Under the increasingly complex urban traffic environment, VANET faces problems such as unstable data transmission and frequent connection interruptions [4]. In the research of this paper, it is mainly aimed at the daily working hours of the complex urban traffic environment. In the scenario where the car is used as a mobile node to transmit messages between buses with fixed trajectories and pedestrians moving randomly, it is hoped that the source node will move towards some specific directions. Route forwarding to avoid routing loops or invalid routes.



## 2 Related Work

In the research of VANET routing algorithm, relevant scholars have achieved certain results. Classical routing algorithms such as Epidemic, DD, and FC have their specific routing transmission modes and applicable scenarios. Some newer routing algorithms proposed recently also have certain practical significance because of their specific ideas and mathematical models.

The Epidemic algorithm was proposed by Vahdat et al. It is a multiple copy (n-copy) routing protocol. The basic idea is that the mobile node stores the message in its data buffer after receiving this message or its copy from the source and carries the message with it as it moves when the next forwarding hop is currently unavailable, once it enters the communication range of another node this node copies the message and forwards the copy to the new node [5]. This technique is suitable for network environments where nodes are more mobile and network resources are more adequate, and are prone to congestion when node network resources are limited, leading to a decrease in message forwarding efficiency [6].

Direct Delivery (DD) algorithm, which is a very simple algorithm that does not use any message base data. It is a single-copy routing algorithm in which the message carrying node forwards messages only when it encounters the final destination node [7]. The disadvantage of this algorithm is that it is very inefficient and overly dependent on limited scenarios, requiring nodes with high caching capacity, and may not exhibit outstanding performance in general scenarios.

First Contact (FC), similar to DD algorithm, is a simple algorithm that does not use any message base, the difference is that the message carrying node will forward the message whenever it encounters a node until it reaches the destination node. FC algorithm embodies the function of forwarding, but it is a kind of purposeless forwarding. Although FC generates a small amount of message copy redundancy, consumes fewer network resources, but because it does not properly select the relay nodes through the relevant utility functions, the message may have problems such as the node carrying the message going backwards from the destination node and loops in the message path during the delivery process, resulting in a suboptimal delivery rate [8].

The ProPHET (Probabilistic Routing Protocol using History of Encounters and Transitivity) routing algorithm uses a forwarding strategy based on encounter prediction, where each node in this algorithm estimates its own encounter probability of reaching other nodes, while at the same time this encounter probability is then used as the utility value of the route [9]. But the disadvantage is that the larger the elected nodes, the higher the reliability of the transmission, which can greatly consume network resources [10]. Similar to the PRoPHET routing algorithm, there is also a routing algorithm MaxProp that uses the history as a utility value to predict the probability of encounter between vehicle nodes. In the MaxProp routing algorithm, each node also needs to calculate and update the probability and history of encounter with other nodes, and when the nodes successfully meet, the encountering nodes exchange the probability information maintained by each node. But also has a certain disadvantage is that the algorithm is computationally more loaded and is not ideal for use in a network environment with high vehicle density. To address the previous problem, the literature [11] proposed MaxProp routing protocol based on group characteristics, which takes advantage of the excellent

connectivity between member nodes in a group, and indirectly increases the encounter probability between nodes in the group and nodes outside the group through intra-group message diffusion, thus increasing the message forwarding opportunities.

The human behavior-based message routing algorithm [12] (Human Behavior Router, HBR) was recently proposed by Andrade et al. The HBR routing algorithm does not rely on the replication of messages between roadside units and nodes, and is based on human behavior to predict the location of nodes at the destination, and obtains location-related information about the network nodes through computation. The HBR is implemented through complex mathematical model constraints to simulate human behavior through the following data: all encountered vehicles have up-to-date information about the destination, its location and speed.

The literature [13] constructs a cluster structure with bus nodes as cluster head nodes and common car nodes as intra-cluster nodes, and this strategy makes reasonable use of information such as fixed bus running routes and speed stability. The literature [14] found some basic characteristics of bus trajectories through statistical analysis of real bus trajectories; using time series analysis, a model for predicting the running time of buses on road sections was given, and the effectiveness of the prediction method was verified through experiments.

### 3 Routing Algorithm Based on Bus Route Information

The characteristics of bus vehicles determine the way they move (such as time, route and other information) with certain regularity. In this paper, we use the information of the bus vehicle travel route to give the routing algorithm based on the bus route information.

#### 3.1 Network Model

The network coverage is a rectangular area, where the length  $a$  meters and the width  $b$  meters. The set of nodes  $V_u = \{u_1, u_2, \dots, u_n\}$ ,  $|V_n| = n$ . Three types of nodes are distributed in the Area: general vehicles  $C$ , buses  $T$ , and pedestrians  $P$ . To distinguish the nodes in the network, we give each node an identity. At any given moment, any node has an identity  $ID$ ,  $ID \in \{C, T, P\}$ . When a node joins the network, the identity is set and it is specified that the identity cannot be changed. The node motion uses a map-based motion model. The bus node  $u$ .  $ID = T$  makes a round-trip movement on a defined path with stops set on the path, and continues the movement after a stopping time  $t$  at the stop. For other types of nodes, after selecting a destination, move along the shortest path to the destination; after reaching the destination, stay for a moment and continue to select the next destination.

The undirected graph  $G = (V_u, E(t))$  represents an Ad hoc network.  $E(t) \subseteq V \times V$  is the set of edges of the connected nodes, represents the set of links that exist in the network at time  $t$ . The edge  $e_{uv}(t) \in E$  (Among  $u, v \in V_u$ ) indicates the existence of a wireless link between nodes  $u$  and  $v$ .

### 3.2 Algorithm Description

The research scenario in this paper is how to route a bus node if the destination node is a bus node when the bus moves according to a fixed movement trajectory in an urban environment. For this scenario, a Bus Behavior Routing algorithm (BBR) based on bus behavior is given.

The mathematical model of the algorithm is shown in Fig. 1. In the environment of simulated urban traffic operation, it is assumed that the two points A and B in Fig. 1 are the starting and ending stations of the bus. During daily working hours, there are different numbers of buses moving along the path at the two points AB. The source node Source and the distance to A, B two points respectively constitute a triangle constraint, then in the process of random movement of the source node Source will do logical operations on its communication range of neighbor nodes To1, To2, only when the neighbor nodes meet the distance to A, B two points are less than the distance from the source node to A, B two points, it means that the current position of the neighbor node tends to the source node to pass to the destination node, then the source node will copy its own copy of the message to the neighbor node.

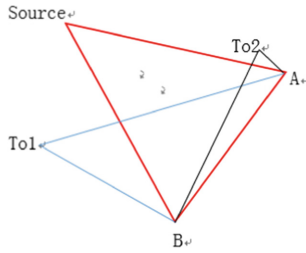


Fig. 1. Bus Behavior routing algorithm

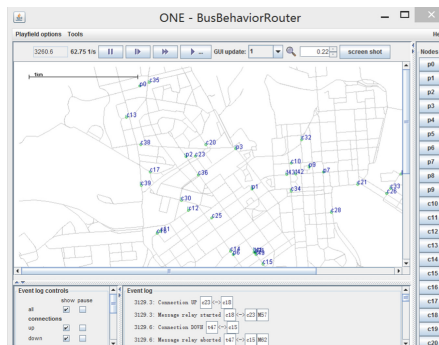


Fig. 2. Simulation interface

After understanding the above mathematical model, we can intuitively see the two neighbor nodes To1 and To2 shown in Fig. 1. It is obvious that the distance from To1 node to point A is greater than the distance from Source node to point A. If one condition

is not satisfied, it is determined to fail, so the source node will not pass its own message data to the To1 neighbor node. In the same way, for the neighbor node To2, it is easy to see that the current position of the To2 node is less than the distance between the Source node and the A and B points, so in the current position, the Source node will prefer to pass its own copy of the message to the neighbor node To2. Subsequently, the To2 node becomes the new source node and once again performs constraint determination, and so on, until the message is delivered to the destination node.

## 4 Simulation Analysis

An open source software environment simulator (Opportunistic Network Environment Simulator, ONE) written in Java at the University of Helsinki, Finland, was used as the simulation tool. Because ONE focuses on the store-carry-transfer network [15]. ONE comes with a variety of node movement models, both random-based and map-based [16, 17].

**Table 1.** Simulation parameters configuration

parameters	values
Simulation area	4500 m * 3400 m
Numbers of Nodes of vehicle	50, 100, 200, 400
Wireless transmission range	10 m
Nodes type	P(Pedestrian), C(vehicle), T(bus)
Movement model	Shortest Path Map Based Movement
Cache Space	5M
Wait time	(0–120) seconds

One of them, Map-Based Movement Model (MBM), is applicable to simulate the vehicle driving along the road on the road. MBM is a derivative of the random waypoint movement model. MBM allows the user to add maps in which the nodes in the model move along streets and roads on the map. The simulation interface is shown in Fig. 2. The simulation parameters are shown in Table 1.

We simulate the implementation of Epidemic, Direct Delivery, First Contact and the routing algorithm BBR proposed in this paper for bus communication environment. The performance is compared in terms of four basic performance metrics (delivery success rate, average delay, routing load factor, and average number of hops).

The variation of transmission success rate with the number of nodes is shown in Fig. 3. Using the strategy proposed in this paper to select the forwarding nodes with reference to the bus operation path; BBR is slightly lower than Epidemic when the number of nodes is 50, and achieves the highest transmission success rate when the number of nodes increases to 100. This is because the algorithm's specify the approximate direction of message forwarding, and the message delivery has directionality, which improves the delivery success rate[18].

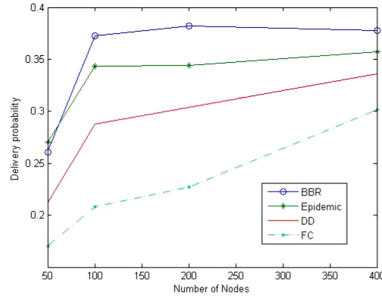


Fig. 3. Delivery probability

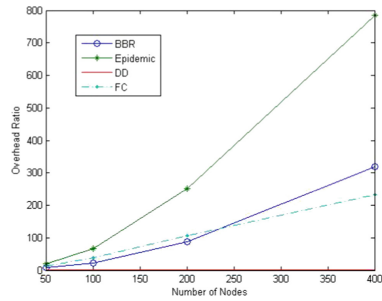


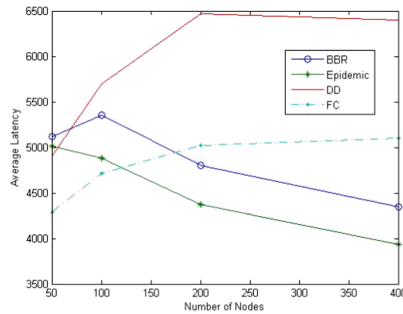
Fig. 4. Overhead Ratio

The routing load ratios vary with the number of nodes, as shown in Fig. 4. With the increase in the number of nodes, the load is on an upward trend. BBR, Direct Delivery and First Contact all have low load ratios, while Epidemic has poor performance in this area. In the following, we analyze the reasons for this phenomenon: BBR, Direct Delivery and First Contact use different strategies to suppress the number of messages forwarded in the process of message forwarding. In the BBR proposed in this paper, the source node does not forward the message to all neighboring nodes, and only a limited number of neighbors participate in forwarding. This strategy reduces the overhead of delivering packets and reduces the routing load ratios.

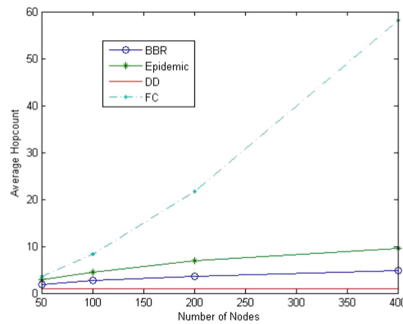
Figure 5 reflects the variation of the average delay with the increasing number of nodes for the four routing algorithms. The Direct Delivery algorithm has a low average delay in the first set of experiments, but the average delay increases rapidly to the highest as the number of nodes increases; The First Contact algorithm has the lowest average latency in both the first and second sets of experiments, but the average latency of the algorithm increases to a higher level as the number of nodes increases. The average latency of the Epidemic algorithm gradually decreases to the lowest among the four algorithms as the number of nodes increases; BBS is second in comparison, with the average latency first rising slightly as the number of nodes increases, and then also decreasing to a lower standard.

Figure 6 reflects the variation of the average number of hops with the increasing number of nodes for the four routing algorithms. As can be seen from the figure, the

Direct Delivery algorithm, due to its simple transmission principle, has an average hop count that does not vary with the number of nodes and is always 1. The average number of hops of the First Contact algorithm shows a nearly 3-fold increase in magnitude with increasing number of nodes in the four experiments. The average hop count of Epidemic algorithm and BBR algorithm increases gradually with the number of nodes, among the two, Bus Behavior routing algorithm appears to have a lower average hop count and shows good performance.



**Fig. 5.** Latency with increasing number of nodes



**Fig. 6.** Hopcount varies with the number of nodes

The graphical analysis shows that the BBR routing algorithm has a good routing performance for the specific scenario of car, bus and pedestrian communication studied in this paper.

## 5 Conclusions

This paper addresses the increasingly complex urban traffic environment where VANETs face problems such as unstable data transmission, frequent connection interruptions, and dependence on fixed infrastructure. A specific bus behavior-based routing algorithm BBR is given after considering the advantages and disadvantages of various classical routing algorithms routing performance in the scenarios of urban cars, buses with fixed

trajectories, and pedestrian mobile communication, and after referring to the corresponding mathematical models. The algorithm proposes an improvement on the classical routing algorithms Epidemic and Direct Delivery by restricting the source nodes to transmit messages only in the direction that meets the constraints of the mathematical model. Four algorithms, BBR, Epidemic, Direct Delivery, and First Contact, are simulated with the help of the open source simulation simulator ONE for different numbers of nodes. The BBR algorithm exhibits the highest message transfer success ratios, lower load ratios, lower average latency and average hop count, with more suitable scenario-specific and usage implications.

**Acknowledgements.** This work is supported by the National Natural Science Foundation of China (No. 61902051).

## References

1. Hu, S., Jia, Y., She, C.: Performance analysis of VANET routing protocols and implementation of a VANET terminal. In: 2017 International Conference on Computer Technology, Electronics and Communication (ICCTEC), pp. 1248–1252 (2017)
2. Yassine, H., Salah, M.: VANET cross-layer routing. In: 2019 International Conference of Computer Science and Renewable Energies (ICCSRE), pp. 1–2 (2019)
3. Khalid Diaa, M., Samer Mohamed, I., Ayman Hassan, M.: OPBRP - obstacle prediction based routing protocol in VANETs. *Ain Shams Eng. J.* (2022)
4. Zhao, H., Chen, S., Li, D., et al.: The establishment of the network connectivity model in VANET. In: Proceedings of the International Conference on Wireless Communications & Signal Processing, Yangzhou, China. IEEE, October 2016
5. Tian, D., Zhou, J., Wang, Y., Zhang, G., Xia, H.: An adaptive vehicular epidemic routing method based on attractor selection model. *Ad Hoc Netw.* **36**, Part 2 (2016)
6. Alenazi, M.J.F. Cheng, Y., Zhang, D., et al.: Epidemic routing protocol implementation in ns-3. In: Proceedings of the 2015 Workshop on ns-3, pp. 83–90 (2015)
7. Rodrigues, J.J.P.C., Soares, V.N.G.J.: An introduction to delay and disruption tolerant networks (DTNs). In: Electronic and Optical Materials, Advances in Delay-Tolerant Networks (DTNs), 2nd edn. Woodhead Publishing (2021)
8. Spaho, E., Barolli, L., Kolicic, V., et al.: Evaluation of single-copy and multiple-copy routing protocols in a realistic VDTN scenario. In: 2016 10th International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS), pp. 284–289. IEEE (2016)
9. Nam, J.C., Kim, E.H., Lee, M.K., et al.: Enhanced PROPHET based on message delivery predictability in delay tolerant networks. In: Proceedings of the Information and Communication Technology Convergence, Jeju, South Korea, pp. 457–459. IEEE (2015)
10. Pathak, S., Gondaliya, N., Raja, N.: A survey on PROPHET based routing protocol in delay tolerant network. In: Proceedings of the ICEI, Pune, India. IEEE, February 2017
11. Harrati, Y., Abdali, A.: MaxHopCount: DTN congestion control algorithm under MaxProp routing. *IJCSNS* **17**(5), 206 (2017)
12. de Andrade, G.E., de Paula Lima, L.A., Calsavara, A., de Oliveira, J.A., Michelon, G.: Message routing in vehicular delay-tolerant networks based on human behavior. In: 2016 10th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP), pp. 1–6. IEEE (2016)
13. Tao, B., Li, D., Zhang, G., Yang, Y.: Research on zone routing protocol based on bus backbone network. *Comput. Eng.* <https://doi.org/10.3969/j.issn.1000-3428.2016.03.002>

14. Zhang, F.-S., Jin, B.-H., Wang, Z.-Y., Hu, J.-F., Zhang, L.-F.: A routing mechanism over bus-based VANETs by mining trajectories. *Jisuanji Xuebao/Chin. J. Comput.* **38**, 648–662 (2015). <https://doi.org/10.3724/SP.J.1016.2015.00648>
15. Mass, J., Srirama, S.N., Chang, C.: STEP-ONE: simulated testbed for edge-fog processes based on the opportunistic network environment simulator. *J. Syst. Softw.* **166** (2020)
16. Ekman, F., Keränen, A., Karvo, J.: Working day movement model. In: Proceedings of the ACM SIGMOBILE Workshop on Mobility Models, Hong Kong, China. ACM (2008)
17. Keranen, A.: Opportunistic network environment simulator. Special assignment report, Helsinki University of Technology. Department of Communications and Networking (2008)
18. Issac, G.A., Mary, A.J.: Validation scheme for VANET. In: 2019 2nd International Conference on Signal Processing and Communication (ICSPC), pp. 11–15 (2019)





# Processing Method of Computer Data

Xiaolin Jiang<sup>(✉)</sup>

Shandong Institute of Commerce and Technology, Jinan, Shandong, China  
dzkbs9@163.com

**Abstract.** This paper studies a data processing method, which is applied to the client. The method includes: determining the first privacy node set and the second privacy node set. The first privacy node set includes one or more nodes in the blockchain network composed of multiple nodes, and the second privacy node set is a subset of the first privacy node set; Acquiring the key, and each node in the first private node set also acquires the key; Encrypt the transaction data based on the key to obtain the encrypted data; Send the encrypted data to each node in the second private node set, and these nodes will endorse and verify the encrypted data based on the key; If the endorsement verification is successful, the encrypted data will be sent to the blockchain network. Each node in the blockchain network will store the encrypted data in the blockchain. Each node in the first private node set can decrypt the encrypted data based on the key. This paper also provides a data processing device and a computer system.

**Keywords:** Computer · Data Processing · Encrypted Data · Network

## 1 Introduction

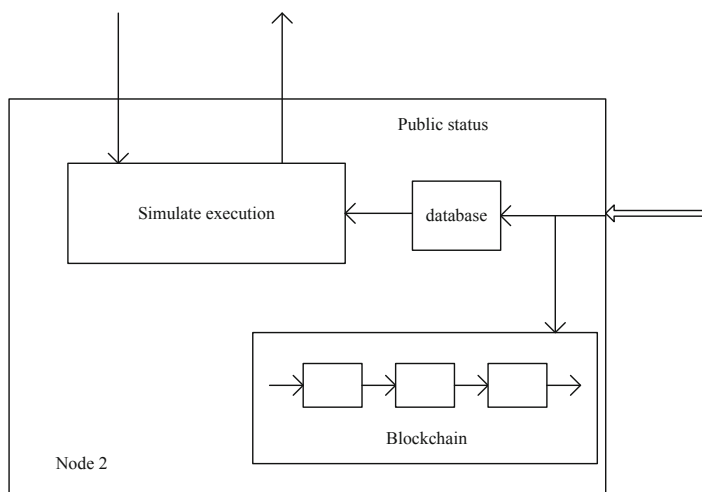
With the popularity of computers and the unique computing power of computers, people increasingly like to use computers to assist their work, one of which is to use computers to process data with tables [1]. At present, when you want to select part of the data in a column of the table, there are generally the following methods: first, walk through all the rows of the table to determine whether the row is the data you want to select in turn; [2] when the row is the data you want to select, click the row to select the row until you walk through all the rows of the table to select all the data you want to select; second, when you are in the table, When the data you want to select is more than the data you don't want to select, it is the opposite of the first method [3]. Go through all the rows of the table, select all the data you don't want to select in turn, and then click the Invert Selection option to select all the data you want to select [4]. In the prior art, whether the first method or the second method, when the number of rows in the table is large, it is very easy to miss or make a wrong selection when selecting data [5].

One aspect of this paper provides a data processing method, which is applied to the client [6]. The method includes: determining the first privacy node set and the second privacy node set [7]. If the endorsement verification is successful, the encrypted data will be sent to the blockchain network, and the encrypted data will be stored in the blockchain by each node in the blockchain network, where each node in the first privacy node set

can decrypt the encrypted data based on the key [8]. This paper also provides a data processing method, which is applied to a node in a blockchain network composed of multiple nodes [9]. The method includes: when a node belongs to the first private node set determined by the client, obtaining the key, the client and other nodes in the first private node set also obtain the key, and the first private node set includes one or more nodes in the blockchain network; [10] When a node belongs to the second private node set determined by the client, it receives the encrypted data from the client and performs endorsement verification based on the key. The encrypted data is generated by the client to encrypt the transaction data based on the key. The second private node set is a subset of the first private node set; Receive the encrypted data with successful endorsement verification and perform consensus verification based on the key; In addition, other nodes in the blockchain network will store the encrypted data verified by consensus into the blockchain.

## 2 Data Processing

Figure 1 and Fig. 2 illustrate the data processing method. Wherein, Fig. 1 schematically shows the data processing process in the prior art, and Fig. 2 schematically shows the data processing process.



**Fig. 1.** Schematic diagram of data processing in prior art

As shown in Fig. 1, the above shows the interaction process between the client and the node in the blockchain network in the existing technology. The client sends the transaction data to the endorsement node respectively [11]. The endorsement node performs endorsement verification on the transaction data in its own local area. The endorsement result of the endorsement node determines whether the transaction data has been successfully endorsed [12]. The rules for determining the endorsement verification

result can be set according to actual needs, and there is no restriction here [13]. After confirming that the endorsement verification of transaction data is successful, the client sends the transaction data to the sorting node [14]. The sorting node does not approve the content of the received transaction data. The sorting node sorts the transaction data A and other received transaction data according to the predetermined sorting rules, and determines the time to submit the transaction data according to the sorting position. When the corresponding time arrives, the sorting node sends the transaction data to multiple nodes in the blockchain network. After consensus verification, each node stores the transaction data in their respective blockchains, realizing the uplink of transaction data.

In the data processing process shown in Fig. 1, the transaction data stored in the blockchain is visible to all nodes, and the transaction data is visible to all endorsement nodes during the endorsement verification process of each endorsement node, which is prone to the privacy disclosure problem of transaction data.

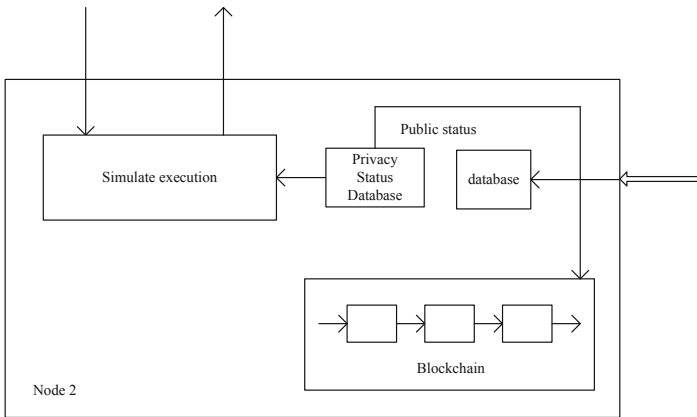


Fig. 2. Schematic diagram of data processing in this paper

The improved data processing process shown in Fig. 2 can solve the above problems. As shown in Fig. 2, the upper part shows the interaction process between the client and the nodes in the blockchain network.

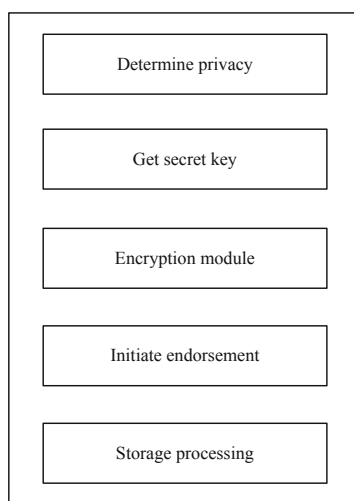
After confirming that the encrypted data endorsement verification is successful, the client encapsulates the message of successful endorsement and sends it to the sorting node [15]. The sorting node does not approve the received content. The sorting node sorts the encrypted data and other received transaction data according to the predetermined sorting rules, and determines the time to submit the encrypted data according to the sorting position. After the consensus verification, each node stores the encrypted data in its corresponding blockchain, realizing the uplink of encrypted data. At the same time, the transaction data corresponding to the encrypted data can be stored in the local privacy database.

It can be seen that for transaction data A in the endorsement verification process, in the uplink process, and after the uplink, it is visible to the node holding the key, but not to other nodes. The endorsement verification process is performed in the node, saving

the computing resources required in the endorsement process. According to the data processing method in this paper, a retrieval mechanism is also set. When the node that should hold the key, such as Node 2, accidentally loses the key or the transaction data in the privacy database and other privacy messages, Node 2 can send a privacy information request message to Node 1 or 3, request corresponding privacy data according to the specific situation, and ensure the consistency of the internal privacy data of each node in the first privacy node set designated by the client.

### 3 Data Processing Device

Figure 3 schematically shows a block diagram according to a data processing apparatus. The data processing apparatus is applied to the client.

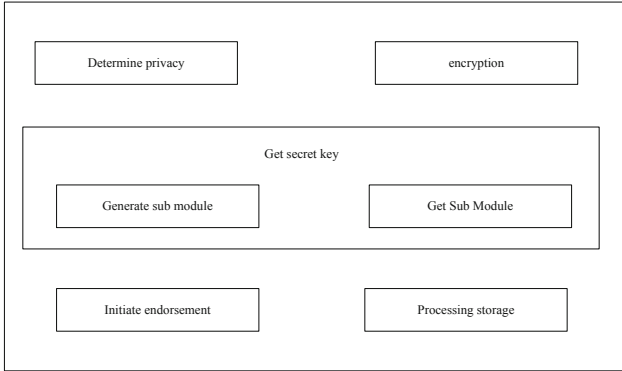


**Fig. 3.** One of the block diagrams of the data processing apparatus

As shown in Fig. 3, the data processing device includes a privacy determination module, a key acquisition module, an encryption module, an endorsement initiation module, and a storage processing module. The privacy determination module is used to determine the first privacy node set and the second privacy node set. The first privacy node set includes one or more nodes in the blockchain network composed of multiple nodes, and the second privacy node set is a subset of the first privacy node set.

The key acquisition module is used to acquire the key, and each node in the first private node set also acquires the key. The encryption module is used to encrypt the transaction data based on the key to obtain the encrypted data. Endorsement processing module is used to send encrypted data to each node in the second privacy node set, and each node in the second privacy node set will endorse and verify the encrypted data based on the key. The storage processing module is used to send the encrypted data to the blockchain network if the endorsement verification is successful, and each node in

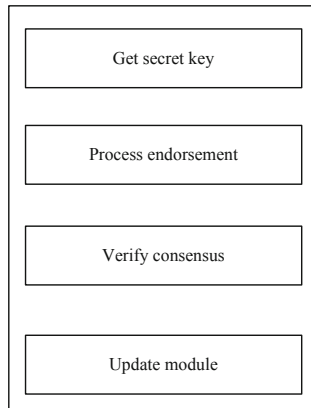
the blockchain network stores the encrypted data in the blockchain, where each node in the first private node set can decrypt the encrypted data based on the key. Figure 4 schematically shows a block diagram according to a data processing apparatus. The data processing apparatus is applied to the client.



**Fig. 4.** Block Diagram 2 of Data Processing Device

As shown in Fig. 4, the data processing device includes a privacy determination module, a key acquisition module, an encryption module, an endorsement initiation module, and a storage processing module. The privacy determination module, key acquisition module, encryption module, endorsement initiation module, and storage processing module respectively have the same functions as the privacy determination module, key acquisition module, encryption module, endorsement initiation module, and storage processing module. The key acquisition module includes a generation sub module and an acquisition sub module. Either of them can realize the above operation of obtaining the key.

The blockchain network also includes sorting nodes. The storage processing module is specifically used to send the encrypted data to the sorting node, which sorts the encrypted data, and sends the encrypted data to other nodes in the blockchain network except the sorting node at a specified time based on the sorting position, so that other nodes store the encrypted data in their respective blockchains. Figure 5 shows a block diagram of a data processing apparatus. The data processing device is applied to a node in a blockchain network composed of multiple nodes.



**Fig. 5.** Block Diagram 3 of Data Processing Device

As shown in Fig. 5, the data processing device includes a key acquisition module, an endorsement processing module, a consensus verification module, and an update module. The key acquisition module is used to acquire the key when a node belongs to the first private node set determined by the client. The client and other nodes in the first private node set also obtain the key, and the first private node set includes one or more nodes in the blockchain network.

The endorsement processing module is used to receive encrypted data from the client and perform endorsement verification based on the key when a node belongs to the second private node set determined by the client. The encrypted data is generated by the client encrypting the transaction data based on the key, and the second private node set is a subset of the first private node set; The consensus verification module is used to receive the encrypted data with successful endorsement verification and perform consensus verification based on the key.

The update module is used to store encrypted data verified by consensus into the blockchain with other nodes in the blockchain network. Figure 6 schematically shows a block diagram of a data processing apparatus. The data processing device is applied to a node in a blockchain network composed of multiple nodes.

As shown in Fig. 6, the data processing device includes a key acquisition module, an endorsement processing module, a consensus verification module, and an update module. Among them, the key acquisition module, endorsement processing module, consensus verification module, and update module have the same functions as the key acquisition module, endorsement processing module, consensus verification module, and update module, and the repeated parts will not be repeated.

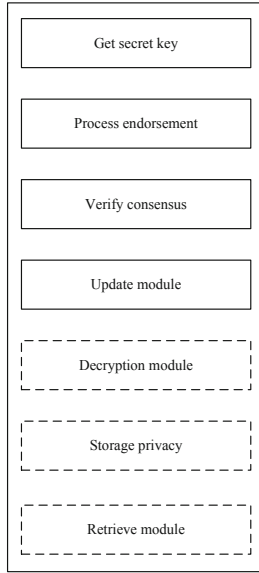


Fig. 6. Block Diagram 4 of Data Processing Device

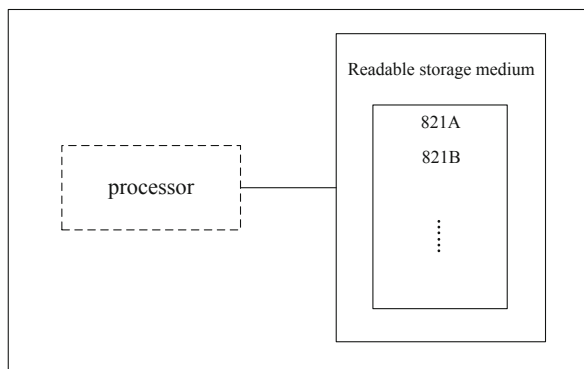
The data processing device also comprises a decryption module and a privacy storage module. The decryption module is used to decrypt the encrypted data based on the key to obtain transaction data when it is stored in the blockchain with other nodes in the blockchain network through consensus verification. The privacy storage module is used to store transaction data in a local privacy database of a node. The data processing device also includes a retrieval module.

### 4 Computer System

Figure 7 schematically shows a block diagram of a computer system adapted to implement the methods described above.

As shown in Fig. 7, a computer system includes a processor and a computer-readable storage medium. Processors may include, for example, general-purpose microprocessors, instruction set processors and/or related chipsets and/or dedicated microprocessors, and the like. The processor may also include on-board memory for caching purposes. The processor may be a single processing unit or a plurality of processing units for performing different actions.

Computer readable storage medium, such as nonvolatile computer readable storage medium, including but not limited to: magnetic storage device, such as tape or hard disk; Optical storage device, such as CD-ROM; Memory, such as RAM or flash memory; wait. The computer-readable storage medium may include a computer program, which may include code/computer executable instructions that, when executed by a processor, cause the processor to execute a method or any variation thereof.



**Fig. 7.** Block Diagram of Computer System

At least one of the privacy determination module, key acquisition module, encryption module, endorsement initiation module, and storage processing module can be implemented as a computer program module described with reference to Fig. 7, which can implement the above knowledge-based answer generation method when executed by the processor.

## 5 Conclusions

This paper provides a data processing device, which is applied to the client. The device includes: privacy determination module, key acquisition module, encryption module, endorsement initiation module, and storage processing module. The storage processing module is used to send the encrypted data to the blockchain network if the endorsement verification is successful, and each node in the blockchain network stores the encrypted data in the blockchain, where each node in the first private node set can decrypt the encrypted data based on the key. In addition, on the other hand, this paper provides a data processing device, which is applied to a node in a blockchain network composed of multiple nodes. The device comprises a key acquisition module, an endorsement processing module, a consensus verification module, and an update module. The consensus verification module is used to receive the encrypted data with successful endorsement verification and perform consensus verification based on the key. In addition, the update module is used to store encrypted data verified by consensus into the blockchain with other nodes in the blockchain network.

## References

1. Mohd Jamaludin, S.Z., Mohd Kasihmuddin, M.S., Md Ismail, A.I., Mansor, M.A., Md Basir, Md.F.: Energy based logic mining analysis with hopfield neural network for recruitment evaluation. *Entropy* (1) (2020)
2. Woo, H., Kim, J., Lee, W.: Analysis of cross-referencing artificial intelligence topics based on sentence modeling. *Appl. Sci.* (11) (2020)



3. Papagiannopoulou, E., Tsoumakas, G.: A review of keyphrase extraction. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* (2) (2020)
4. García-Pérez, A., Sorribes-Palmer, F., Alonso, G., Ravanbakhsh, A.: Overview and application of FEM methods for shock analysis in space instruments. *Aerosp. Sci. Technol.* (2018)
5. Molontay, R., Horvath, N., Bergmann, J., Szekrenyes, D.L., Szabo, M.: Characterizing curriculum prerequisite networks by a student flow approach. *IEEE Trans. Learn. Technol.* (2020)
6. Knorn, S., Varagnolo, D., Staffas, K., Wrigstad, T., Fjllstrm, E.: Quantitative analysis of curricula coherence using directed graphs. *IFAC Papers OnLine* (9) (2019)
7. Reimers, N., Gurevych, I.: Sentence-BERT: sentence embeddings using siamese BERT-networks. *CoRR* (2019)
8. Vlaovi, B., Vree, A.: Discrete time model for process meta language with fictitious-clock. *Appl. Sci.* (6) (2022)
9. Basile, D., Fantechi, A., Rucher, L., Mandò, G.: Analysing an autonomous tramway positioning system with the Uppaal statistical model checker. *Form. Asp. Comput.* (2021). (prep)
10. Anand, A., Bhatt, N., Alhazmi, O.H.: Vulnerability discovery modelling: a general framework. *Int. J. Inf. Comput. Secur.*, 1–2 (2021)
11. Belardinelli, F., Lomuscio, A., Yu, E.: Model checking temporal epistemic logic under bounded recall. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, no. 05 (2020)
12. Basu, T., Aggarwal, K., Wang, C., Chattopadhyay, S.: An exploration of effective fuzzing for side-channel cache leakage. *Softw. Test. Verif. Reliab.* (1) (2020)
13. Zarezadeh, M., Mala, H., Ladani, B.T.: Efficient secure pattern matching with malicious adversaries. *IEEE Trans. Dependable Secure Comput.* (99) (2020)
14. Palazzi, L., Li, G., Fang, B., Pattabiraman, K.: Improving the accuracy of IR-level fault injection. *IEEE Trans. Dependable Secure Comput.* (2020)
15. Mokshin, A.V., Mokshin, V.V., Sharnin, L.M.: Adaptive genetic algorithms used to analyze behavior of complex system. *Commun. Nonlinear Sci. Numer.* (2018)



# Storage System of Computer Data

Peng Wang<sup>(✉)</sup>

Shandong Institute of Commerce and Technology, Jinan, Shandong, China

jwusp7@163.com

**Abstract.** This paper studies a data storage method, device and computer storage medium, including: obtaining the current operation of the back-end storage engine Objectstore, and determining the target category of the current operation according to the operation ID corresponding to the current operation; The current operation is compiled according to the compilation rules corresponding to the target category, and the data corresponding to the compiled current operation is stored accordingly. The corresponding compilation rules refer to the format requirements matching the specified storage protocol. This technical implementation mode adds a new data storage method to CEPH ecology, and also provides a bridge for object storage devices and services as a distributed unified storage.

**Keywords:** Computer · Storage Data · Storage Protocol

## 1 Introduction

At present, the methods to realize fast storage of big data include: receiving multiple types of data to be warehoused through a unified data warehousing interface; Staging the received data to be received into the message queue; [1] The polling service is used to dequeue the temporary data in the message queue and store the data to be warehoused in the database [2]. This paper realizes the fast storage of big data, which is mainly designed and used for the big data industry, and can be used independently [3]. As the main means of data storage and collection, it can also be used together with data services, so that big data can be applied in practical applications and play a greater role in data [4].

However, for some public computer-based public data storage systems, many people will use such public storage systems for file storage; [5] This will cause a problem, that is, there is too much content stored in it [6]. The content stored by each person and the content that needs to be read out later are different [7]. Each person's needs are different, and the storage location is different [8]. In this way, for some public storage systems, how to effectively store the data associated with each other and provide users with convenient access; It is convenient for users to obtain the corresponding storage content in time; In order to solve this technical problem, a solution is provided. In view of this, this paper studies a more widely used data storage method, device and computer storage medium.

## 2 Process of Data Storage Method

As shown in Fig. 1, the data corresponding to the compiled current operation is stored accordingly, including: the data corresponding to the compiled current operation is stored in at least one of the following ways: to a local object storage device, a local object storage system, or a network object storage service.

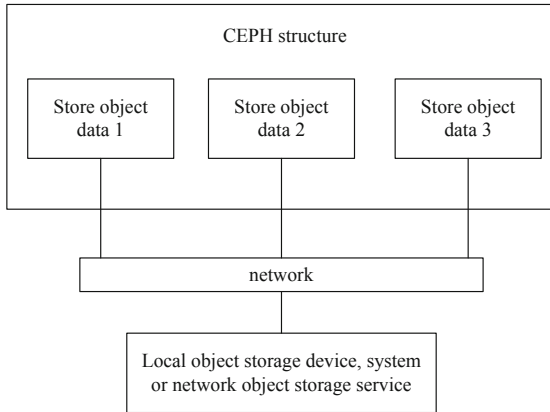


Fig. 1. One of the flow diagrams of the data storage method

It can be seen from Fig. 1 that in the technical implementation mode, the current operation of obtaining the object store is classified into target categories by obtaining the current operation of the back-end storage engine Objectstore [9]. When the current operation is determined to be a target category operation, the current operation is compiled according to the compilation rules corresponding to the target category, and the data corresponding to the compiled current operation is stored to the local object storage device. The local object storage system or network object storage service does not need to change the architecture of CEPH itself [10]. It can still use the RADOS structure [11]. The OSD created in this way can be mixed with the OSDs of other back-end media without impact [12]. Specifically, it includes OSD process under CEPH, which is sent to object storage device, system or network object storage service through network after compilation [13]. In an optional embodiment, the object storage device, system, or network object storage service may be any one or more combinations of the local object storage device, local object storage system, or network object storage service. Therefore, the creation of OSDs can no longer be limited to using local disks as data disks, but can use network resources.

When the target category is the first target category, the current operation is compiled according to the compilation rules corresponding to the target category, including the prefix format conversion of the data corresponding to the current operation, and the converted data corresponding to the current operation matches the data format requirements of the specified storage protocol [14]. The first target category specifically refers to the operation category that needs to be compiled to be accepted by the specified protocol

[15]. The operations of the first target category are also operations where data needs to be stored to a local object storage device, a local object storage system, or a network object storage service. When the target category is the second target category, before compiling the current operation according to the compilation rules corresponding to the target category, it also includes: caching the data corresponding to the current operation. Here, the second target category operation mainly refers to the operation category that requires multiple operations on the same object or intensive data. When the operation belongs to the second category, the data of the operation is cached to improve the response speed and performance. Cache the data corresponding to the current operation, including sending multiple target requests in sequence and processing the return values of multiple target requests asynchronously, where multiple target requests match the operations of the second target category.

When caching the data corresponding to the operations of the second target category, it provides network asynchronous operation management, and sends multiple target requests in sequence. The return value of asynchronous processing multiple target requests can include: adding a data structure pointer to record multiple requests at one time in the compilation process, using this parameter to send multiple requests in sequence for each link at one time, and processing callbacks asynchronously. Aggregate requests can be of different types. In this way, multiple target requests can be issued sequentially and the return values of multiple target requests can be processed asynchronously. As a result, the performance of multiple write operations merging and intensive data read ahead for the same object is improved.

When the target category is a target category that does not meet the set requirements, the method also includes: writing the data of the current operation to the local file system or bare disk based on the solid state hard disk and mechanical hard disk. When the target category is a target category that does not meet the set requirements, you can directly write to the local storage space without compiling it. Operations that do not meet the set requirements can include log writing and cleaning operations. When the current operation can be completed without compiling, the data corresponding to the operation can be processed using the existing logic of the Objectstore and written into the storage space based on the local mechanical hard disk or solid state disk.

The data corresponding to the compiled current operation is stored accordingly, including: according to the object storage device OSD to which the current operation belongs, the data corresponding to the compiled current operation is stored in the data storage area bucket corresponding to the OSD. The current operation refers to the object class operation or transaction class operation that the OSD distributes to the Objectstore. The OSD to which the current operation belongs refers to the corresponding OSD that distributes it to the Objectstore.

As shown in Fig. 2, the data storage method provided in this paper includes the following steps: (A). Get the current operation of the Objectstore; (B). Judge whether the current operation is the first target category. If yes, execute step (C); If not, perform step (G); Step (C) Judge whether the current operation is the second target category. If yes, execute step (D); if no, execute step (E); Step (D). Cache the current operation; Step (E) Compile the data corresponding to the current operation to make it match the format requirements of the specified storage protocol; Step (F), Store the compiled data

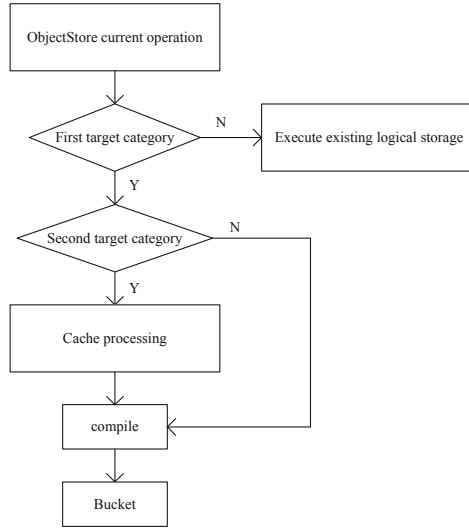


Fig. 2. Flow Diagram 2 of Data Storage Method

corresponding to the current operation in the bucket corresponding to the OSD data storage area; Step (G) Executes an existing logical store. Here, an OSD corresponds to a bucket in the storage area. A bucket in the object storage can be used as an OSD of CEPH. The OSD created in this way can be mixed with the OSDs of other back-end media without impact.

### 3 Hardware Structure of Data Storage

On the other hand, as shown in Fig. 3, this paper also provides a data storage device, which includes an engine module and a connection module. The engine module is used to obtain the current operation of the back-end storage engine Objectstore, and determine the target category of the current operation according to the operation ID corresponding to the current operation; The connection module is used to compile the current operation according to the compilation rules corresponding to the target category, and store the data corresponding to the current operation after compilation. The corresponding compilation rules refer to the format requirements matching the specified storage protocol.

Here, the engine module divides the operations under the storage engine Objectstore into object class operations and transaction class operations. For example, object class operations can include object reading, object writing, object collection creation, object information entry addition, deletion, and modification. Transaction class operations can include thread management classes, log classes, synchronization locks, and valve locks. Different operations have different operation identifiers. The operation identifiers are mainly flag bits, which are enum (data type) sets. The current operations in the Objectstore can be classified according to different flag bits.

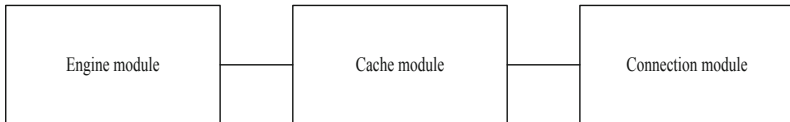
After the target category operation is identified, the connection module will compile the current operation. The compilation rules are based on the format requirements of the

storage protocol. The connection module is specifically used to store the compiled data corresponding to the current operation in at least one of the following ways: to a local object storage device, a local object storage system, or a network object storage service.

Where, when the target category is the first target category, the current operation is compiled according to the compilation rules corresponding to the target category. The engine module is specifically used to convert the data corresponding to the current operation into prefix form, and the data corresponding to the current operation after conversion matches the data format requirements of the specified storage protocol.



**Fig. 3.** Structure 1 of the structures of the data storage device



**Fig. 4.** Structure 2 of data storage device

As shown in Fig. 4, it can also include a cache module. When the target category is the second target category, the cache module is used to cache the data corresponding to the current operation before compiling the current operation according to the compilation rules corresponding to the target category.

Among them, the cache module caches the data corresponding to the current operation, including: sending multiple target requests sequentially, and processing the return values of multiple target requests asynchronously, where multiple target requests match the operations of the second target category. In this way, the cache module can issue multiple target requests sequentially and process the return values of multiple target requests asynchronously. As a result, the performance of multiple write operations merging and intensive data read ahead for the same object is improved.

Where, when the engine module determines that the target category is a target category that does not meet the set requirements, it will call the existing logic module of the Objectstore to write the currently operated data to the local file system or bare disk based on the solid state disk and mechanical disk.

The connection module stores the data corresponding to the current operation after compilation, encapsulates the corresponding interface, and stores the data corresponding to the current operation after compilation in the bucket corresponding to the data storage area of the OSD according to the object storage device OSD to which the current operation belongs.

### 4 Transfer Level of Data Storage Mode

As shown in Fig. 5, in the data storage device provided in this paper, the engine module can be the S3Store module, the connection module can be the S3Conn module, and the cache module can be the S3Cache module. Among them, the S3Store module obtains the current operation from the Objectstore engine, distributes the target category operation, and implements the Objectstore semantic interface. On the other hand, the S3Store module also needs to preprocess these operations; The S3Cache module caches operations of the second target category; The S3Conn module encapsulates the libs3 interface or RESTful interface to complete the flush, read, or list tasks issued by the Cache layer, and then sends them to the object-oriented data storage area.

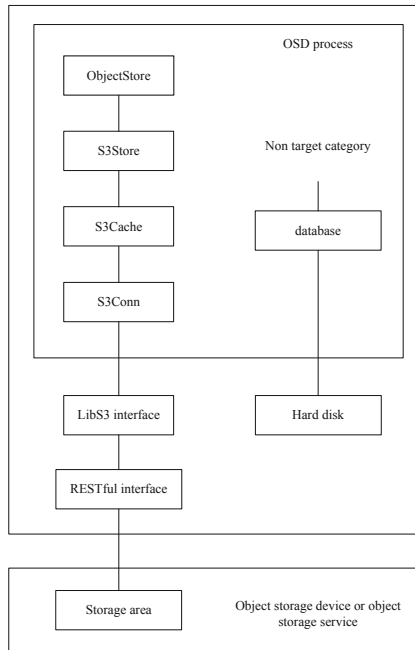


Fig. 5. Transfer hierarchy of data storage mode

In addition, some transaction operations issued by the Objectstore engine are completed by using the database and stored in the hard disk to ensure the atomicity of transaction operations.

Among them, the S3Store module and S3Conn module can distinguish the categories of the current operation of the acquired Objectstore engine, and compile according to different target categories to match the format requirements of the S3 storage protocol. The oS3Cache module is designed as one of the core modules for the performance problems that may arise from the concept of zero fetching based on object storage and the performance problems that may arise from network communication. The S3Cache module can complete the filling and reading of objects during the whole storage. It will

also provide network connection asynchronous operation management, and improve the performance of multiple write operations for the same object and intensive data read ahead.

The S3Store module can be used as a new back-end engine of the Objectstore engine to distribute target categories and non target categories; The S3Cache module is responsible for caching operations of the second target category issued by the Objectstore engine; The S3Conn module completes the communication with the libs3 interface and the downloading and pre reading of the S3Cache module, thereby realizing the unified storage of data operated by different target categories under the CEPH ecosystem, forming a new CEPH object storage back-end, adding a new data storage method to the CEPH ecosystem, and providing a bridge for object storage devices and services as a distributed unified storage. The data storage method and device provided in this paper have at least the following beneficial effects:

- A) A new back-end storage medium has been added to the distributed storage CEPH, which can use local object storage devices, systems, or network object storage services.
- B) Back end object storage is used to store CEPH object data, avoiding re conversion from object to file, and saving performance overhead in interface conversion.
- C) It follows CEPH's principle of storing local objects in object storage, and avoids the write amplification of additional transaction atomicity guarantee caused by interface conversion.
- D) It provides the application space of block storage and file storage for the existing object storage service or object storage system with narrow application scope, and expands the way out for the application of high-capacity stand-alone devices that only provide object interfaces or low-cost network object storage services.
- E) It provides a bridge for distributed unified storage for object storage devices and services, and provides a way for existing object storage to upgrade to unified storage.

## 5 Conclusions

The data storage method, device and computer storage medium provided in this paper determine the target category of the current operation according to the operation ID corresponding to the current operation by obtaining the current operation of the back-end storage engine Objectstore; In this way, target categories are distinguished for the current operation to obtain the Objectstore. When it is determined that the current operation is a target category operation, the current operation is compiled according to the compilation rules corresponding to the target category, and the data corresponding to the compiled current operation is stored accordingly. The corresponding compilation rules refer to matching the format requirements of the specified storage protocol. In this way, after the operations under the Objectstore are distinguished, the operations of the target category are compiled, so that the data corresponding to the compiled operations conform to the format requirements of the specified storage protocol, thereby realizing the unified storage of the data of operations of different target categories, forming a new CEPH object storage backend, and adding a new data storage method to the CEPH ecosystem,



It also provides a bridge for object storage devices and services as a distributed unified storage.

## References

1. Belardinelli, F., Lomuscio, A., Yu, E.: Model checking temporal epistemic logic under bounded recall. In: *Proceedings of the AAAI Conference on Artificial Intelligence (05)* (2020)
2. Basu, T., Aggarwal, K., Wang, C., Chattopadhyay, S.: An exploration of effective fuzzing for side-channel cache leakage. *Softw. Test. Verif. Reliab.* (1) (2020)
3. Zarezadeh, M., Mala, H., Ladani, B.T.: Efficient secure pattern matching with malicious adversaries. *IEEE Trans. Dependable Secure Comput.* (99) (2020)
4. Palazzi, L., Li, G., Fang, B., Pattabiraman, K.: Improving the accuracy of IR-level fault injection. *IEEE Trans. Dependable Secure Comput.* (2020)
5. Mokshin, A.V., Mokshin, V.V., Sharnin, L.M.: Adaptive genetic algorithms used to analyze behavior of complex system. *Commun. Nonlinear Sci. Numer.* (2018)
6. Knorn, S., Varagnolo, D., Staffas, K., Wrigstad, T., Fjellström, E.: Quantitative analysis of curricula coherence using directed graphs. *IFAC Papers OnLine* (9) (2019)
7. Reimers, N., Gurevych, I.: Sentence-BERT: sentence embeddings using siamese BERT-networks. *CoRR* (2019)
8. Vlaovi, B., Vree, A.: Discrete time model for process meta language with fictitious-clock. *Appl. Sci.* (6) (2022)
9. Basile, D., Fantechi, A., Rucher, L., Mandò, G.: Analysing an autonomous tramway positioning system with the Uppaal statistical model checker. *Form. Asp. Comput.* (2021). (prep)
10. Anand, A., Bhatt, N., Alhazmi, O.H.: Vulnerability discovery modelling: a general framework. *Int. J. Inf. Comput. Secur.* (1–2) (2021)
11. Mohd Jamaludin, S.Z., Mohd Kasihmuddin, M.S., Md Ismail, A.I., Mansor, M.A., Md Basir, M.F.: Energy based logic mining analysis with hopfield neural network for recruitment evaluation. *Entropy* **23**(1), 40 (2020)
12. Woo, H., Kim, J., Lee, W.: Analysis of cross-referencing artificial intelligence topics based on sentence modeling. *Appl. Sci.* (11) (2020)
13. Papagiannopoulou, E., Tsoumakas, G.: A review of keyphrase extraction. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* (2) (2020)
14. García-Pérez, A., Sorribes-Palmer, F., Alonso, G., Ravanbakhsh, A.: Overview and application of FEM methods for shock analysis in space instruments. *Aerosp. Sci. Technol.* (2018)
15. Molontay, R., Horvath, N., Bergmann, J., Szekrenyes, D.L., Szabo, M.: Characterizing curriculum prerequisite networks by a student flow approach. *IEEE Trans. Learn. Technol.* (2020)



# Design of Human-Computer Interaction Experience Platform in Virtual Space Under Scientific Information Technology

Yi Fu and Yuzhuo Jiang<sup>(✉)</sup>

School of Shenyang Architecture University, Shenyang, Liaoning, China  
1696515549@qq.com

**Abstract.** With the continuous progress and development of electronic science and information technology and modern information network technology, “virtual reality technology” (VR) has been widely used in various industries. In environmental design, it makes up for the shortcomings of the design scheme when traditional design can only be shown with two D drawings and three D renderings. Virtual reality technology has many application mode, it can simulate the building interior space, real present design characteristics, change the furniture style and interface decoration, let the user get more realistic experience and fun, and it can also timely according to customer demand arbitrary change design, save time cost, avoid the waste of resources.

**Keywords:** Virtual Reality Technology · Indoor Space Design · Human-computer Interaction Experience

## 1 Introduction

### 1.1 Concept of the VR System

The full name of VR is virtual reality technology, which uses a computer to generate an interactive 3 D dynamic view. At present, virtual reality technology can be divided into narrow and broad aspects:

In a narrow sense, the virtual reality:

Virtual reality, also called infinite virtual reality, is the ideal state. This means that computers are used to create a 3-dimensional virtual environment, which the experimenter can invest in using a variety of devices (3D glasses, head-mounted displays, circular projection devices, motion-sensing handles, etc.). In this environment, experiments can observe objects in the environment through natural skills (such as limbs and orbit), obtain multidisciplinary and natural observations, and have huge effects.

Generalized virtual reality:

It is often about creating an interactive environment simulating the real world, focusing on the interaction between humans, which also extends its research and applications.

## 1.2 Historical Evolution of VR Technology

The research and development of VR technology can be roughly divided into three development stages as show in Table 1.

**Table 1.** The development stage of VR technology

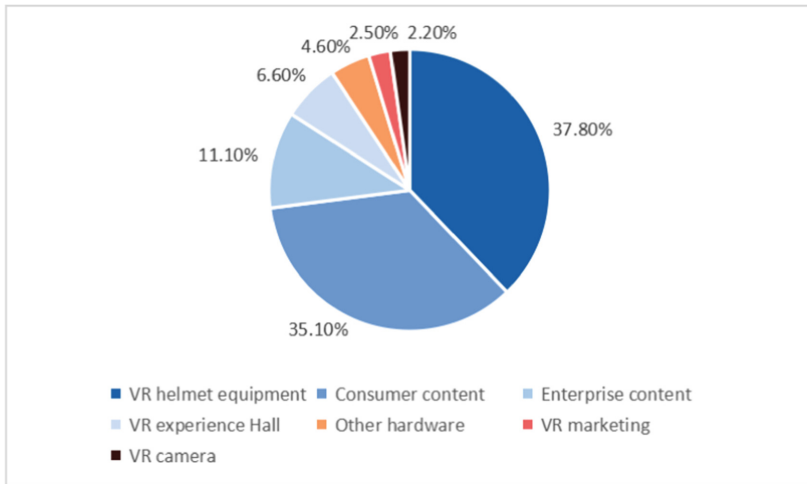
stage	time	figure	Event impact
Early Period: 1950s–1970s	In 1956	The American photographer, Morton Heilig	Has developed and produced the first photography machine to simulate the simulation environment: Sensorama Simulator
	In 1965	“Father of computer graphics” Lvan Sutherland	At the IFIP conference, The Ultimate Display has for the first time explicitly proposed a 3 D virtual sound system composed of sound devices and virtual sound with powerful audio feedback capabilities
Practical phase: 1970s–1980s	In 1973	Myron Krueger	Artificial reality is presented
	In 1989	“The Father of the VR”, Jaron Lanier	“Virtual Reality” was formally proposed for the term virtual reality
Stage of rapid development: from the 1990s to now	In 1993	Boeing Co	The Boeing 777 aircraft was successfully designed using VR technology
	In 1994	Burdea G and Coiffet	“Virtual Reality Technology” published the book “3I” (Imagination, Interaction, and Immersion) properties
	In 2012	Oculus Rift (Virtual reality device)	Starting a wave of civil VR equipment

## 1.3 Present Situation of VR Technology

The range of technologies in virtual augmented reality includes modern computer graphics, sensor processing techniques, dynamics, optics, artificial intelligence, and social psychology [1]. And based on information technology and new human-computer interaction. Modern information technology uses it to create realistic virtual environments

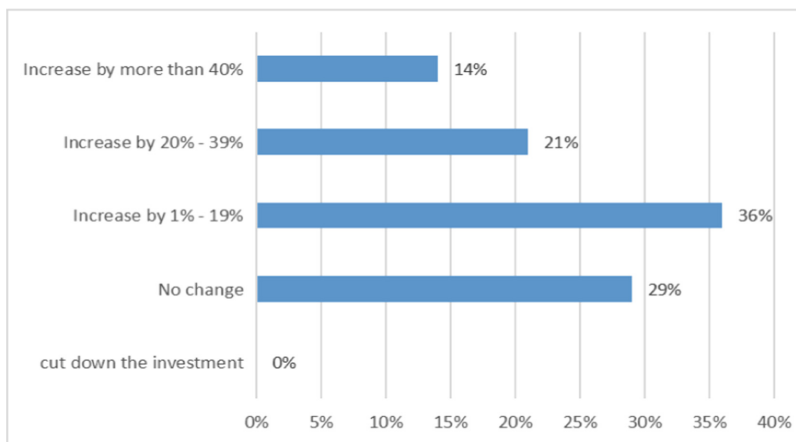
by creating visual, auditory, and touch. Users use the necessary devices in a virtual environment to naturally interact with objects to create the same feel and experience as the real environment.

The year 2016 is the first year of China’s virtual reality related technology. In recent years, VR technology has developed rapidly and steadily in recent years as show in Fig. 1.



**Fig. 1.** Estimated proportion of various Chinese VR market segments in 2021 (unit: %)

According to the big data virtual reality in KPMG 2020 Technology Industry Innovation, many domestic people have promoted their online commercial office and consumption during the epidemic period, so domestic enterprises have also begun to increase their



**Fig. 2.** Changes in enterprise investment in virtual reality technology in 2020 (unit: %)

research and development investment in domestic virtual augmented reality as show in Fig. 2.

#### **1.4 Characteristics of the VR Technique**

Virtual reality technology is widely recognized by the public as an interactive experience of a product, contributing to the research of VR technology [2]. It is based on 3D technology, which creates virtual space to experience it. It is divided into three points:

**Immersion:** now human visual perception of external things especially the visual experience accounted for most of the proportion, so in the virtual real world will pay more attention to the human observation and see the visual authenticity of external things, the virtual world can truly reproduce the participants perceive things, at the same time users can also through hardware equipment, such as data gloves, steering wheel to better perception of the virtual world, so as to achieve a better experience [3].

**Interaction:** In a virtual environment, users can interact with things in the virtual world with VR devices, such as data gloves, force feedback devices, data head voice recognition, etc. These sensors combine virtual and reality with each other, using the human body's sensory system for various actions, represent real experiences of contact, auditory and visual, and realize interactive feelings between the user and the virtual world.

**Lenovo:** in the virtual environment, generate rich imagination, meet the user's imagination and opinions of some design, such as changing the wallpaper pattern and color, display furniture and style, can be in VR data transmission control system to quickly complete the user perception setting, more intuitive show the user's preferences also broaden the vision.

## **2 Link Between VR Technology and Interior Design**

Nowadays, VR technology is deeply popular with the public with its real visual effect and simple and convenient operation, so its application has been expanded and widely used in various fields. Combining VR technology and interior design can not only more intuitively express the designer's design scheme, but also let users truly feel it, and put forward change opinions. This saves a lot of time in the design process and enhances the designers.

### **2.1 Key Points of VR Technology in Interior Design**

If you want to speed the model, you can ignore unnecessary information and consider only the location of walls, doors and windows in the condition. It defines arcs through three points in the wall series, and virtual reality can introduce shortcuts to speed the model and subsequently store the modified data in change records to return it to the previous stages and changes. With the fastest design and convenient operation, the design plan can be designed automatically using 3D data. The implementation of the 3D model will reduce worker measurement time and significantly improve efficiency [4]. Virtual reality technology and environmental design art have great artistic value and practical

value, and the two are inseparable. The combination of the two imperceptibly completes the implementation of environmental planning.

In the interior design, it is usually divided into decorative hard decoration and soft decoration. Hard include: inner wall, door, window, etc. Soft decoration includes: lighting, wallpaper, furniture, and decorative furnishings. In the interior space design, in order to realize the virtual placement of indoor soft and hard decoration, virtual visual reality can be used in the design of indoor wall structure, wallpaper and floor pattern [5]. According to the design of priority to watch the indoor layout, through this way, customers can have a three-dimensional visual experience. Secondly, if there are differences, through virtual technology to make users feel the indoor environment and atmosphere, and then benefit users to join their own style aesthetic, and adjust and improve, it will improve the satisfaction with the required space of customers, thus through the VR technology presents the design concept can achieve the planner and consumers want the best demonstration effect [6].

## 2.2 Phase of VR Technology in Interior Design

### Preparatory phase

If VR technology is used to participate in the preliminary design, the specific design evaluation and budget can be made out, which can make up for the gap that the traditional design technique is only based on hand-drawn and three-dimensional renderings. Therefore, designers can better control the specific implementation of the design concept and the smooth progress of the later planning through VR technology, which largely saves time and resources.

Designer's early conception is accurate is also abstract and erratic, the use of VR technology to generate virtual environment can make designers better immersed, receive the environment comprehensive information stimulation, can touch the inspiration to make thinking more active, and make the idea concrete, but also conducive to the design of all stages of the connection and advance thinking.

### Improvement stage

When a design scheme is clearly conceived, its further design needs to be considered. At this time, we can sort out and compare between the different main design perspectives presented in the simulation space by using VR technology. At the same time, in the interior design, you can also use VR technology to check the connection of the space structure, ceiling, partition and other fixed factors. Designers can also measure the size of the furniture in the VR virtual world, to generate, copy, zoom and other operations. Whether from the whole to local, macroscopic to micro, it can be observed and deliberated through VR technology.

### Display stage

The emergence of VR technology enables designers to intuitively communicate to customer design solutions, communicate with customers, and then to change is not enough to meet customer needs. A good scheme display can play twice the result with half the effort. The comprehensive effect of sound, light and color presented by VR is far beyond the cold drawings of traditional design techniques, It can bring more humanized experience and realize the perfect presentation of human-computer interaction.

### 2.3 Comparison and Analysis of Various Factors in the Design

Comprehensive simulation of spatial feelings. VR technology uses 3D software to design the interior space, according to the user needs, in the virtual space modeling design requirements, and conduct 3-dimensional stretching or synthesis on the basis of the indoor space modeling technology to show a more three-dimensional overall space design structure. After the establishment of the indoor virtual scene model, the virtual reality technology can be further used to render the model to generate the internal scene atmosphere, display, etc. In order to enable users to achieve a better experience, at the same time, designers can also modify the program content in time and reach a consensus with users, and design designs more suitable for the user experience.

Not limited by the time and site factors. Designers can design anytime and anywhere according to user requirements, VR scene in the overall consistent with the real world, through VR, from the outside can see the impact of the floor on lighting, experience the lighting layout and the impact on space, in the material selection, decoration, green plants and surrounding environment design to achieve liberalization [7]. At the same time, in some complex design, it can be more accurately calculated according to VR technology, and timely correct the insufficient links, so as to improve the design concept and the final ideal effect [8].

Reduce the time and the cost of capital. For interior design, designers can reasonably control the real space through VR technology, reduce human resources, but also achieve the sustainable development of low-carbon and energy saving, and directly convey the design concept to users. For example, from the design of the real estate developers' model room, VR technology can save the early investment cost for enterprises, and create the same effect as the real model room in a short time.

## 3 Development Trend of VR Technology

The rise of VR technology runs throughout various industries, and it will also be more and more widely used in environmental design.

### 3.1 The Integration of Various New Technologies

The future trend may be to integrate VR technologies, artificial intelligence technologies and 3D printing technologies and promote each other. As VR technology continues to develop, it breaks not only the limitations of design and creation, it will promote the close connection of various disciplines. Interior design will also be connected with other disciplines, such as digital media and visual communication, so as to create more fresh and dynamic works, constantly satisfying people's operational nature, visual feeling and interactive sense of design works. Under the new mode, it will more promote the expression of artistic expression space and explore new content more suitable for expression [9].

### 3.2 Common Development of Technology and Art

The design work needs comprehensive technology and art. The good and bad technology cannot unilaterally determine the degree of the work, and the artistic processing and aesthetic embodiment are needed, so as to bring superior visual aesthetic to people. If a work wants to lead people to leave a deep impression, it is bound to combine technology with art, with both the aesthetic feeling of art and the scientific and technological innovation. Future design works will combine VR technology, truly in sound, light and color and other aspects of a good feeling [10].

### 3.3 Pay More Attention to the “Humanized” Design

The purpose of the designer is to put people first and put the user experience first to consider the implementation of the overall scheme. For example, it is well known as IKEA, its commodity display forms a warm sense of environmental atmosphere, so that different furniture and accessories constitute the display space of functional areas, bringing users operational experience, rather than simple and tidy placement. One of the reasons why IKEA is popular is its user satisfaction, and it focuses more on humanistic care. Environmental design is people-oriented, human survival, and people are also the designer, builder and user of the environment.

## 4 Conclusions

With the improvement of indoor design level and people’s aesthetic level and demand, the traditional design techniques in some aspects cannot achieve the realistic touch and interactive experience, two-dimensional graphics expression is relatively limited, just from the visual feeling of three-dimensional space, cannot make customers have a sense of immersive. The emergence of VR technology just makes up for this defect. The virtual three-dimensional model of indoor space has the same size as the practical effect, and can render the real scene and display effect, so that the design industry is more technological and diversified, injecting new vitality into the interior design industry.

## References

1. Shan, N.: Research on indoor environment design based on VR technology and wireless sensor network. *Microprocess. Microsyst.* **2021**(83), 103999 (2021)
2. Chen, F.: Construction of a sharing mode based on VR technology in the context of big data. *J. Phys. Conf. Ser.* **1881**(4), 042015 (2021)
3. Kai, C., Lulu, L.: Research on the application of VR technology in interior design. In: *Proceedings of 2nd International Conference on Contemporary Education, Social Sciences and Ecological Studies*, pp. 746–749 (2019)
4. Luo, W., Huang, Y.: Application of VR technology in environmental art design. *J. Phys. Conf. Ser.* **2020**(1648), 032073 (2020)
5. Han, F., Liu, Y.: Indoor intelligent decoration system based on BIM + VR technology. *IOP Conf. Ser. Earth Environ. Sci.* **783**(1), 012121 (2021)



6. Prabhakaran, A., et al.: The effectiveness of interactive virtual reality for furniture, fixture and equipment design communication: an empirical study. *Eng. Constr. Archit. Manag.* **28**(5), 1440–1467 (2021)
7. Gong, M.: Analysis of architectural decoration esthetics based on VR technology and machine vision. *Soft Comput.* **25**(18), 1–13 (2021)
8. Deng, S.: Construction of soft decoration style and artistic aesthetic characteristics in interior design based on VR technology. *J. Phys. Conf. Ser.* **1744**(3), 032222 (2021)
9. Zou, Z.: Development status and trend of indoor VR technology design under the background of big data. *J. Phys. Conf. Ser.* **2020**(1648), 042100 (2020)
10. Bo, Z.: Application of VR technology in the innovation of art design. *Front. Art Res.* **1**(6), 62–66 (2019)



# Design of Intelligent Color Matching System for Cultural and Creative Products Based on Data Analysis Algorithm

Xuan Wan<sup>1</sup> (✉) and Yantao Liang<sup>2</sup>

<sup>1</sup> College of Art Design and Media, Sanda University, Shanghai, China  
wx15901835900@163.com

<sup>2</sup> Dassault Systèmes (Shanghai) I.T. Co., Ltd., Shanghai, China

**Abstract.** In today's era of rapid development of information, it is necessary to maximize the satisfaction of consumers' psychological needs in the process of product design. Pleasant color design is the basis for satisfying consumers' psychological expectations and an essential element of product design. Based on data analysis algorithms, this paper helps to build an intelligent color matching design system for cultural and creative products. It first expounds the research status of data analysis algorithms, then emphasizes the importance of color matching for cultural and creative products, and finally uses big data and artificial intelligence to simulate consumption. Consumers' perception of color after seeing color in computer science, and help design cultural and creative products that meet consumers' consumption preferences within a limited time.

**Keywords:** Data Analysis · Big Data · Artificial Intelligence · Color Matching of Cultural and Creative Products

## 1 Introduction

Product design is the use of specific products to meet people's specific requirements and expectations after reintegrating basic shapes, colors, symbols, and other design elements [1]. This is a way to better describe a process by expressing a specific idea, plan or problem solution through a specific design. Among the three aesthetic elements of product design, color is the most important element to perceive the fastest and convey product information, and can affect the visual information conveyed by shape and material. Therefore, in the design of cultural and creative products, improving the intelligence of color matching and enhancing the efficiency of color matching can promote consumers' satisfaction with cultural and creative products.

## 2 Research Status of Data Analysis Algorithms

### 2.1 Various Forms of Data Analysis

Machine learning is the most important field of artificial intelligence and the most important method of data analysis. Many scientists at home and abroad have studied the application of machine learning technology in big data analysis in various fields.

Big data clustering is often interdisciplinary, cross-functional, and cross-media, and traditional clustering algorithms are difficult to directly apply to big data clustering [2]. Therefore, big data clustering is getting more and more attention. Map Reduce is one of the mainstream frameworks for distributed computing. The parallel operation of traditional clustering algorithms based on MapReduce is an important method for big data analysis. The framework implements data parallelization. Based on the Hadoop platform, the traditional K-Means clustering algorithm is implemented, and the whole process is basically divided into three parts (map, join, reduction) [3]. The Map Reduce programming framework implements the bottom-up algorithm Agglomerative Hierarchical Clustering (AHC). This improves the accuracy and recognition rate of text clustering. A noise-based density-based clustering technique is implemented based on Map Reduce (Density-Based Spatial Clustering Applications with Noise, DB-SCAN) [4]. It is mainly composed of data preprocessing, local DBSCAN and computational intelligence acquisition. It is generally a computational intelligence technology based on mature development networks such as artificial neural, fuzzy systems, and evolutionary computing. They are organically combined with each other and determine the potential of artificial intelligence. Their superior application in big data analysis, as shown in Fig. 1 [5].

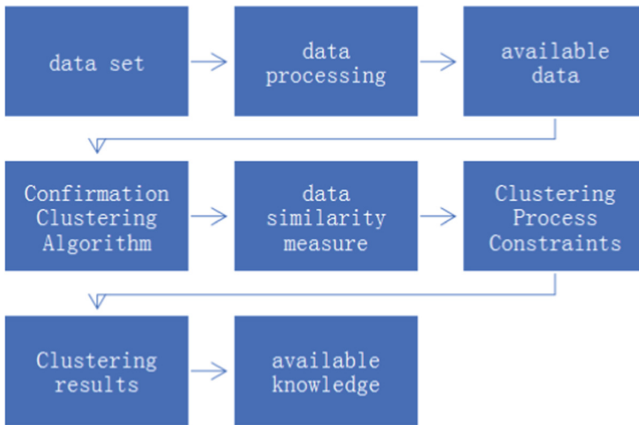


Fig. 1. General Clustering Flowchart

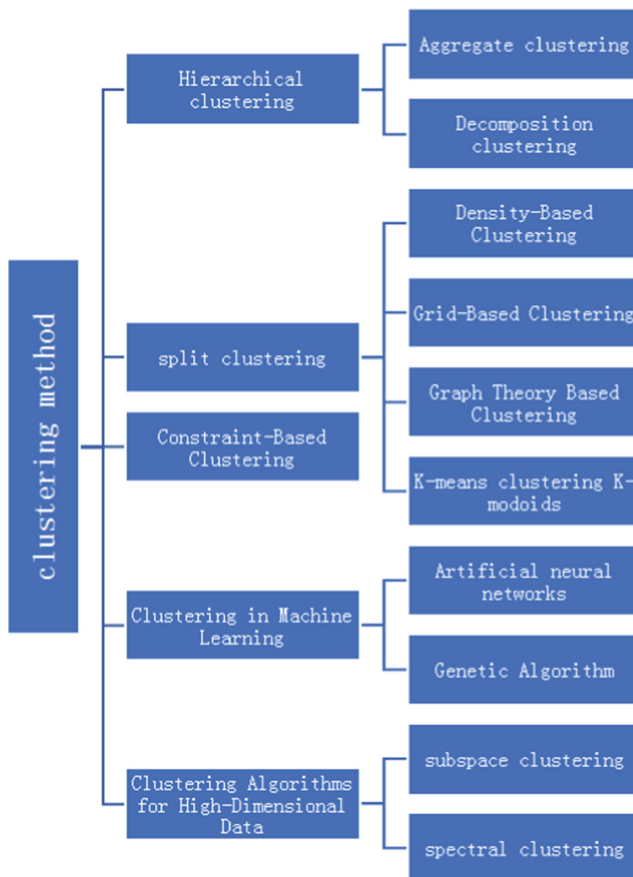
### 2.2 Advantages of Data Analysis Algorithms

First, the diversity and instability of big data determine the limitations of modeling techniques. Given the amount and complexity of big data, it is often difficult to build accurate models based on prior knowledge. Computational intelligence techniques such as evolutionary computing and cluster computing are knowledge-independent and do not require precise problem models, but they do require direct analysis and data processing [6]. This feature is great for analyzing large amounts of data. Big data analysis often involves changes in the environment. This is due to changes in subjective and

objective factors such as the system itself, user needs and goals. Traditional methods are difficult to adapt to changes in the environment. Evolutionary algorithms represented by genetic algorithms have possible solutions between generations. Evolutionary algorithms represented by genetic algorithms have potential solutions between generations [7]. It continuously optimizes the fitness of the population according to the environment, making it easier to adapt to changes in the environment.

Second, accuracy is an important factor in big data. The need to address and manage uncertainty stems from the influence of stochastic factors such as data collection methods, system changes, natural environment, and the specific nature of big data. Therefore, the extraction of potentially uncertain data has become an important topic in today's big data analysis. Computational methods such as fuzzy logic and rough sets improve the objectivity and interpretability of analytical results [8].

Finally, the breadth and complexity of big data means that spatiotemporal computations can be very expensive and accurate solutions cannot be obtained in acceptable



**Fig. 2.** Clustering Algorithms for Data Mining

time [9]. Intelligent computing methods include heuristics that solve problems by mimicking the psychological characteristics of humans and other creatures. These are highly self-organizing, adaptable, general and abstract. Many complex NP problems, such as combinatorial optimization problems, can be solved quickly and roughly, which provides an effective tool for solving large and complex problems, as shown in Fig. 2 [10].

### **3 The Importance of Intelligent Color Matching of Cultural and Creative Products**

Today's commodity world is rich and colorful. In this unpredictable market environment, cultural and creative products can only adapt to the ever-changing trends. The packaging, price, advertisement, quality, etc. of cultural and creative products will change, but no matter how they change, cultural and creative products of the same brand always have a main line supporting each other. However, some cultural and creative products are easy to find among thousands of similar cultural and creative products because they share a common gene (the core value of the brand), which is the brand's DNA. The DNA of the brand is eternal, it is the core element of the brand, and represents the essence of the brand [11]. Brands often reflect the most unique and valuable parts of their core values. In today's visual world, many cultural and creative brands use color to form their core values. Color has gradually become a symbol of brands, and consumers can easily identify some brands by color, but most brands are indistinguishable to consumers. This is the role of the brand's core values. An important indicator of successful brand management is the evaluation of whether it has core values. People receive the color information of cultural and creative products from the outside world, and form the image of the color of cultural and creative products after colliding with personal knowledge and experience. The theory of color value has a long history, such as China's Five Elements and Five Colors Theory, which is the philosophical thought and symbolic meaning of color in the cultural accumulation. Through in-depth research and excavation of color, people have gradually established many scientific and rigorous color science theoretical systems including color intent [12].

### **4 Design Process of Intelligent Color Matching System for Cultural and Creative Products**

The two most commonly used color imagery survey methods are psychophysical measurements and physiological experiments. When looking for the relationship between color and perceptual words, it is confirmed that the semantic difference method is an effective method to evaluate the relevance between different concepts, and then the semantic difference method is used as a general method in the study of color imagery, which describes the investigation and research methods of imagery., the color image survey methods include free association idea, restricted association idea and semantic difference method, and the research methods of color preference cases include selection method, scale evaluation method, etc. In subsequent studies, psych perceptual assessments and physiological measures are generally used together for color imagery research.

Psychological perception evaluation uses more semantic difference method, and combines factor analysis, cluster analysis and other statistical methods to process evaluation data; physiological test mainly uses instruments to obtain relevant physiological data of people, including eye tracker, EEG, ECG, etc., of which eye trackers are the most commonly used tests.

Consider safety experience in color design, use Likert scale method, visual and physiological experiments to obtain people’s perceptual image data, and use TOPSIS method to comprehensively analyze various factors. It will also obtain product color semantics through market research, introduce engineering, and use a combination of quantitative and qualitative methods to obtain users’ tendency to color imagery and establish a color semantic space. After using the semantic analysis method to complete the survey of furniture color imagery, the eye tracker was used to capture the subjects’ gazes to obtain the subjects’ perception of the sample color, so as to support and verify the subjective survey results of color imagery.

In order to solve the problem of non-standard color imagery, the researchers used the semantic difference method to quantify the color intention to obtain more standard and unified color image data, that is, the color image scale. The now well-known color image

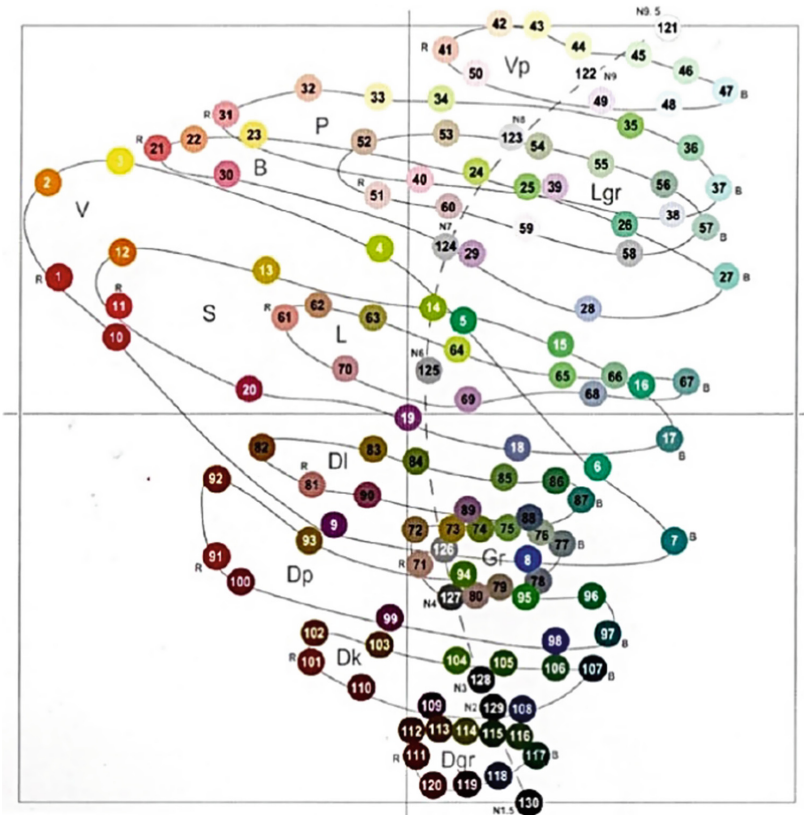


Fig. 3. Monochromatic Imagery Scale Map

scale map was designed by Japanese researchers. The vertical axis represents Soft-Hard, and the horizontal axis represents Warm-Cool. There are color blocks of different colors distributed on the two-dimensional coordinates, and these color blocks are arranged in space with a certain pattern. Inside, the image clearly expresses the emotions that these colors represent. Designers can choose the appropriate color scheme through the image scale map, as shown in Fig. 3.

To understand how people perceive color, researchers use a variety of intelligent algorithms. The main methods are gray theory, fuzzy theory, neural network, sensitivity engineering, mathematics, physical and chemical theory, joint analysis and so on. For example, by combining the emotional evaluation method of grey theory and the aesthetic evaluation method of color harmony, a quantitative evaluation and search method based on the RGB color system is proposed to evaluate a suitable product color scheme. A grayscale clustering process is introduced to predict overall color image ratings for multicolor products. Using fuzzy neural network theorem and grey theory, two color scoring models are developed and applied to a case study of electronic door lock design. The evaluation model of product shape and color is established by combining neural network, quantitative theory and perceptual engineering theory. For example, building a gray correlation model can evaluate the color scheme of a product in multiple images. Describes the observer's effect on the perception of an image due to differences in color area and spatial location on the sample. This study designs an intelligent color matching system for cultural and creative products based on image evaluation adjectives. The usefulness of each element is obtained by collecting the image satisfaction of the samples from the experimenters, and a color evaluation model is constructed. Use conjoint analysis to find out what each metric means. We created a Dun's Relevance Model correlation model based on the known basic color sample image score data to predict and calculate the image score of the target product.

The current research mainly focuses on color design from three directions: harmony theory, epistemology and image. Related research uses a variety of intelligent algorithms, such as particle swarm optimization, gray theory, neural network, genetic algorithm and so on. On the basis of semantics, style classification is used to solve the problem of inconsistency of word description semantics. Collect product color droplets to extract color balances and use Access to create color scheme libraries. Due to the uncertainty of the user's image description, the system needs to introduce fuzzy expectations into the color image of the product, and establish a fuzzy optimal design model. Particle swarm optimization is used to obtain optimal solutions optimized according to user expectations. The scheme can meet the actual expectations and preferences of users, convert the original color case database scheme on the market into a 3D color model, and use a derivation algorithm to obtain a new color scheme. The rule used in the algorithm is to find identical, adjacent, contrasting and complementary colors for a particular color, get a different monochromatic algorithm, and continue to recombine. The system uses fuzzy theory to create a set of color-identifying words and determines the proportion of each color in the semantics. Considering the area and position of each color in the scheme, a more accurate color harmony evaluation mechanism is proposed. Interactive Genetic Algorithms are used to determine the execution of the patterns, and the results of user selections are also stored as data for optimizing the next algorithmic model.

Studying how to choose colors according to the principles of color adjustment. Then, based on the mapping relationship, a color assignment method for assigning two or more colors to each part of the machine tool is proposed. Based on the above two methods, create color design samples. Finally, design a prototype system and integrate it, starting from the color design thinking of industrial designers, find a good product color frame that matches the consumer's target preferences, create a color frame set, and use the method of color allocation area division as Your product is assigned color and the final color image is evaluated.

## 5 Conclusion

Data analysis algorithms are changing with each passing day, especially in the manufacture of cultural and creative products. The addition of computer technology has greatly shortened the production time and ensured the production quality. Data analysis algorithms improve designer productivity and output designer ideas with high precision. Based on the understanding of data algorithms, this paper uses artificial intelligence and big data models and other data algorithm analysis tools to establish and optimize the design of the intelligent color matching system for cultural and creative products, which has made certain contributions to improving the color matching efficiency of cultural and creative products. The intelligent color matching system for cultural and creative products improves the design accuracy and efficiency of cultural and creative products. From an aesthetic point of view, the color design of cultural and creative products should be harmonious and beautiful in color matching to meet the aesthetic needs of consumers. However, when designers design product colors, they often need to try several color matching attempts and compare the color schemes in order to obtain the best color matching effect. Therefore, it is very necessary to use data analysis algorithms to help cultural and creative designers complete color matching design and improve work efficiency.

**Acknowledgments.** This work was financially supported by the key scientific research project of Sanda University 'Art resources and cultural creative integration design under the empowerment of artificial intelligence' fund.

## References

1. AlHabdan, A., AlShamrani, A., AlHumaidan, R., AlFehaid, A., Eisa, S.: Color matching of universal shade resin-based composite with natural teeth and its stability before and after in-office bleaching. *Int. J. Biomater.* **2022** (2022)
2. Cui, P., Xue, Y., Liu, Y., Sun, X.: Intelligent manufacturing of color blended yarn: color matching algorithm and manufacturing process through computer numerically controlled ring spinning system. *J. Eng. Fibers Fabr.* **16** (2021)
3. Alqahtani, F.: In-vitro efficacy of two spectrophotometer-based shade matching systems for color matching of ceramics after artificial accelerated aging Eficácia in vitro de dois espectrofotômetros baseados em escalas de cor para avaliação de cor de cerâmicas após envelhecimento artificial acelerado. *Braz. Dent. Sci.* **24**(2) (2021)



4. Ohio State Innovation Foundation: Patent issued for image color data normalization and color matching system for translucent material (USPTO 10,387,746). *Biotech Bus. Week* (2019)
5. Anonymous. New system allows custom liquid-color matching by molders. *Plast. Technol.* **65**(7) (2019)
6. Axalta Coating Systems IP Co. LLC: Patent application titled “systems and methods for matching color and appearance of target coatings”. Published Online (USPTO 20190172228). *Comput. Wkly. News* (2019)
7. Wang, D., Liu, C., Chu, F., Wang, Q.-H.: Full color holographic display system based on intensity matching of reconstructed image. *Opt. Express* **27**(12) (2019)
8. Justiawan, et al.: Comparative analysis of color matching system for teeth recognition using color moment. *Med. Devices (Auckland, N.Z.)* **12** (2019)
9. Walowit, E., Tahoe, L.: Determination of individual-observer color matching functions for use in color management systems. *Electron. Imaging* **2019** (14) (2019)
10. Sysuev, I.A., Varepo, L.G., Trapeznikova, O.V.: Revisiting measuring colour gamut of the color-reproducing system: interpretation aspects. *J. Phys. Conf. Ser.* **998**(1) (2018)
11. Ramanna Pavithra, K.: Standardizing shade matching with technology-based shade matching systems and color formulation. *Int. J. Prev. Clin. Dent. Res.* **5**(1) (2018)
12. Patents: Patent application titled “system, method and apparatus for performing colour matching”. Published Online (USPTO 20160216156). *Polit. Gov. Week* (2016)



# Design of an Intelligent Educational Virtual Classroom Platform Based on the Mobile Terminal APP of the Internet of Things

Xu Xing<sup>1</sup>, Xilin Zhang<sup>1</sup>(✉), and Qiong Zhang<sup>2</sup>

<sup>1</sup> Department of Foreign Languages, DaLian University of Science and Technology, Dalian 116052, Liaoning, China  
77138119@qq.com

<sup>2</sup> Department of Education and Psychology, Hubei Engineering University, Xiaogan 432100, Hubei, China

**Abstract.** While traditional education and curriculum models are under increasing pressure, the development of intelligent terminal technology based on the Internet of Things provides practical technical support for new curriculum models. This is the need for social education development and the trend of platform curriculum development. The purpose of this article is to study the intelligent education virtual classroom platform (CP) based on the mobile terminal APP of the Internet of Things. This article introduces the basic technology and scientific principles involved in the development of an intelligent education virtual CP based on the Internet of Things mobile terminal APP, introduces the system development process, and introduces the database and web server technology used in the background. Experimental research shows that after the use of virtualization, the response time is significantly reduced by about 20% when the number of concurrent users is large, which effectively improves the response speed of the intelligent education virtual classroom system. Prove the effectiveness and correctness of the system.

**Keywords:** Internet of Things · Mobile Terminals · Virtual Classrooms · Education Platforms

## 1 Introduction

With the advancement of science and technology, the development of mobile terminal applications has become more and more intelligent, and the applications have become more and more extensive, which has brought reforms to the education platform. In the late 20th century, with the popularization of the Internet in developed regions such as Europe and North America, it provided an opportunity for smart terminals to enter the campus education system [1, 2]. The combination of the Internet of Things and mobile terminal technology can improve the effectiveness of school education, improve the classroom learning environment, increase students' enthusiasm for education, and also improve the quality of college teaching.

In the research on the design of the intelligent education virtual CP based on the mobile terminal APP of the Internet of Things, many scholars have studied it and achieved good results. For example, Lu X has launched multiple mobile learning platforms, and students can access educational resources [3]. Liu T C uses servers to optimize the business needs of data interaction. The integration of technology increases the impact of classroom learning and improves student participation through classroom changes [4]. In this regard, it is of great significance to study the intelligent education virtual CP based on the mobile terminal APP of the Internet of Things.

This article conducts an in-depth and complete analysis of the Android system from the perspective of the system and application space, and finally gives the basic steps for setting up the environment. According to the system design, this paper uses SSH technology and VMware vSphere software to realize the various functional modules of the platform, and introduces the realization of the intelligent education virtual CP of the Internet of Things mobile terminal APP in detail.

## 2 Research on the Design Method of Virtual Classroom Platform

### 2.1 Teaching Platform Requirements

#### Sign-in management

Student sign-in (electronic roll call) is a very valuable function of the entire software. The specific functions implemented include student sign-in, server-side receiving data, and client-side query and statistics sign-in data.

Student sign-in: Simulate sign-in behavior by entering the user name and password to log in to the software. If the network cannot be connected at the time, the student information is temporarily stored in the student's local sqlite database, and will be actively submitted to the server after the network is restored, and the local computer's data will be deleted at the same time.

Server-side receiving data: The expected goal of this system is to integrate with the school's existing educational administration management system. The student sign-in data is sent to the server side through a Web Service connection in a network environment, and the server queries and stores the data, The data requested by the client is returned to the client after the operation is over [5, 6].

Client query statistical check-in data: Students can check the check-in status of each subject according to the student number, and the teacher can check the check-in status of each course according to the course number. The client side displays the data records returned by the server in the optimal way.

#### File sending and receiving

The realization of document transmission and reception has practical significance. In this module, the teacher is the document transmission initiator and the students are the document receiver. The specific functions implemented include selecting files, sending files, and receiving files.

Select file: Since the Android API does not provide the function of a file browser, you can select a file through a file browser written by yourself.

Sending files: After confirming the sending file object, select the receiving object of the file, and transmit the file through multicast or TCP connection.

Receive files: After confirming, start to receive files, which are saved in the SD card directory of the student's mobile phone by default.

#### Non-functional requirements

Stability requirements: For this system, especially the screen allocation module, the application itself should not occupy too much mobile phone system resources. However, when the application runs for too long, it will cause the mobile phone to run slowly and even memory leaks, resulting in The process becomes slower [7, 8].

Security requirements: Since the system needs to run in an environment with a network, the security defense of the mobile phone must be strengthened during use to avoid virus infection. When processing user login, only simple identity verification is required. Identity verification is the same as the school's existing educational administration management system.

Scalability requirements: The software application architecture should be extensible. When the current functional modules are small, the software performance can meet the requirements. When the functions continue to increase, we should consider the further expansion of the software. Some functional interfaces are Can be expanded.

Interface requirements: The limitations of mobile devices in terms of hardware configuration and network speed make the user interface design of mobile applications more important than traditional WEB applications and desktop applications. Therefore, when designing mobile applications, you need to follow these additional rules: simple and easy to learn, functionally focused, and interactive.

## 2.2 Design of the Virtualization Platform of Distance Education

### (1) Physical layer

This layer contains some low requirements for computer applications. Through virtualization and other system storage technologies, computers can be used to connect different types of storage devices and data servers to realize the integrated management of big data. Central control and status monitoring of storage devices are identified. At the same time as capacity expansion, it is essentially a work-based allocation and utilization [9, 10]. Using the lower-level vSphere platform architecture, using vMotion, HA and other services to achieve high-density remote learning platform search clusters, you can easily implement the largest and most efficient system on the original application server platform.

### (2) Logic layer

The same layer provides the upper layer with a view of sharing control between different functions. Through task management, security management and other design tasks, the lower-layer storage and upper-layer applications are seamlessly integrated and connected to realize the collaborative work between multiple storage devices.

### (3) Application layer

The application resource manager stores many application modules on the Web server, and can choose different learning resource methods for flexible configuration and customization, so that applications and behavior teaching can work together, so that teachers can be fully expressed in teaching. Ordinary users can also choose resources according to their training needs to assist in training and testing. The

performance of the application layer directly determines the quality of the entire system, so the design of this layer is very important.

(4) Access layer

The access layer is the entrance for users to receive services, and users can access the server through a browser to directly run educational courses. According to the user's authority, users can be divided into 3 main roles: system administrator, source administrator and simple user. The program administrator has the highest authority and can perform tasks such as program setting, information statistics, user management, and distribution. You can publish and retrieve the source and other permissions.

(5) Construction of vSphere virtualization platform

In this powerful architecture, VMware vCenter Server is used to manage the entire architecture. Administrators can connect to VMware vCenter Server or ESX Server through various VI clients running on other computers in the corporate network and perform multiple control functions on them.

(6) VSwitch (virtual switch) network management

The network is the core of the power supply. All VMs in the cluster must communicate through the network. These functions are implemented through virtual switches. Standard vSphere switches detect smart virtual machines connected to their virtual ports and use this information to direct traffic to the virtual machines. Use personal Ethernet adapters to connect the virtual network to the physical network, and connect a standard vSphere switch to the physical switch.

## 2.3 System Detailed Design

(1) Exam management module

For the intelligent education virtual CP of the Internet of Things mobile terminal APP, the specific topic management module is mainly carried out by four aspects of work. The main task of daily teachers is the intelligent education virtual CP of the Internet of Things mobile terminal APP. The question bank is maintained. When the test is about to take place, the teacher logs in to the intelligent education virtual CP of the Internet of Things mobile terminal APP, selects the automatic test paper management module, automatically generates test papers, and then manually maintains the test papers that are automatically generated. To determine the final test paper [11, 12].

(2) Score management module

For the grade management module, it needs to associate the course with the relevant students first, so as to confirm the grade information and student information. For the student's performance information, the input is mainly the responsibility of the college's academic affairs office teacher or the teacher. The general principle is that the college teacher is responsible for the input of courses opened by the school, and the courses taught by the college teacher are mainly responsible for the college teacher.

(3) Design and implementation of video classroom functions

For the teacher's intelligent education virtual CP, the video classroom management module is an important function. It mainly provides a learning platform

for students. Through this functional module, students can watch online teaching videos of various subjects online, and can also view the video. The downloading of the materials can effectively improve the students' self-study ability. In the video classroom management function module, the main functions include multiple aspects, mainly including video data search, video data paging display, video data online browsing and video data download. When the user clicks on the function module, the system first determines whether the user is logged in. If the user is not logged in, you need to return to the login page to input personal information.

## 2.4 The Quality Evaluation Algorithm of Teaching Platform

The MIN method represents the shortest distance. Namely

$$D_{KL} = \min_{i \in G_k, j \in G_L} d_{ij} \quad (1)$$

The group average method uses the average of the squared distances of all samples, and takes the average of the distances between each sample in the  $G_K$  class and each sample in the  $G_L$ , which can represent the distance between classes more accurately than the MIN method and the MAX method.

$$D_{KL} = \frac{1}{n_k n_L} \sum_{i \in G_K, j \in G_L} d_{ij}^2 \quad (2)$$

## 3 Experimental Research on the Design Method of Virtual Classroom Platform

### 3.1 Platform Performance Test Analysis

The test and analysis of the platform are carried out by means of stress testing. The learning platform is deployed on a physical server Tomcat, and two virtual machines are created on the same physical server at the same time to run the learning platform Tomcat server. Load balancing between. Use Loadrunner to simulate concurrent access to the learning platform Tomcat server in the two server architectures, and record the test data.

### 3.2 Experimental Method

System response time is an important parameter for evaluating the service quality of a Web system. Therefore, this article will test and compare the system before and after virtualization according to the system response time. Response time refers to the time elapsed after the user sends a request to receive the result returned by the server. The test method in this article is to simulate the concurrent operation of the same request from multiple users, gradually increase the number of concurrent operations, and record the average response time of the original system and the virtualized system respectively.

## 4 Experimental Research and Analysis of Virtual Classroom Platform Design Methods

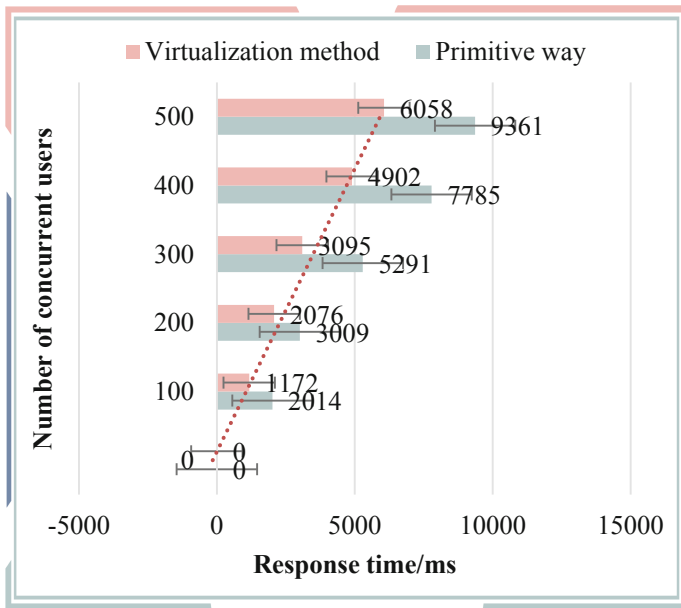
### 4.1 System Test Analysis

In real life, user access generally obeys Poisson distribution, but Jmeter does not currently support Poisson distribution simulation, so in order to simplify, this article adopts a uniform distribution simulation method, all user requests are sent out evenly within one second To visit the homepage of the site. The number of concurrent users ranges from 0 to 500, and the speed of increasing the number of concurrent users by 100 is performed ten times. Finally, the test results of the original system and the virtualized system are calculated, as shown in Table 1.

**Table 1.** Response time comparison

Number of concurrent users	Primitive way	Virtualization method
0	0	0
100	2014	1172
200	3009	2076
300	5291	3095
400	7785	4902
500	9361	6058

It can be seen from Fig. 1 that after the virtualization method is adopted, the response time is significantly reduced by about 20% when the number of concurrent users is large, which effectively improves the response speed of the system. Prove the effectiveness and correctness of the system.



**Fig. 1.** Response time comparison

#### 4.2 Response Time Test Analysis of the Course On-Demand Page

The test case is selected as the course on-demand loading page and the score query function. Among them, course on-demand is the most important function of the learning platform, and the performance of concurrent access affects the user experience. The score query requires a large number of searches in the database, and it is necessary to test the response time of the platform. The experimental results are shown in Table 2.

**Table 2.** Comparison of response time of course on-demand page

Visits	Before virtualization	After virtualization
0	0	0
50	327	739
100	874	715
150	1312	869
200	1993	1126
250	3126	2015



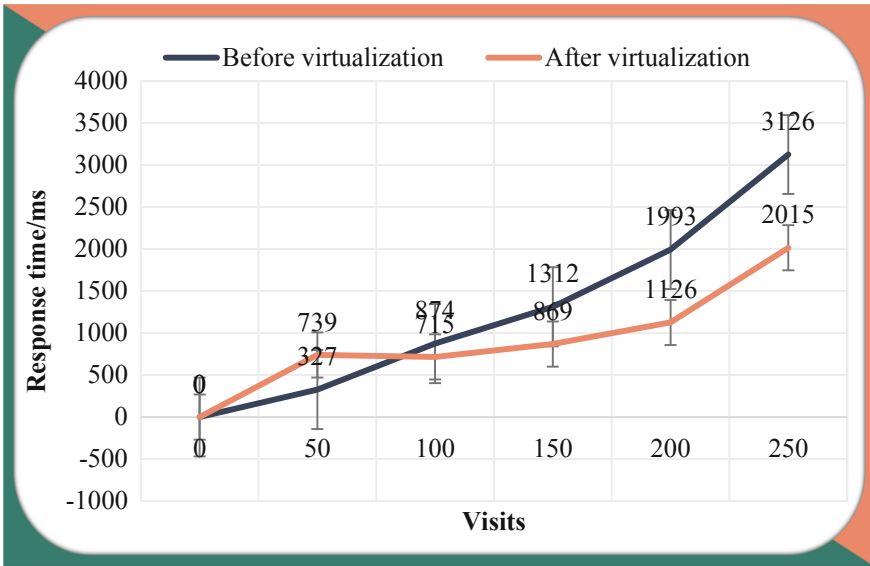


Fig. 2. Comparison of response time of course on-demand page

As shown in Fig. 2, according to the line chart data, it can be analyzed that after virtualization, under the same physical server, a stress test is performed. After virtualization, the response time of concurrent access is reduced. When the number of concurrent access is large, the response time is the reduction is obvious. By comparing the maximum number of TCP connections, CPU utilization, and memory utilization before and after virtualization, it can be seen that virtualization technology can improve resource utilization.

### 5 Conclusions

This article has carried out the whole process of demand analysis, system design, implementation and testing of the intelligent education virtual CP of the Internet of Things mobile terminal APP. The demand analysis mainly gives the content of performance requirements and functional requirements. System design, The realization and testing are mainly aimed at the specific realization of the various functional modules, system architecture and network architecture of the system.

### References

1. Liu, L., Xu, J., Huan, Y., et al.: A smart dental health-IoT platform based on intelligent hardware, deep learning, and mobile terminal. *IEEE J. Biomed. Health Inform.* **24**(3), 898–906 (2020)
2. Ning, W., Chen, X., Lan, Q., et al.: A novel wiki-based remote laboratory platform for engineering education. *IEEE Trans. Learn. Technol.* **PP**(3), 1 (2017)

3. Lu, X., Pan, Z., Xian, H.: An integrity verification scheme of cloud storage for Internet-of-Things mobile terminal devices. *Comput. Secur.* **92**, 101686.1–101686.8 (2020)
4. Liu, T.-C., Lin, Y.-C., Wang, T.-N., Yeh, S.-C., Kalyuga, S.: Studying the effect of redundancy in a virtual reality classroom. *Educ. Tech. Res. Dev.* **69**(2), 1183–1200 (2021). <https://doi.org/10.1007/s11423-021-09991-6>
5. Salazar, A.N., Nobrega, F.L., Anyansi, C., et al.: An educational guide for nanopore sequencing in the classroom. *PLOS Comput. Biol.* **16**(1), e1007314 (2020)
6. Radovan, M.N.: Acceptance of technology and its impact on teachers' activities in virtual classroom: integrating UTAUT and Coi into a combined model. *Turk. Online J. Educ. Technol. TOJET* **16**(3), 11–22 (2017)
7. Kuznetcova, I., Glassman, M., Lin, T.J.: Multi-user virtual environments as a pathway to distributed social networks in the classroom. *Comput. Educ.* **130**, 26–39 (2019)
8. Xin, Y., Zuo, X., Huang, Q.: Research on the construction of seamless learning platform based on open education. *Asian Assoc. Open Univ. J.* **13**(1), 88–99 (2018)
9. Zhou, X., Tang, L., Lin, D., et al.: Virtual & augmented reality for biological microscope in experiment education. *Virtual Real. Intell. Hardw.* **2**(4), 316–329 (2020)
10. Kolluru, S., Varughese, J.T.: Structured academic discussions through an online education-specific platform to improve Pharm. D. students learning outcomes. *Curr. Pharm. Teach. Learn.* **9**(2), 230–236 (2017)
11. Yuan, Y.: To build a solid education platform and give full play to the educational role of student apartment – a case study of Sichuan university of arts and science. *Int. J. Soc. Sci. Educ. Res.* **3**(5), 40–44 (2020)
12. Lee, J.Y., Patel, S.J.: Innovating business model in the higher education industry: a platform-based approach in university career services. *Ind. High. Educ.* **34**(2), 91–99 (2020)



# Design and Development of English Autonomous Learning Database Framework Based on Artificial Intelligence Algorithm

Ximeng Wen and Shaoxian Hong<sup>(✉)</sup>

Hainan College of Vocation and Technique, Haikou 570216, Hainan, China  
maryhong686868@163.com

**Abstract.** With the continuous development of computer science and technology, artificial intelligence has become a hot topic in the field of computer research, and has attracted more and more attention in practical applications. This paper analyzes and designs the English autonomous learning database based on the existing construction ideas. First of all, this paper introduces the theoretical basis of autonomous learning, expounds the learning characteristics and functions of English autonomous learning, and then studies the artificial intelligence algorithm, and designs the Database framework of English autonomous learning based on the algorithm, and tests the operation of the database framework. The final test results show that the independent English learning system runs smoothly with short operation time and delay time, both within 3 s. At the same time, the database can also be compatible with about 1000 users, with low CPU consumption. This shows that the system can fully meet the needs of autonomous learning.

**Keywords:** Artificial Intelligence Algorithm · English Autonomous Learning · Database Framework · Framework Design

## 1 Introduction

With the continuous development of computer technology, network, digital information processing and other information technology is more and more concerned by the country and people. On the Internet, various data types emerge in an endless stream. How to efficiently use these databases to meet students' learning needs has become a hot topic [1, 2]. Artificial intelligence is a discipline that produces a new pattern after combining new things with the existing environment and is widely applied in practical fields to solve practical problems. It involves biotechnology, neural network theory, information processing and many other aspects, and is of great significance to human social life and the development of science and technology [3, 4].

There have been many researches on artificial intelligence. The research on artificial intelligence started earlier in foreign countries, and there have been rich and mature research results. In the United States, its researchers put forward two learning platforms based on neural network and natural language processing technology [5, 6]. Domestic scholars have carried out theoretical and practical research on English autonomous

learning system based on artificial intelligence algorithm. Some scholars put forward the construction of English autonomous education system based on neural network by analyzing the current situation of students and teachers in China. Some scholars study the teaching mode of artificial intelligence and apply the intelligent algorithm to the auxiliary teaching. In addition, some scholars combined natural language processing technology, ontology cognition, multidimensional data platform, self-organization and other methods to construct an intelligent network framework with high practical value [7, 8]. Through the analysis and summary of relevant literatures at home and abroad, some scholars put forward that optimizing the keyword table structure of English word database by using neural network is the core idea of realizing artificial intelligence learning system. The above research has laid the research foundation for this paper.

With the continuous development of computer technology, computer science is increasingly mature, its powerful function has been widely used in various fields. This paper takes the theory and method of artificial intelligence as the basic framework to conduct the research of English autonomous learning. Based on the application of intelligent algorithm in neural network and relevant literature and journal papers, the platform is developed and designed with reference to the excellent research results at home and abroad. At the same time, the model and implementation steps of language feature classification system based on artificial intelligence algorithm are summarized by combining with the actual case analysis. On this basis, a complete set of student-oriented curriculum database, teaching test case database and exercise database is put forward.

## **2 Discussion on the Design of English Autonomous Learning Database Framework Based on Artificial Intelligence Algorithm**

### **2.1 Independent English Learning**

#### 1) Theory

Humanistic psychology emphasizes the subject status of students in classroom teaching and attaches importance to the needs, emotions and will of students in the learning process. Social cognition theory is developed based on cognitive psychology. It makes an in-depth study of the learning process and puts forward that knowledge acquisition is mainly discovered by learners themselves, which is conducive to studying autonomous learning from students' inner psychological process. Operationalism does not pay much attention to the internal adjustment process of learners, but focuses on the influence of external environment on independent learning. The concept of "autonomous learning" was first proposed by a famous American psychologist. Individuals can change their original knowledge and experience through their own efforts after being stimulated by the external environment. Therefore, students make use of various resources and information spontaneously and actively and carry out activities in a certain order, thus forming a unique and stable structural form [9, 10].

## 2) Learning Characteristics

Accumulation of knowledge. After a certain period of time, each learner will form his own unique and valuable content. These resources are determined by the characteristics of learning difficulty, uncertain factors and difficult storage.

Students can communicate, cooperate and discuss with each other. Everyone may have different levels of language proficiency, and there are also cognitive conflicts or personal emotional experiences under different national cultural backgrounds.

Autonomy of learning. In traditional teaching, students passively accept the knowledge provided by teachers and do not freely play and express their own personal characteristics such as needs, opinions and opinions. In the process of English self-study based on artificial intelligence algorithm, various and targeted solutions or suggestions can be put forward according to students at different levels [11, 12].

Diversity of learning objectives. First of all, it is necessary to correctly understand the content of the article to a certain extent. Second, it is necessary to solve the problem through independent choice and finally achieve the expected results. Finally, it is necessary to have self-analysis ability and innovation spirit. The goal of autonomous learning is to build a learning environment with high flexibility, which can reflect students' personality characteristics and development potential and can adapt to changing constantly, so that students can live in a relatively loose, active and challenging spirit and creativity without affecting their future growth.

## 3) The role of independent learning

In traditional English teaching, most students pay more attention to the mastery of grammar and words, but ignore their application. This leads many students to develop the "dumb" mindset, that is, when confronted with a problem, they only know that language cannot solve the problem. Autonomy, not knowledge in the traditional sense, is the most important thing in English learning. Therefore, it is of vital importance and one of the necessary prerequisites to develop a resource base system platform based on artificial intelligence algorithm, which is intelligent and humanized, has a certain depth of understanding, and can carry out higher level thinking ability training and self-study and other learning activities. In addition, attention should also be paid to the improvement of the existing data information processing process and way, so as to better meet the personalized learning needs of students and achieve richer teaching content.

## 2.2 Artificial Intelligence Algorithm

### 1) Basic principles

Artificial intelligence algorithm mainly refers to the artificial neural network as the basis, through the existing data set and knowledge processing, forming rules with autonomous learning significance. It uses the existing computer system, carries on the information processing and reasoning on the basis of the existing data, and realizes the fitting operation of the relation between the input and output vector and the eigenvalue. At present, intelligent machine learning methods are commonly used: based on statistical theory, neural network and so on. These traditional methods are

based on existing computing resources as the center and developed a new technical means, such as random search method, clustering algorithm (LDA) and so on, and some heuristic design based on artificial intelligence research, such as genetic algorithm and simulated annealing condition method. Artificial intelligence algorithm by simulating human behavior, thinking and learning, automatic completion of corresponding goals in the computer system. Artificial neural network can train a large number of organisms with specific information characteristics. It can combine human brain structure and brain function to form a new model, and achieve a certain degree of parallelism. At the same time, it can use neural mechanism to control the interaction force between individuals to achieve the optimal goal. It can also be through the simulation of human behavior and get some data in the computer system to automatically complete the corresponding target task.

## 2) Algorithm support

The application of BP algorithm has become a very important and active research topic of great significance in the field of computer science and artificial intelligence technology development. On the basis of neural network theory, different types of problems (such as locality or nonlinearity) and state space complexity in artificial systems are realized through training and design of training sample data sets. BP artificial neural network is a kind of nonlinear topological structure formed by connecting many neurons. It mainly includes three parts: forward hidden layer, weight hiding and back propagation. In the practical application of BP algorithm, a large number of training samples will be generated, and these data sets often have various complex relationships. The neural network is nonlinear, time - varying and polymorphic.

When the actual output of the network is inconsistent with the expected conclusion, the output error will be generated, and its definition formula is as follows:

$$F = \frac{1}{5}(e - u)^2 \quad (1)$$

By putting the data into the above equation, the relationship between the hidden layer and the output error can be obtained:

$$F = \frac{1}{5} \sum_{k=3}^l (s_k - e(net_k)) \quad (2)$$

It can be seen from Formula 1 and 2 that the errors of the network are closely related to the weights of each layer. Therefore, in order to reduce the estimation error of the network, it is necessary to adjust the weight of the network. The training process of BP algorithm is to modify the weight of the network and reduce the error. The purpose of adjustment weight is to reduce the error, so the adjustment of weight should be in the direction of gradient descent, and is proportional to the gradient descent of error.

### 2.3 The Effect of Database Construction on Independent English Learning

In traditional database, English learners can only through their knowledge to read and understand the article content, but with the development of computer technology, network information technology, more and more advanced, and the increasing demand for information, under the influence of factors such as the students need to look up information from different angles to get more useful information. Therefore, for the present, how to combine the existing teaching resources with the theory of artificial intelligence becomes the key point to solve the above problems, which is one of the important goals of data scalability. With the integration and development of network information technology and intelligence technology, as well as the research direction and methods of related disciplines are becoming more and more mature and scientific, the system model gradually forms a new theoretical architecture, realizing innovative improvement and optimization and upgrading of its application development research work. In addition, as computer hardware and software resources, database can provide students with more learning space and improve their independent learning ability.

## 3 Experimental Design of English Autonomous Learning Database Framework Based on Artificial Intelligence Algorithm

### 3.1 Design of Independent English Learning Database Framework

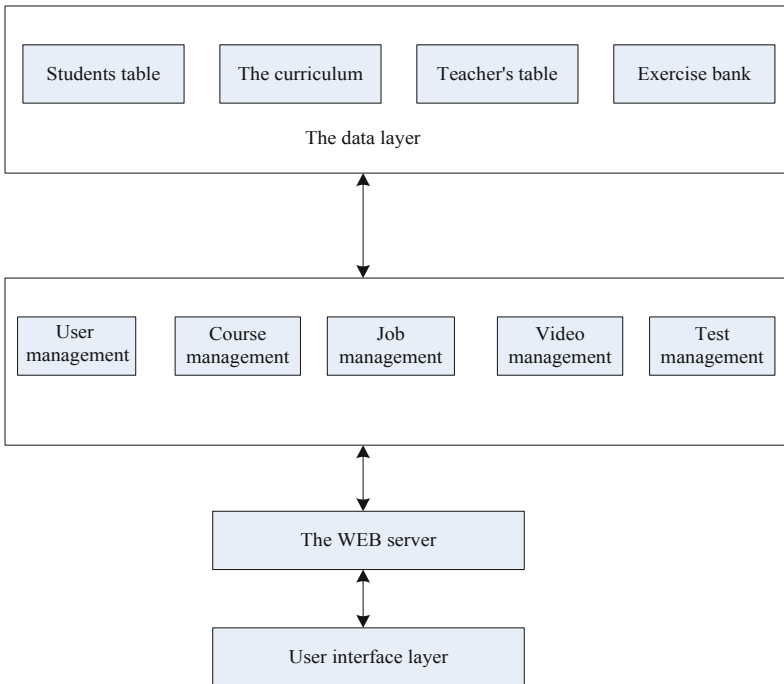


Fig. 1. The English Autonomous Learning Database Framework

As can be seen from Fig. 1, the independent learning database contains data modules of student table, class schedule, teacher table and question bank. When learners enter the theme, they will store their own relevant keywords and some relevant information. When students need to consult literature. Teachers according to the resources provided by content to search and retrieve relevant keywords, at the same time can also be based on the existing data analysis and selection and other auxiliary function, help students to understand the latest news and hot topics, and generates a complete and reasonable, scientific and effective and easy to understand the needs of the learners in the knowledge base of frame structure model. And combined with relevant theoretical knowledge and practical experience and their own discipline characteristic, professional design can meet the demand of the field and has the high feasibility of personalized learning environment and its framework, but also should make full use of network resources to realize data sharing in the process of learning, so as to provide students with a more convenient and quick, efficient and practical, the intelligent English teaching system can meet users' requirements for information retrieval and communication.

### 3.2 Running Test of English Autonomous Learning Database

In the development process, the database running test is an important stage, mainly including the following aspects: (1) data entry and modification. Convert a text file to a digital format. (2) Analyze the input information of learners. According to the needs of different users for personalized design and learning resources allocation, etc. Through the system function modules, performance requirements and other related technical parameters under the conditions of writing the corresponding operation program, complete the database logical structure test, and finally achieve the goal, but also need to pay attention to the input data must meet certain standard conditions to successfully output results.

## 4 Experimental Analysis of the Design of English Autonomous Learning Database Framework Based on Artificial Intelligence Algorithm

**Table 1.** The Database runs the test data

Test times	Operate time (s)	Delay time (s)	Occupying the CPU usage rate (%)	Number of concurrent users
1	2	1	1	1102
2	2	1	1	1215
3	3	2	2	1025
4	2	1	1	1130
5	3	1	1	1055



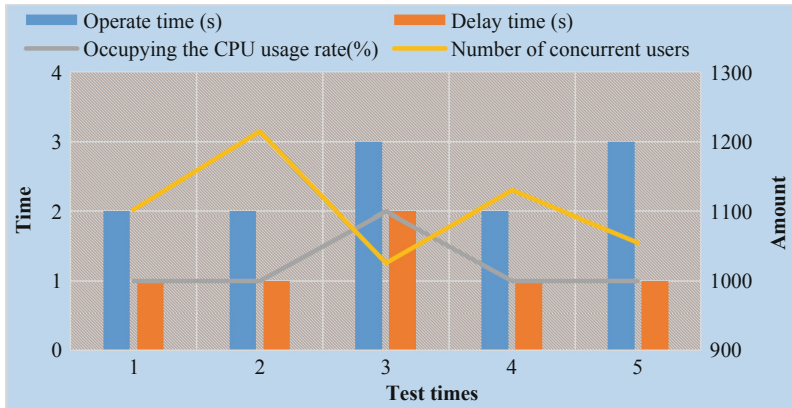


Fig. 2. The Database runs the test data

After the development and design of the English autonomous learning database based on artificial intelligence algorithm, the test results show that the system can better achieve the expected function. As can be seen from Table 1 and Fig. 2, the English autonomous learning system runs smoothly with short operation time and delay time, both within 3 s. At the same time, the database is compatible with about 1000 users and occupies low CPU consumption. This shows that the system can fully meet the needs of autonomous learning.

## 5 Conclusion

With the continuous development of computer technology, artificial intelligence has been widely used in various fields, and the autonomous learning method based on intelligent learning algorithm has been paid more and more attention. This paper analyzes and introduces it from the current research situation. It includes the current language knowledge and patterns in The language discipline in China, as well as the aspects based on data mining and classification, and puts forward corresponding strategy suggestions combined with the actual situation, and applies artificial intelligence technology to the design of autonomous teaching database.

**Acknowledgment.** Foreign Language Teaching Reform Research Project of Hainan Colleges and Universities in 2021-2022: Research on “AI” Learning Mode in Improving English Autonomous Learning Ability of Vocational College Students. Project Number: Hnjgw2022ZD-4.

Teaching Reform Research Project of Hainan Colleges and Universities in 2020: Research on English Curriculum Design and Development of Higher Vocational College Enrollment Expansion Based on Dynamic Language Curriculum Model. Project Number: Hnjg2020ZD-45.

## References

1. Radanliev, P., De Roure, D.: Review of algorithms for artificial intelligence on low memory devices. *IEEE Access* **9**, 109986–109993 (2021)

2. Alsaade, F.W., Aldhyani, T.H.H., Al-Adhaileh, M.H.: Developing a recognition system for diagnosing melanoma skin lesions using artificial intelligence algorithms. *Comput. Math. Methods Med.* **2021**, 9998379:1–9998379:20 (2021)
3. Abdollahizad, S., Balafar, M.A., Feizizadeh, B., Sangar, A.B., Samadzamini, K.: Using hybrid artificial intelligence approach based on a neuro-fuzzy system and evolutionary algorithms for modeling landslide susceptibility in East Azerbaijan Province, Iran. *Earth Sci. Inform.* **14**(4), 1861–1882 (2021)
4. Shaikh, J., Singh, R.S., Gelmecha, D.J., Ayane, T.H.: Artificial intelligence-based algorithm to track the probable COVID-19 cases using contact history of virus infected person. *Int. J. Comput. Appl. Technol.* **66**(2), 145–153 (2021)
5. García-Peñalvo, F.J., et al.: Application of artificial intelligence algorithms within the medical context for non-specialized users: the CARTIER-IA platform. *Int. J. Interact. Multimed. Artif. Intell.* **6**(6), 46–53 (2021)
6. Esfandyari, S., Rafe, V.: GALP: a hybrid artificial intelligence algorithm for generating covering array. *Soft. Comput.* **25**(11), 7673–7689 (2021)
7. Altamirano-Flores, J.S., et al.: Identification of HIV-1 Vif protein attributes associated with CD4 T cell numbers and viral loads using artificial intelligence algorithms. *IEEE Access* **8**, 87214–87227 (2020)
8. Bao, J., Liu, X., Xiang, Z., Wei, G.: Multi-objective optimization algorithm and preference multi-objective decision-making based on artificial intelligence biological immune system. *IEEE Access* **8**, 160221–160230 (2020)
9. Kaur, S., et al.: Medical diagnostic systems using artificial intelligence (AI) algorithms: principles and perspectives. *IEEE Access* **8**, 228049–228069 (2020)
10. Singler, B.: Blessed by the algorithm: theistic conceptions of artificial intelligence in online discourse. *AI Soc.* **35**(4), 945–955 (2020)
11. Goli, A., Moeyini, E., Shafiee, A.M., Zamani, M., Touti, E.: Application of improved artificial intelligence with runner-root meta-heuristic algorithm for dairy products industry: a case study. *Int. J. Artif. Intell. Tools* **29**(5), 2050008:1–2050008:20 (2020)
12. Yuan, K.-C., et al.: The development an artificial intelligence algorithm for early sepsis diagnosis in the intensive care unit. *Int. J. Med. Inform.* **141**, 104176 (2020)



# Devise and Application of Intelligent Guidance System in Exhibition Space

Heng Shen (✉)

Xi'an Fanyi University, Xi'an 710105, Shanxi, China  
524237276@qq.com

**Abstract.** In the exhibition space, the guidance system is the link to sort out the complex space environment, so that visitors can get better communication with exhibits and space environment. The research goal of this paper is to design, apply and study the intelligent guidance system in the public display space, improve the singleness and imperfection of the traditional guidance system, and provide diversity and possibility for the design of guidance system in the display space. Starting with the classification characteristics of display space and traditional guidance system, this paper summarizes the construction method of guidance system in display space and the design objectives, principles and Strategies of various elements of intelligent guidance system, so as to provide new ideas for the design of intelligent guidance system in the same type of display space and enhance the emotion of information dissemination.

**Keywords:** Display Space · Intelligent Guidance System · Design and Application

## 1 Overview

With the deepening of urbanization, there are more and more public display spaces in our life. In the face of complex spatial information, the role of guidance system becomes more and more obvious. With the advent of the mobile Internet era, the media for people to obtain and transmit information has changed greatly, especially the wide application of new media technology. The service object of the guidance system is people. The complexity of people determines that the needs of different audiences must be considered in the design of the guidance system. With the rapid development of information technology, the traditional visual guidance design has been unable to meet the changing space environment and people's needs, Relying on the Internet of things technology to create a functional integrated intelligent guidance system has become a breakthrough in the design [1]. According to the characteristics of different audiences, the use of intelligent display can more efficiently transmit information.

## 2 Characteristics and Current Situation Analysis of Guidance System in Exhibition Space

### 2.1 Characteristics of Exhibition Space

With the rapid development of science and technology, people have more and more ways to obtain information, display functions are becoming more and more diversified, and display categories are becoming more and more diversified. There are many classification methods of display space, and the frequently used classification methods are as follows Table 1.

**Table 1.** Shows the classification and characteristics of the space

Classification standard of display space		Type of presentation space	Characteristics of exhibition space	Example
Length of presentation	Short	Temporary display	Short display time, high flexibility and immediate removal after display	Auto show and Commodity Fair
	Long	Permanent display	Long display time, low flexibility and high cost	The Palace Museum
Display purpose	For the purpose of selling goods	Commercial exhibition space	Dense flow of people and prominent commodities	Shopping malls and supermarkets
	For the purpose of culture and education	Public welfare exhibition space	Strong logic between spaces	Art galleries and other types of Museums
	For entertainment	Entertainment display space	The space is relatively open and the streamline is free	Children's playground, KTV
	For publicity purposes	Publicity display space	It can accommodate many people to gather and discuss	Enterprise press conference and Expo

(continued)

**Table 1.** (continued)

Classification standard of display space		Type of presentation space	Characteristics of exhibition space	Example
	For demonstration purposes	Presentation space	There are plenty of seats and a main display point	Conference speech and show
Exhibition place	Indoor	Indoor display space	The space is relatively closed and the line of sight is easy to be blocked	Shopping malls, specialty stores, art galleries
	Outdoor	Outdoor exhibition space	Open space and good line of sight	Concert, Music Festival
Show the closeness of the space	Opening to the outside world	Open exhibition space	Open space and low line of sight resistance	Open square
	Half open	Semi open exhibition space	The space is relatively open, and some lines of sight are blocked	Square in front of shopping mall and street market
	Close	Closed display space	The space is relatively closed and the line of sight is blocked	Shopping malls, museums

## 2.2 Current Situation Analysis of Guidance System in Exhibition Space

In the form of field investigation and questionnaire survey, this paper makes a field investigation on several public exhibition spaces, and summarizes the shortcomings of the guidance system in the current exhibition space.

- 1) Incomplete information. In a complete guidance system, the environmental information carried by each sign is different and has internal continuity. At present, the guidance design of most display spaces only stacks the information on a single sign. The fracture of information transmission not only fails to guide the user to the corresponding place, but also increases the difficulty for the user to obtain and screen information, which is bound to cause the unclear dynamic line of some visitors[2].
- 2) Single function. The specific functions of the guidance system can be divided into basic indication function, environment optimization function, auxiliary management function, etc. The existing guidance system in the exhibition space only has basic

indication function, and does not have auxiliary management functions combined with Internet of things related technologies, such as real-time prompt of the flow of people in the exhibition hall, overall planning of the moving line of visit, etc.

- 3) Lack of characteristics. The guidance system can not only meet the functional needs of users, but also reflect the regional cultural characteristics. However, at present, the design form of the guidance system in most exhibition spaces is slightly single, and the modeling lacks characteristics, which can not reflect the artistic atmosphere and modern flavor of the modern exhibition hall [3].

### **3 Characteristics of Intelligent Guidance System in Exhibition Space**

#### **3.1 Scientific Orientation**

Due to the large number of elements and complex contents in the display space, the display system must have good logic, be able to logically summarize the complex elements in the display space, and make the information organized. As an indispensable part of the display space environment, the primary function and fundamental attribute of the guidance system is guidance and guidance. The scientific and reasonable guidance system can help the visitors entering this space to visit the whole space more conveniently and form a scientific and reasonable moving line. Therefore, scientific guidance is the foundation of the guidance system [4]. It is necessary to make an in-depth analysis of the space, take into account the visitors' psychological and behavioral characteristics, and integrate it into scientific guidance.

#### **3.2 Clear Hierarchy**

In the early stage of the design of the guidance system, the visiting habits of the audience should be taken into account, and then the display space should be divided into primary and secondary levels. Through the primary and secondary division of the exhibition content, the divided information should be collected and planned, and then transmitted to the audience through the guidance system. The information planned according to the hierarchy can help the audience have a clearer understanding of the space and quickly help the audience understand the structural relationship of the display space, so as to reach the destination more clearly and intuitively, save time and space utilization, and make the display space orderly.

#### **3.3 Unique Identifiability**

The guidance system is to better express the spatial information. It will use a lot of design language in the design to enhance the visual expression, so that the audience can quickly understand the whole display space. The guidance system of the display space with different nature content will also change. The individual logo derives different types and functions, which can bring different visual feelings to the audience through different colors, materials and shapes [5]. The characteristic guidance system can provide the audience with novel visual feelings and obtain a unique sense of experience.

### **3.4 Uniformity**

In a display space, the guidance system should be consistent with the content and style of the whole display space, and the guidance system must adapt to different types of display spaces. In the guidance system, each sign has different functions and meanings, but as an integral part of the guidance system, they should be unified to some extent. For example, we can express them by means of color, material and modeling, and apply them in the guidance system in combination with the characteristics of display space and display content. The unity of the guidance system will also greatly enhance the dissemination of the display theme.

### **3.5 Generality**

With the development of economy and the improvement of living standards, more and more people participate in all kinds of exhibition spaces. Among many visitors, we should not only consider the needs of normal visitors, but also the needs of disabled people. Compared with normal people, their demand for guidance system will be more urgent. The concept of universal design emphasizes paying attention to the equality of all people in space and time [6]. The design should truly reflect the “people-oriented” of the design according to the “concept of design for all” of universal design.

## **4 Design of Intelligent Guidance System for Public Exhibition Space**

### **4.1 Design Objectives**

By combing and analyzing the current situation of the display space guidance system, it can be seen that the existing guidance system not only has poor continuity of information transmission and lack of characteristics in design, but also has the prominent problem of single function.

To solve these problems, intelligent guidance system came into being. The intelligent guidance system is committed to solving the problem that people in the exhibition space have a better sense of experience in the sub space. Combined with relevant technologies of the Internet of things, it realizes an efficient man-machine feedback mechanism, makes the situation in the exhibition space clear at a glance by means of visualization, and displays the three-dimensional environment of the public space as a seamless plane environment, which is convenient for visitors to observe the public space intuitively, Realize the vision matching of public space. Improve people’s sense of efficient participation in this space, so as to achieve the organic combination of function and aesthetics, humanization and intelligence [7].

### **4.2 Design Principles**

First of all, the purpose of the design is “people-oriented”. The design of the intelligent guidance system in the display space should fully reflect the understanding and respect

for human nature, and establish a clear and perfect visual guidance combined with users' psychology and behavior habits, so as to form a benign man-machine interaction. Guide facilities of different indication and guidance types shall be set according to local conditions in different space environments to ensure the continuity of guide information during use, guide users to the correct position, and improve the interactive experience between users and intelligent guide system in combination with Internet of things and other technologies. For example, users can receive and sort guide information only through smart phone devices, Achieve the guidance goal.

Secondly, the diversity of users' requirements for the functions of the guidance system urges the intelligent guidance system to comb and integrate its functions in the design, so as to ensure the maximization of the system effect. This process integrates and develops appropriate functions based on users' behavior habits and demands [8]. While meeting the basic functions of instruction and guidance, it can also feedback information to users in real time, so as to establish a close human-computer interaction relationship.

Finally, compared with the traditional guidance system, the intelligent guidance system should not only consider the visual and modeling design in the physical space, but also consider the interactive interface design relying on new media. In the design, we should consider not only the needs of normal people, but also the special needs of the elderly and the disabled, but also their needs for information density. For example, the orderly arrangement of colors, symbols, graphics, text and other contents can improve the sense of experience in the interaction process, as shown in Table 2.

**Table 2.** Shows the combined shape of the space

Combined form of display space	Effect
Enclosure and septum	Enclosure and partition are important means of spatial segmentation. Rich enclosure can increase the density of display, so as to enrich the visual sense of visitors
Inclined space	The inclined space can give the audience a different feeling and provide a new form of exhibition

### 4.3 Design Strategy

The guidance system in the exhibition space can help visitors quickly get familiar with the venue environment, guide and standardize the behavior trend of visitors. Even in a completely unfamiliar environment, they can obtain information accurately and quickly through the intelligent guidance system.

Firstly, the intelligent guidance system consists of hardware and software. The hardware consists of three parts: central system processor, three-dimensional vision guide and induction prompt light in important parts of the exhibition hall. The corresponding software also consists of three parts: cloud terminal for data integration and analysis, memory card reservation system built-in in offline three-dimensional reservation machine and online mobile phone reservation program [9].



Secondly, the intelligent guidance system can not only meet the basic guidance function of instruction guidance, but also integrate functions such as real-time information feedback, guidance route guidance, crowd density analysis, etc. when using this system, visitors can scan the QR code next to the exhibits with their mobile phone when standing in front of any interested exhibits. When the central processing unit receives the information, The relevant information of this exhibit will be fed back to the mobile phone. When intelligent guidance signs are laid in public space, their alternative locations shall include the guidance information of the nearest destination as far as possible [10]. By constructing a multi-objective optimization model including the layout location and destination of guidance signs, after determining the intelligent guidance signs in public space, ensure the optimal value of the layout location of guidance signs in the intelligent guidance system in public space and the optimal distance between the location guided by guidance signs and the current location, so as to realize the intelligent guidance of the whole display space.

Finally, while the intelligent guidance system has advanced functions, it also needs to have the function of displaying the style of the whole exhibition hall and spreading art and culture. Therefore, its visual design needs to have cultural characteristics. For example, in the selection of graphic symbols and font design, we should not only convey their symbolic significance, but also combine with the space environment, add more artistic elements and reflect the spirit of the place; In color matching, appropriate color matching is used to give users correct psychological cognition. For example, red represents large traffic here, and blue represents small traffic. At the same time, the matching of red and blue also reflects an advanced sense of science and technology.

## 5 Conclusion

A perfect intelligent guidance system is complex, logical, scientific and humanized. It needs to consider all aspects of factors, coordinate with rational design planning and perceptual spatial cultural connotation, and make it more in line with the value of public display space itself. Therefore, the introduction of Internet of things, cloud computing, artificial intelligence and other technologies plays an important role in the design and development of intelligent guidance system [11].

In the exhibition space, the guidance system is the summary and interpretation of the architectural space layout and guidance. It summarizes and integrates the complex spatial information and transmits it to the visitors, so that the visitors can understand the structural relationship of the space in a short time and quickly reach the set destination. At the same time, the design of intelligent guidance system can also improve the competitiveness and influence of the brand to a certain extent. It is an important expression of the cultural elements of public display space. It should not only meet the basic function of guidance, but also create a cultural and artistic atmosphere, realize the intelligent management function, and give full play to the maximum utility of the intelligent guidance system of public display space. The design of intelligent guidance system based on Internet of things, cloud computing, artificial intelligence and other technologies has become the trend of display space guidance system design in the future. In this process, only based on the characteristics of the display space and rooted in the user's behavior

habits and emotional demands, can we create an intelligent guidance system with the perfect integration of function and characteristics, humanization and intelligence, so as to truly realize the people-oriented design.

## References

1. O'Hagan, L.: Instagram as an exhibition space: reflections on digital remediation in the time of COVID-19. *Mus. Manag. Curatorship* **36**(6), 610–631 (2021)
2. Ke, L., Xiaodong, L.: Study on the design of luxury exhibition space based on narrative theory. *E3S Web Conf.* **236**, 05060 (2021). EDP Sciences
3. Castellano, C.G., Lopez, M.: Inside and outside the exhibition space: the poetics and politics of Colectivo Quintapata. *Small Axe A Caribb. J. Crit.* **24**(3), 31–52 (2020)
4. Yuki, H.: The electronic exhibition space as a catalyst for engagement and inquiry. *Art Libr. J.* **44**(4), 174–179 (2019)
5. Shen, X., Shao, W., Zhang, Z., Xu, P.: Hotel intelligent guidance system based on ZigBee technology. *Microprocess. Microsyst.* **77**(9), 173–177 (2020)
6. Li, H., Miao, X.: Application of intelligent guidance system in American school district based on cognitive guidance algebra I. *J. Intell. Fuzzy Syst.* **35**(3), 2879–2885 (2018)
7. Lu, J., thahrocc Valaec: College multimedia teaching system devise and application evaluation. *Int. J. Electr. Eng. Educ.* **20**(3), 78–85 (2020)
8. Huang, Z.M., Wu, Q.F.: The design and implementation of P2P traffic intelligent guidance system. *Appl. Mech. Mater.* **475**, 798–802 (2014). Trans Tech Publications Ltd.
9. Zolotova, A.S., Likholet, E.V., Fedotova, N.N.: Current trends in the museums and exhibitions exhibition space formation. *IOP Conf. Ser. Mater. Sci. Eng.* **698**(3) (2019)
10. Wulandari, A.A.A., Fajarwati, A.A.S., Latif, F.: The relationship of exhibition space design and the success of delivering messages to museum visitors in Jakarta. *Humaniora* **8**(3), 219 (2017)
11. De Lima, J.T.M., Borges, E.A., de Sousa, L.N., Vilar, R.S., Akonde, S.B.: The scientific communication of the gondwana project in an exhibition space at the geodiversity museum–UFRJ [A disseminação científica do projeto gondwana no espaço expositivo do museu da geodiversidade–UFRJ]. *Anuario do Instituto de Geociencias* **44**(1), 89–93 (2021)



# Configuration of Crop Supply Chain Based on Gale-Shapley Algorithm

Wenting Kan and Bin Xiao<sup>(✉)</sup>

Guangzhou City Construction College, Guangzhou, Guangdong, China  
champagne\_bin@126.com

**Abstract.** In today's society, The circulation of crops between markets is a problem that cannot be ignored. How to improve the efficiency of crop circulation, reduce transportation costs, increase the added value of products and promote the development of crop trade has become an important research. Based on the Gale-Shapley algorithm, this paper studies the supply chain configuration of crop sources. Its purpose is to improve the circulation and trade of crops. This paper mainly uses the field inspection method and the questionnaire survey method to explore the supply chain scheme of crop supply. The survey results show that 74.4% of people agree with the role of crop manufacturing, distribution, supply chain design and inventory requirements in the crop supply chain.

**Keywords:** Gale-Shapley Algorithm · Crops · Supply Chain · Configuration Research

## 1 Introduction

At present, due to the continuous advancement of technology, The production and circulation of crops are also gradually changing to a modern way. In order to improve the production efficiency of crops, improve the dissemination and circulation of crops, and promote the sales of crops, it is necessary to adjust and improve their supply chain. Therefore, this paper studies the application of Gale-Shapley algorithm in it.

There are many theoretical achievements in the research of crop supply chain configuration based on Gale-Shapley algorithm. For example, Nasia Zacharia et al. propose an efficient implementation of Gale and Shapley's "Offer and Reject" algorithm when applied to the case of identifying candidates [1]. According to I. A. Omar, the exchange of information is essential for the coordination and integration of stakeholders and for the sustainable operation of the supply chain. Additionally, improved information sharing helps retailers spend less time finding alternative suppliers to meet unexpected demand surges during the pandemic [2]. R. Rajesh said that supply chain risk management involves management decisions under conditions of many uncertainties. As modern supply chains are complex networks with many vulnerabilities, resilient supply chains with built-in capabilities to deal with unexpected events can be very important [3]. Therefore, the research on the supply chain of crop sources has practical significance and value.

The Gale-Shapley algorithm is first studied in this paper, and it is analyzed and described. Secondly, it systematically analyzes the competition relationship of supply chain. Then, the synergy between crops and supply chain is expounded. After that, a brief discussion of supply chain performance is given. Finally, farmers' suggestions on the optimal configuration of crop supply chain are collected through investigation, and relevant conclusions are drawn.

## 2 Crop Supply Chain Configuration Based on Gale-Shapley Algorithm

### 2.1 Gale-Shapley Algorithm

Bilateral matching is to match the subjects in two different sets, and try to make the matching subjects of both sides get a satisfactory matching result.

Each party A subject applies for his favorite party B subject. Each Party B subject temporarily accepts applicants with higher priority according to its own priority to Party A's subject until the quota is used up. The remaining applicants will be rejected. The algorithm ends when no Party A subject is rejected [4, 5].

Gale-Shapley mechanism is a dominant strategy for Party A to express their true preferences, and this mechanism is stable.

Based on the traditional Gale-Shapley mechanism, combined with the actual matching mechanism in realistic bilateral matching problems, a constrained Gale-Shapley mechanism is proposed. According to the characteristics of the constrained Gale-Shapley mechanism, a recommendation problem in two situations is proposed: The first type of situation is when the main body of Party A has not selected any main body of Party B, and it is recommended according to its basic attributes such as risk preference and priority; the second type of situation is when the main body of Party A has selected some of the main parties of Party B. Under the circumstance, the priority of Party A's main body and the selected Party B's main body are combined, and then Party A's main body is recommended [6, 7].

Different real background problems, the matching mechanism will be different, so the recommendation algorithm will also be different.

In order to describe the relative position of a crop in a certain ranking sequence, two indicators of ranking advantage and ranking disadvantage are firstly proposed. These two indicators go up and down, so the F-score is introduced to synthesize them [8, 9]. The two indicators of ranking advantage and ranking disadvantage are represented by A and D respectively, and their calculation methods are as follows:

$$A(s) = A(s_i) = \frac{p_b - i + 1}{p_b} \quad (1)$$

$$D(s) = D(s_i) = \frac{1}{M} \quad (2)$$

Among them,  $s$  represents the priority order of Party A's main body. In order to comprehensively measure the ranking advantage and ranking disadvantage, this paper

introduces the F score. According to the definition of F-score, its expression is:

$$F_{\alpha}(s) = (1 + \alpha)^2 \cdot \frac{A \cdot D(s)}{(\alpha^2 \cdot A(s)) + D(s)} \quad (3)$$

In effect, the F-score can be thought of as a weighted harmonic mean of A and D.

## 2.2 Systematic Analysis of Supply Chain Competition

The supply chain takes customer demand as the internal driving force, and realizes the accumulation of its own profits through the production, transportation and sales of products while meeting customer needs. Under the guidance of changing customer needs, the supply chain continuously adjusts and completes the entire operation process, and the supply chain must be coordinated and closely follow the changes of customer needs in order to win [10, 11].

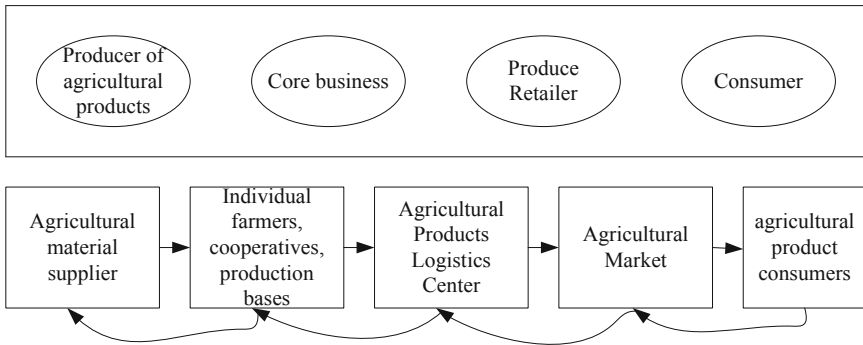
In addition to the physical elements in the supply chain, there are also elements such as information flow, logistics, and capital flow. First of all, an important factor for the success of the supply chain operation is whether the enterprises within the supply chain can achieve coordination and meet the changing customer needs, especially with the rapid development of information technology, the information flow has become an important factor related to the success or failure of the supply chain. Secondly, a typical manufacturing supply chain is also filled with a large number of supply and demand relationships, such as upstream companies meeting the ordering needs of downstream companies, downstream companies meeting consumer needs, etc. In the process of meeting supply and demand, a large amount of logistics is generally accompanied. Finally, the supply behavior between enterprises within the supply chain is not gratuitous, nor is it imperative (administrative), but a commercial transaction behavior. Therefore, a large amount of capital flow is bound to be generated in the process of these transactions [12].

## 2.3 Collaboration Between Crops and Supply Chain

Crops are various plants cultivated in agriculture. There are many types of crops with different characteristics. China has a large planting area, and the temperature, humidity, and climate in the north and south are different. Therefore, different regions are suitable for planting different crops.

Different types of vegetables have different planting time, receiving time, growth cycle, suitable temperature and storage time. However, it can be seen that vegetables can be grown every month. Therefore, farmers' supply of crops and farmers' demand for agricultural materials are perennial. Properly planning the production of crops can simultaneously reduce the total cost of the supply chain of agricultural materials and crops, improve the satisfaction of final consumers, and realize the integration of agriculture and materials.

The crop supply chain is a crop logistics chain that connects agricultural material suppliers, crop producers, crop logistics centers, crop retailers and crop consumers. Its structure is shown in Fig. 1:



**Fig. 1.** Agricultural Product Supply Chain with Logistics Center as the Core

Crop supply chain collaboration is the main body of the crop supply chain. Using modern technology, in the process of crops from the upstream to the downstream of the supply chain, the total cost is reduced by sharing information and mutual cooperation. Crop supply chain collaboration emphasizes the relationship between various subjects, including contract formulation, information sharing, risk sharing, and revenue sharing. As the raw materials of crops, the production cost, circulation cost and quality of agricultural materials are closely related to the crop supply chain.

The collaborative model of agricultural materials-crop supply chain, builds a collaborative model based on agricultural materials suppliers, agricultural cooperatives, crop demanders and third-party logistics companies.

The generation of supply chain collaboration is aimed at the conflict of production and sales plans that each subject in the supply chain aims to maximize their own interests. Negotiation is an effective means of resolving conflicts and realizing information sharing, and it is also the premise of making contracts. Therefore, negotiation is the key to realizing supply chain coordination. The traditional negotiation methods among the various entities in the supply chain are mostly offline negotiation, and the online negotiation technology is developed with the development of e-commerce. However, the traditional agent negotiation mode is still not suitable for large enterprises such as farmers, processing enterprises and supermarkets. The main reason is the inherent instability of the crop supply chain. Therefore, this paper first proposes a third-party logistics-centered agricultural material-crop supply chain collaboration model. And then constructs a model based on third-party logistics between third-party logistics enterprises and processing enterprises, supermarkets, between third-party logistics enterprises and agricultural material suppliers. A negotiation model between 3PLs and agricultural cooperatives is to support supply chain stability and contract enforcement.

## 2.4 Supply Chain Performance

The crop supply chain is very different from other supply chains because the crop supply chain is based on the supply chain of this particular product of crops. Characteristics of the crop supply chain include: high degree of specialization of facilities and equipment,

high operational risks, high requirements for logistics levels, large differences in time and space between crop supply and demand, and relatively fragile crop supply chains.

Using factor analysis method, combined with data, the performance level of crop operation was evaluated. The crop performance evaluation index system is constructed from four aspects: supply chain node performance level, supply cooperation level, supply chain operation level and supply chain green level, and the validity and practicability of the evaluation system are verified through the model.

### **3 Investigation on the Configuration of Crop Supply Chain**

#### **3.1 Purpose of the Investigation**

In order to promote the development of crops and crop supply chains, and promote the level of agricultural modernization, it is necessary to grasp the existing technology and related industries, and optimize the design of the configuration of the supply chain of crops. The sources of goods in this article mainly include seeds, fertilizers and pesticides of crops. In realizing the optimal allocation problem, in order to make the whole system operate effectively, it must be ensured that all nodes can be utilized to the maximum. Therefore, the existing crop supply chain is optimized and configured based on the Gale-Shapley algorithm.

#### **3.2 Investigation Method**

(1) On-the-spot investigation

This paper firstly investigates the rural soil and suitable seeds for planting in this city. Secondly, investigate the number and quality of local farmers. Then carry out relevant investigations on the farming tools and purchasing methods of the land. Finally, an unannounced visit was made to the township government's support for agricultural land-related policies.

(2) Questionnaire

After on-the-spot investigation, a questionnaire survey was conducted among the villagers. The questionnaire survey was mainly carried out around issues related to the supply chain of crops. The method of questionnaire survey is offline questionnaire, that is, villagers are invited to answer the questionnaire on the spot, and they will give guidance on the side.

#### **3.3 Investigation Process**

The survey took one week on the field trip, and the questionnaire survey design and development lasted three days. A set of questionnaires was designed, and a total of 100 questionnaires were distributed. Since the questionnaires were collected on site, after three days of filling in the questionnaires, 90 completed questionnaires were obtained, and the effective filling rate reached 100%.

### 4 Analysis of Questionnaire Results

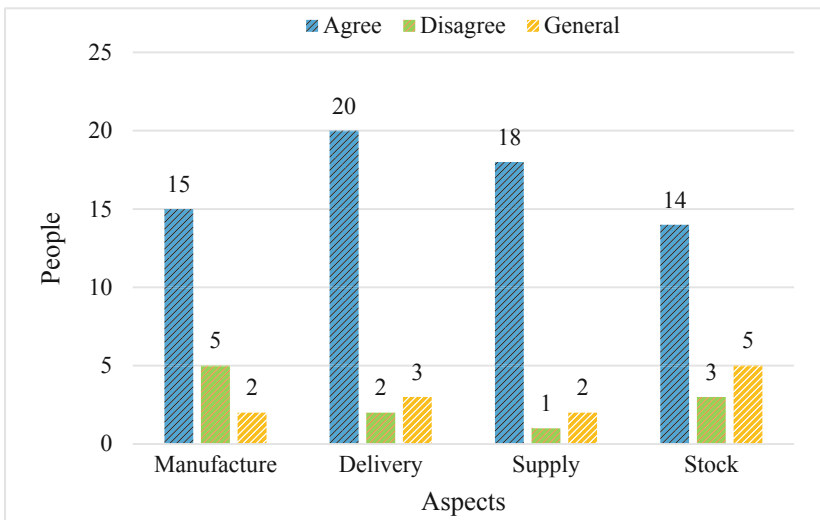
Farmers’ Attitudes towards the Allocation of Crop Supply Chains.

The results of this paper based on the questionnaire survey are shown in Table 1. The villagers believe that the manufacturing, distribution, supply chain design and inventory requirements must be taken into account in the configuration of the supply chain of crops.

**Table 1.** Farmers’ Attitude towards the Allocation of Crop Supply Chain

	Agree	Disagree	General
Manufacture	15	5	2
Delivery	20	2	3
Supply	18	1	2
Stock	14	3	5

As shown in Fig. 2, we can find that 15 villagers agree with the role of crop production in the supply chain. 20 villagers agreed with the role of crop distribution in the supply chain. Thirty-two villagers agreed with the supply of crops and the role of inventory in the supply chain.



**Fig. 2.** Farmers’ Attitude towards the Allocation of Crop Supply Chain



## 5 Conclusions

The crop supply chain is an organic whole, which includes multiple stakeholders such as crop producers, processing companies and retailers. In the whole system, farmers and processing enterprises as intermediate links can realize the forecast of product demand through information exchange. In order to improve circulation efficiency, it is also necessary to strengthen effective communication with upstream suppliers to obtain more resources and opportunity costs. The crop supply chain is a complex system. It consists of the logistics and information flow involved in the production, processing and sale of crops. Since the Gale-Shapley algorithm model has certain defects, it is necessary to further improve and perfect it.

## References

1. Zacharia, N., Papaioannou, E., Kaklamanis, C.: An efficient implementation of the Gale and Shapley “propose-and-reject” algorithm. *Electron. J. Graph Theory Appl.* **8**(1), 29–57 (2020)
2. Omar, A., et al.: Supply chain inventory sharing using ethereum blockchain and smart contracts. *IEEE Access* **10**, 2345–2356 (2022). <https://doi.org/10.1109/ACCESS.2021.3139829>
3. Rajesh, R.: A novel advanced grey incidence analysis for investigating the level of resilience in supply chains. *Ann. Oper. Res.* **308**(1–2), 441–490 (2020). <https://doi.org/10.1007/s10479-020-03641-5>
4. Lejarza, F., Baldea, M.: An efficient optimization framework for tracking multiple quality attributes in supply chains of perishable products. *Eur. J. Oper. Res.* **297**(3), 890–903 (2022)
5. Anzoom, R., Nagi, R., Vogiatzis, C.: A review of research in illicit supply-chain networks and new directions to thwart them. *IISE Trans.* **54**(2), 134–158 (2022)
6. Mondal, A., Roy, S.K.: Application of Choquet integral in interval type-2 Pythagorean fuzzy sustainable supply chain management under risk. *Int. J. Intell. Syst.* **37**(1), 217–263 (2022)
7. Kraus, M., Folinás, D.: The impact of a manufacturing execution system on supply chain performance. *Int. J. Appl. Logist.* **12**(1), 1–22 (2022)
8. Faasolo, M.B., Sumarliah, E.: An artificial neural network examination of the intention to implement blockchain in the supply chains of SMEs in Tonga. *Inf. Resour. Manag. J.* **35**(1), 1–27 (2022)
9. Lima, C., Relvas, S., Barbosa-Póvoa, A.: A graph modeling framework to design and plan the downstream oil supply chain. *Int. Trans. Oper. Res.* **29**(3), 1502–1519 (2022)
10. Castellano, D., Gallo, M., Santillo, L.C.: A periodic review policy for a coordinated single vendor-multiple buyers supply chain with controllable lead time and distribution-free approach. *4OR* **19**(3), 347–388 (2021)
11. Fatemi Ghomi, S.M.T., Karimi, B., Behnamian, J., Firoozbakht, J.: A multi-objective particle swarm optimization based on pareto archive for integrated production and distribution planning in a green supply chain. *Appl. Artif. Intell.* **35**(2), 133–153 (2021)
12. Ahmad, R.W., Salah, K., Jayaraman, R., Yaqoob, I., Omar, M., Ellahham, S.: Blockchain-based forward supply chain and waste management for COVID-19 medical equipment and supplies. *IEEE Access* **9**, 44905–44927 (2021). <https://doi.org/10.1109/ACCESS.2021.3066503>



# Application of Particle Swarm Optimization in BIM Building Modeling

Guang Yang<sup>1</sup>(✉) and Xin Guo<sup>2</sup>

<sup>1</sup> QianNan Polytechnic for Nationalities, Guizhou 558022, Duyun, China  
1007153730@qq.com

<sup>2</sup> Anhui Agricultural University, Hefei 230036, Anhui, China

**Abstract.** BIM technology is defined as a software tool for building models, which is now used by the architecture industry as an extremely important modeling tool for architectural drawings. This tool combines architectural theory and knowledge, as well as many software plug-ins for practical applications. This technology can complete typical 3D drawings, which is one of the basic prerequisites for building construction. BIM technology is a tool with great application value. This tool can carry out 3D stereoscopic of building model and visual design of building structure at the same time. The biggest advantage is that it can ensure the optimal accuracy of each data point of the building, so as to ensure that the building becomes a very accurate building body after the actual completion. This paper studies the application of particle swarm optimization in BIM building modeling and analyzes the application of BIM building modeling. The data test shows that the application of particle swarm optimization algorithm in BIM building modeling has excellent performance in building modeling accuracy.

**Keywords:** PARTICLE Swarm Optimization · BIM Building · Building Modeling · Application Research

## 1 Introduction

The construction project is one of the projects with large scale and complex working procedure, and the completion of its construction work must rely on information technology. The use of BIM technology solves most of the work problems of construction engineering, from architectural drawing design, architectural detail setting, architectural quality parameter control, building information construction, etc., can be solved very efficiently. In the actual construction work, the application of BIM technology effectively solves a lot of technical problems in construction engineering, and solves a number of work procedures in construction. BIM technology adopts the latest technical architecture of software, which well solves some modern application problems and uses information technology to solve construction engineering problems. The application research of particle swarm optimization algorithm in BIM building modeling promotes the deepening of the application research of BIM building modeling.

Many scholars at home and abroad have studied particle swarm optimization. In foreign studies, NabiS proposed a dynamic load balancer based on resource and deadline awareness based on particle swarm optimization, which is suitable for cloud tasks with deadline constraints. These schedulers can optimize a single target or multiple targets with non-conflicting parameters. However, schedulers need to be able to provide balanced solutions for conflicting parameters such as time and cost [1]. In this paper, we propose a novel algorithm for simultaneous adjustment of optimal parameters of automatic voltage regulator (AVR) and power system stabilizer (PSS) using oscillating exponential decay particle swarm optimization technique. The algorithm is applied to single - machine infinite bus (SMIB) power system and 9 - bus multi - machine power system. Simulation results show that the proposed method is effective and can ensure the dynamic stability and convergence speed of simulation [2]. AlajmiMS proposed to investigate the performance of advanced machine learning and quantum behavior evolution calculation methods for predicting aluminum surface roughness in face milling machine. Quantum particle swarm optimization (QPSO) and least square gradient Booster integration (LSBoost) were used to simulate a large number of surface milling experiments and predict surface roughness values with high precision [3].

At this stage, there are a huge number of BIM enterprises. At the beginning of the 21st century, with the acceleration of urbanization in China, such companies were rapidly generated, although many enterprises have different qualities and business consciousness [4, 5]. But these companies, like a hundred schools of thought, have promoted the development of the theory and application of the building modeling industry, and promoted the rapid progress of the building industry. Research on the application of particle swarm optimization algorithm in BIM building modeling is conducive to the technological breakthrough in the field of BIM building modeling [6, 7].

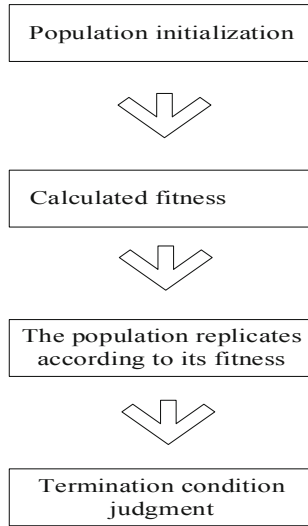
## 2 Design and Exploration of the Application Research of Particle Swarm Optimization in BIM Building Modeling

### 2.1 Particle Swarm Optimization

Particle swarm optimization algorithm (PSO) is a swarm intelligence optimization algorithm that simulates the social behavior of animal groups and has become an important branch of natural computing [8, 9].

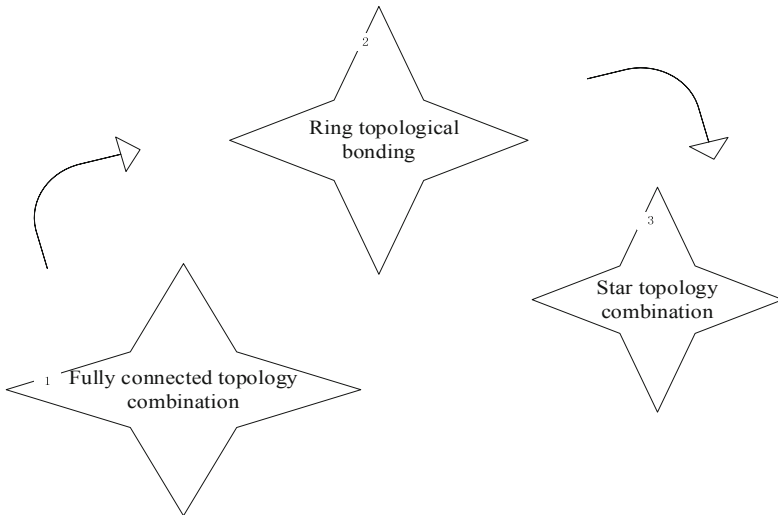
The similarities with genetic algorithm are shown in Fig. 1: First, sample initialization; Second, to obtain individual fitness in the population; Third, according to the above fitness, population replication; (4) Judge whether it can be stopped. If the expected result is achieved, stop. If not, skip to the second step [10, 11].

As can be seen from the above, THERE are many similarities between PSO and genetic algorithm. Both algorithms are initialized first, evaluated systematically by fitness, and searched in space according to fitness. Neither algorithm can guarantee the optimal solution. PSO does not have crossover, mutation, etc., in the steps of genetic algorithm, and the particle can have memory function.



**Fig. 1.** Genetic algorithms having in common

Considering the impact of population topology on algorithm performance, it is combined with LAPSO algorithm to further improve the optimization performance of the proposed LAPSO algorithm [12, 13]. Because different topologies have different characteristics, they have different effects on algorithm performance. Therefore, the fully connected, ring and star topologies shown are combined with LAPSO algorithm respectively, and their advantages and disadvantages are analyzed, as shown in Fig. 2.



**Fig. 2.** Three kinds of population topology

- (1) Fully connected topology is combined with LAPSO algorithm. In the early stage of search, because each particle is not only affected by the gravity of particles with better fitness than its own, but also repelled by particles with worse fitness than its own, the speed of approaching the current optimal particle is slow, and the algorithm shows good population diversity. In the later stage of search, this interaction makes the aggregation of particles become worse to the globally optimal particles, so the weakness of local search capability is weak.
- (2) Combination of ring topology and LAPSO algorithm [14, 15]. Since the contemporary particle of the ring topology is only related to its two adjacent particles, in the early search of the algorithm, the contemporary particle is only affected by the interaction of the two adjacent particles, which reduces the influence of the global optimal particle to a certain extent, but its population diversity is not as good as that of the fully connected topology. In the later stage of search, the particle interaction is obviously smaller than that of the fully connected topology, so it has a faster convergence rate and better local search ability.
- (3) Star topology is combined with LAPSO algorithm. In the star topology, the central particle moves in the direction of the resultant force of the interaction of all other particles, while the other particles are only affected by the central particle and quickly move towards it. Although the convergence speed is fast, it is easy to fall into local optimization.

## **2.2 Research on the Application of Particle Swarm Optimization in BIM Building Modeling**

The application research of particle swarm optimization in BIM building modeling can be divided into the following points:

First of all, BIM technology adopts software technology to mainly solve the problem of building modeling. By establishing a building model, data simulation is carried out on the construction project, so that engineers can improve the aesthetic, safety, information and other modern building requirements of the construction project according to the data set [16, 17]. Among them, the software platform is a very key part, the software involves a lot of technical realization of architectural modeling, such as the realization of architectural modeling interface, the realization of the underlying computing module of architectural modeling, data analysis, data processing, creative design, the adoption of data model through computer tools. Among them, the technology includes computer hardware part and software part, which can comprehensively process the data of building modeling and promote the optimization of building model.

Secondly, in the large-scale application of BIM building model, software plug-in is a very important aspect that cannot be ignored. Software processing is the building model data collection, data processing, so as to realize the building model construction. BIM construction model deals with all aspects of construction engineering, data simulation, data collection, data processing, etc., which are very important aspects of building model establishment. Information technology has accelerated the process of building model datalization.

Finally, BIM technology promotes BIM software to be more scientific and efficient from the input of model data and adjustment of parameters [18]. Architectural model

should pay attention to the establishment and modification of data model, so as to produce a very realistic architectural model. In the actual data model construction, BIM technology is very key, this technology covers a lot of building data modules, especially building data information, so that it can be directly introduced to build the building model. In the process of building a building model, the computer makes comprehensive use of many algorithm formulas, such as splitting algorithm, dynamic R-tree generation algorithm, Bayesian classification algorithm and so on.

### 3 Research on the Application Effect of Particle Swarm Optimization in BIM Building Modeling

When entering the region of global optimal solution, enhance the gravitational effect and reduce the repulsive effect, and consider the gravitational effect of particles with better fitness than themselves and particles of global optimal solution to improve the local search ability:

The speed and position update formula of LAPSO algorithm is

$$v_{ik}(t+1) = w_{v_{ik}}(t) + \sum_{j \in B(i)} C_j r_{jk} [p_{jk}(t) - x_{ik}(t)] - (1 - \varepsilon) \sum_{z \in W(i)} C_z r_{zk} [p_{zk}(t) - x_{ik}(t)] + \varepsilon r_{gk} [p_{gk}(t) - x_{ik}(t)] \quad (1)$$

$$X_{ik}(t+1) = X_{ik}(t) + V_{ik}(t+1) \quad (2)$$

Where,  $v_{ik}(t+1)$  and  $v_{ik}(t)$ ,  $x_{ik}(t+1)$  and  $x_{ik}(t)$  are respectively the velocity and position of the KTH dimension of the  $i$ th particle of the  $t+1$  generation and the KTH dimension of the  $t$  generation;  $W$  is inertial weight;  $c_j$  and  $c_z$  are learning factors;  $r_{jk}$ ,  $r_{zk}$  and  $r_{gk}$  are random factors ranging from 0 to 1.  $p_{jk}(t)$  and  $p_{zk}(t)$  represent the positions of the  $k$ -dimension of the  $t$ -generation particle  $J$  and particle  $Z$  respectively;  $p_{gk}(t)$  represents the position of the  $k$ -dimension of the global optimal particle  $G$  of the  $t$ -generation;  $B(I)$  represents the collection of particles with better fitness than particle  $I$ ;  $W(I)$  represents the collection of particles with lower fitness than particle  $I$ ;  $\varepsilon$  is the intervention factor, which is used to intervene the influence factor of the global optimal solution.

$$\varepsilon = \begin{cases} 0 & t < t_g^1 \\ 1 & t < t_g^1 \end{cases} \quad (3)$$

Where,  $t_g^1$  is the minimum iteration number of LAPSO algorithm when it enters the global optimal solution region.

In this paper, particle swarm optimization algorithm is used to build BIM building modeling. The algorithm is used to input sample data to the building that needs modeling. Since the building is a huge overall system, the system is composed of countless key data groups. However, during the modeling process, these data sets need to be set at speed and position. The velocity and position are calculated by formulas (1) and (2). At the same time, intervention factors are introduced to adjust the overall parameters of BIM building modeling, so as to achieve the BIM building model that meets the actual requirements.

### 3.1 The Advantages of BIM

- (1) 3d modeling visualization can more intuitively show the appearance and characteristics of the building, virtual technology allows people to enjoy the beauty, investors can more clearly know the basic situation of the construction project.
- (2) 3D modeling has technical innovation, with CAD drawing space can not meet the solution ability. It can integrate the advantages of multiple disciplines to solve very detailed and complicated small problems and reduce the completion time and cost of the project.
- (3) The information technology adopted at the bottom of BIM technology introduces computer tools, which can deal with data model and geometric and mathematical problems more efficiently, and at the same time promote the optimization of construction engineering, improve efficiency and reduce error.

## 4 Investigation and Research Analysis of the Application of Particle Swarm Optimization in BIM Building Modeling

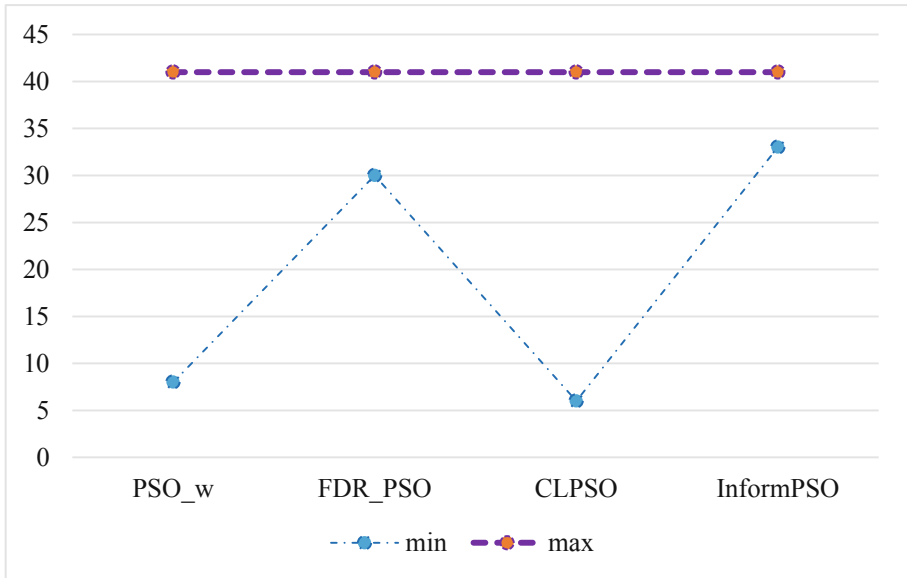
In this section, PSO\_w, FDR\_PSO, CLPSO and InformPSO algorithms are implemented. InformPSO is the particle swarm optimization algorithm adopted in this paper. The application of BIM in building modeling is tested, and the same parameters are set for the four algorithms. This paper realizes the platform is Matlab.

**Table 1.** Table of application algorithm test effect in BIM building modeling

Function	Best/worst fitness value			
	PSO_w	FDR_PSO	CLPSO	InformPSO
F1	7.6843e-12 2.1076e-8	3.7173e-33 6.7132e-32	6.5218e-15 1.0543e-14	9.7123e-41 8.7345e-41
F2	9.6843e-41 9.1076e-38	4.7043e-41 9.2135e-41	4.0432e-35 1.6783e-32	7.9432e-41 7.6872e-41
F3	8.6103e-41 8.9086e-41	2.7657e-41 5.0657e-41	1.2076e-41 3.4012e-41	7.7032e-41 8.3425e-41
F4	9.3643e-20 4.7096e-17	5.7064e-33 4.0643e-30	9.6582e-15 7.1534e-14	1.3298e-35 1.9802e-33
F5	5.1226e-20 1.5506e-16	9.7053e-41 9.7547e-41	7.9820e-11 4.9125e-6	9.6643e-41 9.8398e-41

As shown in Table 1, four algorithms, PSO\_w, FDR\_PSO, CLPSO and InformPSO, are listed in the table. F5 carries out function optimization, and the accuracy achieved is shown in the table. For example, the optimization result of PSO\_w on function F1 is that the optimization accuracy range is (2.1076E-8, 7.6843E-12). It can be seen from the table that the optimal accuracy  $10^{-41}$  and  $10^{-8}$  minimum accuracy of PSO\_w optimization for the five functions are achieved. The minimum precision of FDR\_PSO,

CLPSO and InformPSO is  $10^{-30}$ ,  $10^{-6}$ ,  $10^{-33}$  respectively; The optimal precision is  $10^{-41}$ ,  $10^{-41}$ ,  $10^{-41}$  respectively. InformPSO has the strongest optimization capability and can better optimize BIM building modeling. At the same time, the optimization accuracy of all four algorithms can be kept above average  $10^{-6}$ , so all algorithms have high precision processing ability.



**Fig. 3.** Minimum and maximum accuracy images

Figure 3 shows that the purple dotted line represents the maximum optimization accuracy of BIM building modeling optimized by four algorithms. The blue dotted line represents the minimum optimization accuracy of BIM building modeling optimized by four functions. As can be seen from the figure, InformPSO algorithm has the highest optimization ability. At the same time, the average optimization accuracy of the four algorithms exceeds, and the overall optimization ability is very strong.

The test shows that the application of particle swarm optimization algorithm in BIM building modeling has a high application effect in BIM building modeling.

## 5 Conclusions

The promotion of BIM technology is actually people's pursuit of new technology, and people choose BIM technology in the information age more wisely. This technology has many advantages, such as building structure drawing, building knowledge storage and so on. The addition of new computer tools further enhances the progress of BIM modeling ability. The actual construction project, there will generally be a lot of difficulties and deficiencies in the field of architecture, and then people can use computer technology



to solve this problem. BIM solves the problem of poor communication of buildings and makes use of computer software to smooth the communication of information. The application research of particle swarm optimization algorithm in BIM building modeling is conducive to higher level research and development and innovation of BIM building modeling application research.

## References

1. Nabi, S., Ahmed, M.: PSO-RDAL: particle swarm optimization-based resource - and deadline-aware dynamic load balancer for deadline constrained cloud tasks. *J. Supercomput.* **78**(4), 4624–4654 (2021). <https://doi.org/10.1007/s11227-021-04062-2>
2. Rodrigues, F., Molina, Y., Silva, C., et al.: Simultaneous tuning of the AVR and PSS parameters using particle swarm optimization with oscillating exponential decay. *Int. J. Elec. Power Energy Syst.* **133**(4), 107–215 (2021)
3. Alajmi, M.S., Almeshal, A.M.: Least squares boosting ensemble and quantum-behaved particle swarm optimization for predicting the surface roughness in face milling process of aluminum material. *Appl. Sci.* **11**(5), 2126 (2021)
4. Qt Eat, H., Awad, M.: Using hybrid model of particle swarm optimization and multi-layer perceptron neural networks for classification of diabetes. *Int. J. Intell. Eng. Syst.* **14**(3), 11–22 (2021)
5. Alshahir, A., Molyet, R.: Improving the reconfiguration of hybrid power networks by combining Genetic Algorithm (GA) with Particle Swarm Optimization (PSO). *Amer. J. Elec. Power Energy Syst.* **10**(1), 6 (2021)
6. Rezgui, S.E., Benalla, H., Bouhebel, H.: Hybrid bacteria foraging-particle swarm optimization algorithm in DTC performance improving for induction motor drive. *Indonesian J. Elec. Eng. Comput. Sci.* **22**(2), 660 (2021)
7. Kotla, R.W., Yarlagadda, S.R.: Comparative analysis of photovoltaic generating systems using particle swarm optimization and cuckoo search algorithms under partial shading conditions. *J. Europ. J. Res. des Syst. 'mes Automatis'és* **54**(1), 27–33 (2021)
8. Nabavi, S.R., Eraghi, N.O., Torkestani, J.A.: Wireless sensor networks routing using clustering based on multi-objective particle swarm optimization algorithm. *J. Intell. Proced. Elec. Technol. (JIPET)* **12**(47), 49–67 (2021)
9. Wijayanti, E.A., Rahmadanti, T., Enri, U.: Perbandingan Algoritma SVM dan SVM Berbasis Particle Swarm Optimization Pada Klasifikasi Beras Mekongga. *Gener. J.* **5**(2), 102–108 (2021)
10. Malik, G., Upadhyaya, S., Sharma, R.: Particle swarm optimization and maximum entropy results for MX/G/1 retrial G-Queue with delayed repair. *Int. J. Math. Eng. Manag. Sci.* **6**(2), 541–563 (2021)
11. Ramadhani, B., Garside, A.K.: Particle swarm optimization algorithm to solve vehicle routing problem with fuel consumption minimization. *Jurnal Optimasi Sistem Industri* **20**(1), 1–1 (2021)
12. Saeed, A.A., Jameel, N.: Intelligent feature selection using particle swarm optimization algorithm with a decision tree for DDoS attack detection. *Int. J. Adv. Intell. Inform.* **7**(1), 37–48 (2021)
13. Prasetyo, T.A.: Particle swarm optimization and genetic algorithm for big vehicle problem: case study in national pure milk company. *Int. J. Comput. Sci. Appl. Math.* **7**(1), 28 (2021)
14. Vijayakumar, T., Vinothkanna, R.: Efficient energy load distribution model using modified particle swarm optimization algorithm. *J. Artif. Intell. Capsule Netw.* **2**(4), 226–231 (2021)

15. Setiami, R., Maulana, A.: Development of E-modules in engineering drawing courses with the BIM system building modeling application. *Jurnal PenSil* **10**(1), 1–7 (2021)
16. Sheward, H.: BIM Based Analysis of Spatial Properties in Building Layouts. *American Journal of Civil Engineering and Architecture* **9**(4), 142–155 (2021)
17. Sriyolja, Z., Harwin, N., Yahya, K.: Barriers to implement Building Information Modeling (BIM) in construction industry: a critical review. *IOP Conf. Ser. Earth Environ. Sci.* **738**(1), 012021012021 (2021)
18. Omayer, H.M.: Building Information Modeling BIM as a development tool for the management of construction projects. *Fayoum Univ. J. Eng.* **3**(2537–0626), 9 (2021)



# Intelligent Prediction Model of Agricultural Environment Based on Intelligent Algorithm

Shuxin Zhang and Fangbo Hou<sup>(✉)</sup>

Electrical and Information Engineering College, Jilin Agricultural Science and Technology University, Jilin 132101, China  
houfangbo2016@163.com

**Abstract.** Intelligent computing technology is to describe the problem object through a specific mathematical model, making it operable, programmable, computable and visual. Intelligent computing technology uses its characteristics of parallelism, self-adaptability and self-learning habit to conduct law mining and knowledge discovery for massive data in various disciplines. The intelligent algorithms accumulated in the application of intelligent computing technology have the common characteristics of simulating natural processes. In this paper, the intelligent algorithm is applied to the agricultural environment prediction model, and its application in the agricultural environment prediction model is studied.

**Keywords:** Intelligent Computing Technology · Intelligence Algorithm · Agricultural Environment Prediction Model

## 1 Establishment and Determination of Agricultural Mathematical Models

Agricultural mathematical models usually construct a certain specific process in agricultural problems, and use a few mathematical formulas for illustration [1]. Generally speaking, agricultural mathematical models have simple forms and simple connotations [2]. Agricultural mathematical models are usually established to rationally analyze a certain mechanism or simulate a specific objective process. There are multiple ways to establish and determine agricultural mathematical models.

- (1) Mathematical models can be established based on a large amount of experimental data obtained from serious field experiments or breeding experiments. Since agricultural mathematical models mainly serve agriculture, and agriculture is a process conducted in fields or breeding grounds, it is essential to establish basic agricultural mathematical models (such as tillering dynamic model, leaf area model and crop growth period model, etc.) in fields or breeding grounds. It should be noted that long-term farmland experiments, breeding experiments, or feeding experiments are required before modeling. To make the model universal, farmland experiments or feeding experiments must be diverse. If conditions permit, experiments should be

conducted in a multi-point range. Before farmland experiments or feeding experiments, it is crucial to have a theoretical understanding of mathematical models. The theoretical understanding is not arbitrary imagination, but thinking based on the simulation process for thinking. Certainly, excellent research and development results completed by predecessors must be absorbed [3]. After integrating with own actual situations, agricultural mathematical models can be completed through innovative research. Based on this concept, preliminary mathematical models should be established first. After obtaining accurate data from actual farmland experiments or breeding experiments, tentative mathematical models can be used to perform rigorous statistical analysis of the model results. If the model results are not good, tentative mathematical models must be improved or even rejected. It is necessary to reconsider experimental mathematical models. This process should be repeated until a good agricultural mathematical model is obtained [4].

- (2) For situations that cannot be studied on farmland fields or breeding grounds, such as photosynthesis and respiration of crops, necessary and possible laboratory experiments must be carried out. To establish agricultural mathematical models, it is vital to control the lighting, temperature, and humidity conditions of the laboratory. In addition, the model should reflect the law of change in a process when only 1 or 2 factors change while other factors remain unchanged. If all factors are changing, it is difficult to establish single-factor or two-factor mathematical models. Normally speaking, mathematical models established in the laboratory are more reasonable than those established in farmlands (or breeding grounds).
- (3) The scientific literature related to the model should be extensively consulted. The mathematical models generally recognized should be selected. Considering the complexity and diversity of the agricultural production process, agricultural models may contain dozens or even hundreds of mathematical models. The agricultural mathematical models used are not all models that can be developed by modelers. Generally speaking, mathematical models recognized by the scientific community are reliable and applicable after rigorous scientific verification, which can be applied [5].

## 2 Application Examples of Agricultural Mathematical Models

### 2.1 Regression Prediction Model

The regression model is usually used to analyze the correspondence between one or multiple variables and dependent variables. Finally, the relationship function is determined through statistical testing and then applied to the calculation of predictive analysis. The formula of the linear regression model is:

$$Y = \sum_{i=1}^n a_i x_i + b \quad (1)$$

Where  $Y$  is the dependent variable;  $x_i$  is the  $i$ -th independent variable;  $a_i$  is the regression coefficient;  $b$  is the constant (namely, the regression intercept). When  $n = 1$ , it is called a single variable linear regression model.

In the field of agricultural environment, the application of a regression model can be described as diverse, including predicting agricultural water use in a region, studying the relationship between energy consumption and agricultural output value, as well as the impact of environmental factors on agricultural output [6]. For example, the linear regression method can be used to establish a quantitative model of agricultural nitrogen and phosphorus loss rate [7].

The establishment of the model can be divided into the following steps:

- (1) Determination of regression factors: after extensive literature reading, fertilization, rainfall, vegetation coverage, slope, and soil geology are used as the main factors affecting the soil nitrogen and phosphorus loss rate.
- (2) Collection of data: 63 groups of data on the loss of nitrogen and phosphorus from agricultural land by use types in 32 academic papers published in 1993 and 2007 are collected [8].
- (3) Modeling: the statistical analysis software SPSS is applied to analyze the data, including linear regression analysis, significance test, correlation test, etc.

Through multiple linear regression analysis, it is found that there is a good linear relationship between the agricultural nitrogen and phosphorus loss rate and its influencing factors. The regression equation is:

$$Y = -12.507 + 0.063x_1 + 0.143x_2 - 1.257x_3 + 4.796x_4 - 11.060x_5 \quad (2)$$

Where  $x_1$  is the precipitation;  $x_2$  is the total nitrogen input;  $x_3$  is the total phosphorus input;  $x_4$  is the slope  $x_5$  is the ratio of the total nitrogen input to the total phosphorus input;  $Y$  is the ratio of the total nitrogen loss to the total phosphorus loss.

## 2.2 Calculus Model

The calculus mathematical model can be used to describe the process of rice and wheat tillering. The process of rice and wheat tillering can directly affect the number of panicles and the yield of crops in the later period. The research on the number of rice and wheat tillering is of guiding significance for the improvement of the crop yield per hectare. The results show that the tillering number of rice and wheat in the early stage changes following the parabolic law [9]:

$$x = x_0 + k_1(1 - a_1x_0)t - k_2(1 - a_2x_0)t^2 \quad (3)$$

Where  $x$  is the number of tillers,  $x_0$  is the number of basic seedlings,  $t$  is the time, and the rest are parameters. The condition for the maximum tillering should meet the following requirement

$$\frac{dx}{dt} = k_1(1 - a_1x_0) - 2k_2(1 - a_2x_0)t = 0 \quad (4)$$

After solving the equation, the maximum value of the peak period of tillering can be obtained

$$t_{\max} = \frac{k_1(1 - a_1x_0)}{2k_2(1 - a_2x_0)} \quad (5)$$

The change rate of the peak tillering maximum value to the original seedlings is

$$\frac{dt_{\max}}{dx_0} = \frac{d}{dx_0} \left[ \frac{k_1(1 - a_1x_0)}{2k_2(1 - a_2x_0)} \right] = \frac{k_1(a_2 - a_1)}{2k_2(1 - a_2x_0)^2} \quad (6)$$

If  $\frac{dt_{\max}}{dx_0} < 0$ , the function decreases monotonously. In other words, the more basic seedlings, the earlier the peak of tillering will be reached, and the less tillering will increase. The rate of change of the tillering speed to the basic seedlings is:

$$\frac{\partial^2 x}{\partial x_0 \partial t} = \frac{\partial}{\partial x_0} [k_1(1 - a_1x_0) - 2k_2(1 - a_2x_0)t] = -k_1a_1 + 2k_2a_2t \quad (7)$$

If  $\frac{\partial^2 x}{\partial x_0 \partial t} < 0$ , the more basic seedlings, the lower the growth rate of tillering.

During the crop growth and development, in addition to the tillering of main stems, new tillering can also produce another tillering. Therefore, the crop tillering ability can be expressed as:

$$\frac{dx}{dt} = b(x_0 + b't) \quad (8)$$

Where  $b$  is the parameters related to the tillering of main stems, and  $b'$  is the parameter related to new tillering. Then, the total number of tillering can be obtained by integrating, namely:

$$\int_0^t dx = \int_0^t b(x_0 + b't)dt \quad (9)$$

$$x = bx_0t + \frac{1}{2}bb't^2 = bt(x_0 + \frac{1}{2}bb')$$

This is the mathematical model used to calculate the total number of tillering.

### 2.3 Power-Exponent Model

The power-exponent model can be used to reflect the relationship between pest density and yield loss rate [10].

It is unreasonable to use a linear function  $L = a + bX$  or a power function  $L = bX^a$  to express the quantitative relationship between the pest density  $X$  and the crop yield loss rate  $L$ . After improving the model, the power-exponent model is introduced

$$Y = Y_M \exp(-bX^a) \quad (11)$$

$Y$  refers to the crop yield after pests;  $Y_M$  refers to the maximum crop yield without pests;  $a$  and  $b$  are parameters.

The yield loss rate  $L$  is:

$$L = \frac{Y_M \exp(-bX^a)}{Y_M} \quad (12)$$

Or

$$L = 100 - 100 \exp(-bX^a) \quad (13)$$

After analyzing the model, the following can be obtained:

- (1) If the pest density is 0, the yield loss rate is also 0.
- (2) When the pest density is low, that is, there are no intraspecific competitions of pests, the yield loss rate increases rapidly with the increasing pest density.
- (3) When the pest density is high, that is, there are intraspecific competitions of pests, the yield loss rate slows down with the increasing pest density.
- (4) When the pest density is  $\infty$ , the yield loss rate is  $L = 100\%$

It has been verified that the relationship between pest density and yield loss rate reflected in the power-exponent model is consistent with the actual situation. In other words, the power-exponent model has certain characteristics of the pest simulation model.

## References

1. Drewniak, B., Song, J., Prell, J., et al.: Modeling agriculture in the community land model. *Geosci. Model Dev.* **6**(2), 495–515 (2013)
2. Hao, F., Xuan, L., Mu, C.: Research on key technologies of intelligent agriculture based on agricultural big data. In: International Conference on Smart City and Systems Engineering, pp. 598–601. IEEE (2017)
3. Soldatova, E., Guseva, N., Bychinsky, V.: Modelling of redox conditions in the shallow groundwater: a case study of agricultural areas in the Poyang Lake Basin China. *Proc. Earth Planet. Sci.* **17**, 197–200 (2017)
4. Bai, X.Z., et al.: Coal mine personnel positioning algorithm based on improved adaptive unscented Kalman filter with wireless channel fading and unknown noise statistics. *Trans. Inst. Meas. Control.* **44**(6), 1217–1227 (2022)
5. Kaburlasos, V.G., Spais, V., Petridis, V., et al.: Intelligent clustering techniques for prediction of sugar production. *Math. Comput. Simul.* **60**(3–5), 159–168 (2002)
6. Prabakaran, G., Vaithyanathan, D., Ganesan, M.: FPGA based effective agriculture productivity prediction system using fuzzy support vector machine. *Math. Comput. Simul.* 185:1–16 (2021)
7. Hashimoto, Y., Murase, H., Morimoto, T., Torii, T.: Intelligent systems for agriculture in Japan. *IEEE Control Syst. Mag.* **21**(5), 71 (2001)
8. Liu, L., Wang, Q., Li, B.Q.: A system architecture for intelligent agriculture based on edge computing. *Int. J. Comput. Appl. Technol.* **64**(2), 126–132 (2020)
9. Kumar, A.V.S.P., Bhramaramba, R.: Enhanced and improved hybrid model to prediction of user awareness in agriculture sector. *Int. J. Adv. Comput. Sci. Appl.* **9**(8), 338–343 (2018)
10. Solow, A.R., Adams, R.F., Weiher, R.: The value of improved ENSO prediction to US agriculture. *Clim. Change* **39**(1), 47–60 (1998)



# Mathematical Modeling and Optimal Design Based on Information Technology Algorithms

Hua Li<sup>(✉)</sup>

Beihua University Teacher's College, Jilin 132013, Jilin, China  
55485563@bh.edu.cn

**Abstract.** Higher mathematics should reform the teaching system, emphasize more practicality, and more importantly use multimedia technology, computer technology and network resources, so as to improve the efficiency and level of mathematics teaching. The purpose of this paper is to study mathematical modeling and optimal design based on information technology algorithms. Based on school teaching practice and guided by the relevant theories of teaching information technology, this paper conducts an investigation and research on the current situation of the application of information technology in mathematics teaching in schools, grasps the difficult problems and influencing factors that affect the application of teaching information technology, and develops the mathematics curriculum hierarchically. Carry out the application and practice of teaching information technology methods, and form a relatively standardized teaching process. Flexible use of various teaching software to build a teaching environment conducive to teacher-student interaction, build a better learning platform for students, and guide students to effectively use various information technologies to carry out learning. Guided by modern teaching information technology and based on its own teaching practice, this paper further clarifies the teaching theory of educational informatization through literature research, and rationally applies research methods. After two years of practice, it explores effective ways to reform teaching methods. And made an effect analysis, the research conclusions formed are reliable, and it has certain guiding and reference significance for the teaching of mathematics courses in colleges and universities. Experiments have proved that the teaching effect of mathematical modeling based on information technology is excellent. After teaching, 92% of students are interested in mathematical modeling, and the pass rate in the final assessment has reached 100%, of which the excellent rate has reached 94%.

**Keywords:** Information Technology · Mathematical Modeling · Advanced Mathematics · Mathematics Teaching

## 1 Introduction

Under the general trend of informatization of economic and social development in today's world, modern information technology has become inseparable from people's life, work



and study. A new request was made. The reform of mathematics teaching should highlight the applicability of teaching. Mathematical modeling is the breakthrough between mathematical theory and practical problems, and the best combination of mathematical knowledge and application ability. Through mathematical modeling, students can learn to use mathematical thinking to understand and think about the environment and society they live in, which can achieve the purpose of improving students' mathematical application ability and comprehensive quality [1, 2].

The research on mathematical modeling and optimal design based on information technology algorithms, many scholars have studied it and achieved good results, for example: Okubanjo A launched the TED program in the field of education, which requires teachers and students to work together to Explore TED and how this video should be used in classroom teaching [3]. Zhao P pointed out that the majority of teachers should fully understand the role of information technology in education and teaching, and understand how to apply it to the reform of education and teaching in the modern information technology environment. Teachers should learn the teaching design in the process of using information technology., so that in the information age, education and teaching can realize its own information function [4].

Based on school teaching practice and guided by the relevant theories of teaching information technology, this paper conducts an investigation and research on the current situation of the application of information technology in mathematics teaching in schools, grasps the difficult problems and influencing factors that affect the application of teaching information technology, and develops the mathematics curriculum hierarchically. Carry out the application and practice of teaching information technology methods, and form a relatively standardized teaching process. Flexible use of various teaching software to build a teaching environment conducive to teacher-student interaction, build a better learning platform for students, and guide students to effectively use various information technologies to carry out learning. Guided by modern teaching information technology and based on its own teaching practice, this paper further clarifies the teaching theory of educational informatization through literature research, and rationally applies research methods. After two years of practice, it explores effective ways to reform teaching methods. And made an effect analysis, the research conclusions formed are reliable, and it has certain guiding and reference significance for the teaching of mathematics courses in colleges and universities.

## **2 Research on Mathematical Modeling and Optimal Design Based on Information Technology Algorithms**

### **2.1 Characteristics of Information-Based Teaching Mode**

The basic characteristics of the information-based teaching model can be summarized into six aspects: First, it is necessary to emphasize that the center of the teaching process is the students, and to provide more favorable conditions for students' active learning. Second, in the teaching process, we should set up a more realistic learning environment, and learn to introduce students into a good teaching situation. Third, to provide students with more diversified teaching resources, so that students can more autonomously

research various problems encountered in the process of students' learning. Fourth, focus on guiding students to carry out interactive and cooperative learning. Fifth, it is emphasized that students learn to express, and actively use language to express the problems they encounter and the ideas they want to express. Sixth, guide students to carry out active meaning construction [5, 6].

## 2.2 The Innovative Educational Path of Mathematical Modeling Under the Modern Information Technology Environment

(1) Provide guidance in the process of mathematical modeling.

Mathematical modeling teaching is based on the purpose of improving students' interest in learning and cultivating innovative ability. Teach a specific academic program.

First, make a good theme selection. Mathematical models allow students to fully understand and experience the whole process of solving practical problems and feel the practical value of mathematics. Therefore, the topic selection must be effective, emphasizing the discovery of solutions to problems from different perspectives and levels, so as to benefit students at different levels.

The second is to focus on the process. By choosing the right subject, the depth and orientation process should be provided later as support [7, 8].

(2) Set up online courses on mathematical modeling.

The development of information technology has provided more flexible options for mathematical modeling, and it also makes students feel that learning mathematics is not boring, which is very good for cultivating students' creative thinking books. The practice of teaching mathematical modeling tasks online is a very important representation. The introduction of mathematics online learning makes students the main body of learning, not just simple "memorization", but more independent exploration or collaborative communication with peers. The progressive and relaxed learning environment encourages many students to enjoy learning mathematics. In particular, the standard modules such as basic models, functional learning models, and probability and statistics models presented in it make the simplicity and perspective of mathematical models more obvious, allowing students to have the opportunity to understand what they have learned, learn and acquire scientific knowledge and improve their self-learning ability [9, 10].

## 2.3 Algorithm Selection

In this paper, the survey method for mathematical modeling teaching mainly adopts questionnaire survey, and the algorithm for this survey questionnaire mainly uses Bayesian algorithm.

Definition 3.1: Let  $(C, E, P)$  be a complete event space, where  $P$  is the probability set, and the event  $A_i \in E (i = 1, 2, \dots, n)$ ,  $p(A_i) > 0$ , type set  $C = \{C_1, C_2, \dots, C_m\}$ , has a formula for any event  $B$  in the classification set:  $U_{i=1}^n A_i = C$ ,

$$P(B) = P(A_i) \sum_{i=1}^n P(B|A_i) \quad (1)$$

Definition 3.2: Suppose there is a complete set of events  $(C, E, P)$ , where and the event probability.

$P(A_i) + P(B) = 1$ ,  $P(B) > 0$ , For  $B \in E$  having the formula:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^n P(B|A_j)P(A_j)} \quad (2)$$

Since in reality, many actual environments and temporary situations are difficult to fully control and predict, Bayes theory cannot fully achieve objectivity and absoluteness in event statistics and expectation analysis, resulting in some computational differences [11, 12].

### 3 Mathematical Modeling and Optimization Design Based on Information Technology Algorithms Research and Design Experiments

#### 3.1 Survey Design

The application scope of information technology has become more and more extensive, and in many vocational colleges, information technology has been introduced to carry out information-based education. Although the theory of information-based education is very clear, there are still some problems in the process of teaching practice, especially many practical problems at the operational level. In order to have a clearer understanding of various issues, this paper conducts investigation and analysis, and interviews some teachers who work in the front line, and obtains a lot of first-hand information.

The purpose of this survey is mainly to analyze the problems faced in the current information-based education process, and to clarify the problem. In order to make the investigation activities more pertinent and obtain more valuable results, we first select the subjects of the interview, and then design the questionnaire after the interview. There are two types of questionnaires, one is for students, the other is for teachers. The questionnaire for students is mainly to understand the current students' computer and mathematical modeling level, and the questionnaire for teachers is mainly to understand the current teachers' information technology literacy, teaching literacy, professional ethics literacy and so on.

In order to investigate the current situation of informatization application of mathematical modeling teaching in schools, the survey objects selected in this paper are teachers and students of a college in Xi'an. There are 50 teachers in different grades and different professional shifts. Students were divided into experimental class and control class with a total of 200 students (134 boys and 66 girls).

#### 3.2 Questionnaire Survey

In this study, the survey method of intentional sampling was adopted. According to the author's own years of experience in teaching mathematics in secondary vocational schools, combined with the setting of the mathematics curriculum in our school, we

conducted separate interviews with teachers and students. A total of 200 students of different majors from a college in Xi'an were selected to conduct a questionnaire survey on mathematics learning, in order to comprehensively understand the mathematics learning situation of students in secondary vocational colleges. And 50 teachers were randomly selected to conduct a questionnaire survey on teaching information technology in order to clarify the teachers' teaching information technology literacy. The questionnaire is divided into two parts: student questionnaire and teacher questionnaire. (1) Student Vol. It includes the following 6 contents: A student's gender ratio; B student's situation when learning mathematical modeling; C student's use of information technology in the learning process; D student's computer ability level; E mathematics classroom learning The situation of F students' teaching methods and evaluation of teachers. (2) Faculty Vol. It includes the following 5 contents: A teacher's understanding of information technology teaching; B the integration effect of information technology in teaching practice; C teacher's teaching methods; D teacher's understanding of students' learning attitude; E teacher's information technology literacy understand the situation. 200 questionnaires were distributed for students, 200 were recovered, and the recovery rate was 100%; 50 questionnaires were distributed for teachers, and 50 were recovered, with a recovery rate of 100%. The returned questionnaires are all valid. In order to analyze different questionnaires for easier analysis, Excel software was used to integrate all the data for comparison.

## 4 Experiment Analysis of Mathematical Modeling and Optimization Design Based on Information Technology Algorithm

### 4.1 Comparison of Teaching Effects

In order to test the effect of the teaching method of mathematical modeling based on information technology, this paper conducts experimental teaching through two different classes in the same school. In ordinary classes, traditional teaching methods are used to conduct after-class surveys on students in two classes. The survey data are shown in Table 1.

**Table 1.** After-school survey of students in two classes

	Interested	Generally	Not interested
Experimental class	92	4	4
Classic class	6	42	52

It can be clearly seen from Fig. 1 that the teaching effect of the assisted teaching method based on information technology is far superior to the traditional conventional teaching method. After the experimental class was taught, 92% of the students were interested in mathematical modeling, but only 6% of the students in the conventional teaching method were interested in mathematical modeling.

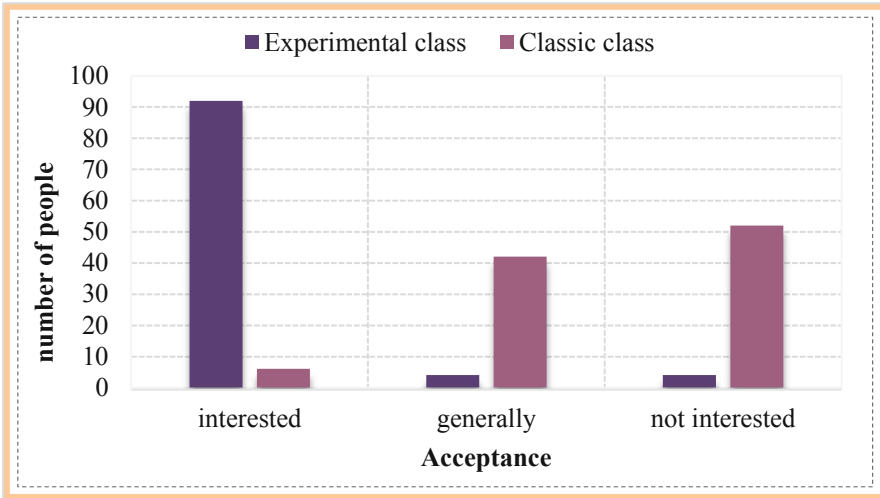


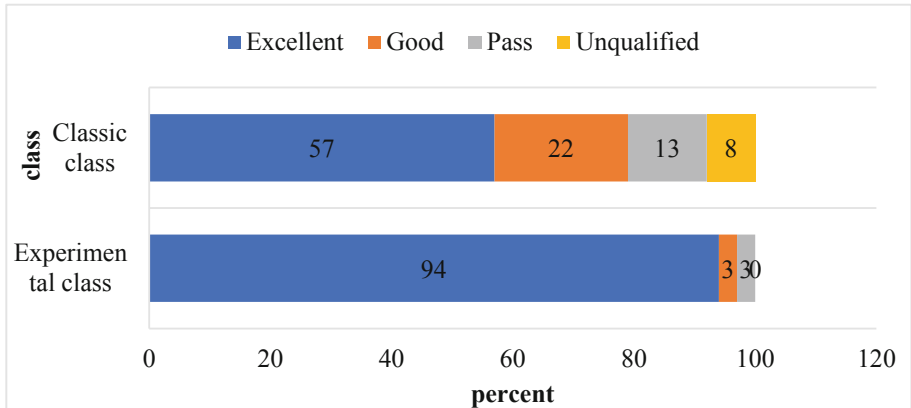
Fig. 1. After-school survey of students in two classes

#### 4.2 Comparison of Teaching Achievements

In order to confirm the teaching results of mathematical modeling based on information technology, this paper makes statistics on the final exam results of the two classes mentioned above. The statistical results are shown in Table 2.

Table 2. Statistics of final exam results for two classes

	Excellent	Good	Pass	Unqualified
Experimental class	94	3	3	0
Classic class	57	22	13	8



**Fig. 2.** Statistics of final exam results for two classes

From Fig. 2, it can be seen that the teaching results of mathematical modeling based on information technology are relatively good. Among them, 94% of them have excellent grades, and the pass rate has reached 100%, but the excellent rate of ordinary classes is only 57%, and the pass rate is only 92%. The comparison between the two is obviously that the mathematical modeling teaching method based on information technology is more excellent.

## 5 Conclusions

With the continuous development of information technology, it continues to influence traditional ideas, methods and teaching methods. Mathematics teaching organically combines traditional teaching methods with information technology, which changes the nature of mathematics teaching to a certain extent. To cultivate students' innovative ability, knowledge transfer alone is not enough. In the mathematics teaching explored in this paper, the establishment of students' mathematical modeling awareness and the cultivation of students' creative thinking ability are complementary and inseparable quality teaching requirements. Therefore, it is necessary to further mobilize students' subjective initiative, cultivate students' innovative thinking ability, and improve students' thinking ability and innovation ability.

## References

1. Xu, X., Zeng, Y., Li, Y., et al.: Minimum-latency FEC design with delayed feedback: mathematical modeling and efficient algorithms. *IEEE Trans. Wirel. Commun.* **PP**(99), 1(2020)
2. Redreev, G.V., Chervenчук, V.D., Chervenчук, I.V., et al.: Interface modeling algorithms for dispatch control. *IOP Conf. Ser. Earth Environ. Sci.* **624**(1), 012092 (6pp) (2021)
3. Okubanjo, A., Oluwadamilola, O.: Turkish journal of engineering dynamic mathematical modeling and control algorithms design of an inverted pendulum system. *Turkish J. Eng.* **3**(1), 14–24 (2019)

4. Zhao, P., Guan, H., Wei, H., et al.: Mathematical modeling and heuristic approaches to optimize shared parking resources: a case study of Beijing, China. *Transp. Res. Interdisc. Perspect.* **9**(4), 100317 (2021)
5. Yang, Z.: Hybrid global optimization method based on dynamic kriging metamodel and gradient projection method for optimal design of robot. *J. Mech. Eng.* **55**(11), 61 (2019)
6. Nahayo, F., Bagorizamba, A., Bigirimana, M., et al.: Predictive mathematical and statistical modeling of the dynamic poverty problem in Burundi: case of an innovative economic optimization system. *Open J. Optim.* **10**(4), 25 (2021)
7. Santibanez-Aguilar, J.E., Lozano-Garcia, D.F., Lozano, F.J., et al.: Sequential use of geographic information system and mathematical programming for optimal planning for energy production systems from residual biomass. *Ind. Eng. Chem. Res.* **58**(35), 15818–15837 (2019)
8. Khashirova, T.Y., Lamerdonov, Z.G., Zhaboev, S.A., et al.: Information technologies and mathematical modeling in solving environmental problems in the design of shore protection structures. *Ecol. Indust. Russia* **23**(9), 13–17 (2019)
9. Reymov, K.M., Turmanova, G.M., Makhmuthonov, S.K., et al.: Mathematical models and algorithms of optimal load management of electrical consumers. *E3S Web Conf.* **216**(1), 01166 (2020)
10. Lucay, F.A., et al.: Design of Flotation Circuits Using Tabu-Search Algorithms: Multispecies, Equipment Design, and Profitability Parameters. *Minerals* **9**(3), 181 (2019)
11. Penkovskii, A., Stennikov, V., Postnikov, I.: Unified heat supply organization: mathematical modeling and calculation. *Energy Proc.* **158**, 3439–3444 (2019)
12. Darandis, N., Nazari, M.: A new mathematical modeling and sub-optimal chemotherapy of cancer. *J. Biol. Syst.* **29**(03), 647–685 (2021)



# Internet of Things Data Mining Technology in Sports Physical Health Analysis

Qian Zhang<sup>(✉)</sup>

Physical Education School, Wuhan Business University, Wuhan, Hubei, China  
3609296@qq.com

**Abstract.** To improve the physical health (PH) of students, it is necessary to carry out physical exercise scientifically on the basis of ensuring the physical exercise time and intensity of students. This paper takes students from a certain university as the research object, and collects statistics on the physical fitness (PF) test data of the students in the school from 2017 to 2021. The data includes students' BMI index, vital capacity, 50 m and standing long jump scores, and use the Internet of Things data mining technology (DMT) to analyze the effective data information, and use Excel software to draw charts, to obtain the PH qualification level of the students in the school, the results of the PH of boys and girls are qualified and above accounted for 65.5% and 66.6% respectively, but there are still many students who are physically healthy unqualified, so the school should strengthen students' physical exercise. By analyzing the reasons that affect students' PH level, this paper puts forward relevant suggestions, hoping that the school and other colleges and universities (CAU) can learn from the PH improvement path of this paper to improve students' awareness of exercise and physical quality.

**Keywords:** Internet of Things Data Mining Technology · Physical Health · College Students · Physical Exercise

## 1 Introduction

Schools play a very important role in the education process of students' PH and are the main body of responsibility for students' PH education. Schools need to pay attention to the importance of sports, and let students actively participate in sports, which can not only exercise and promote PH, but also allow students to relax when facing the pressure of cultural courses [1].

At present, the data mining technology based on the Internet of Things has been widely used in the analysis of PF and health. For example, some medical scientists have done research on lung capacity. The research shows that the lung capacity must reach a certain target at a certain age. If it is small, it will have an impact on people's PH. Therefore, lung capacity is a very important indicator to measure whether a country or a nation is healthy and long-lived [2, 3]. A school uses DMT to analyze teachers' cognition of the importance of students' PH. From the analysis results, basically all physical education teachers believe that physical exercise is very important for students,



and more than 90% of school leaders and other teachers believe that physical education is important to students. This shows that the relevant leaders and teachers of the school clearly realize that physical exercise and physical fitness are urgent for students. To improve students' PH and enhance physical quality can ensure learning progress [4, 5]. Although the application research results of DAT in PH analysis are good, the analysis results can be more accurate only by combining data with other technologies or software.

This paper first introduces the concept of IoT DMT, and then uses this technology as a support to analyze the PH data of college students in the past five years, and then propose solutions according to the reasons that affect students' PF.

## 2 IoT Data Mining Technology

Generally speaking, the concept of data mining refers to the process of extracting undiscovered, user-interested and valuable knowledge from a large amount of irregular data. Gain potential value [6]. The Internet of Things DAT combines the Internet of Things and big data technology to mine effective information and data from the vast information chain of the Internet of Things [7]. Common data mining methods include cluster analysis, regression analysis, Web page mining, decision tree algorithm, association rule algorithm, etc. This article introduces the association rule algorithm, which represents the association between things.

The support formula is:

$$Sup(A) = \frac{Sum(A)}{N} \quad (1)$$

The confidence formula is:

$$Conf(A \Rightarrow B) = \frac{Sup(A \cup B)}{Sup(A)} = \frac{Sum(A)}{N} \quad (2)$$

Where A and B are transactions, and N is the total number of data.

## 3 Experimental Research

### 3.1 Research Content

This paper takes a student in a general college as an example, analyzes the student's BMI index, lung capacity, 50-m score, standing long jump score and the overall situation of PH, reflects the student's PH level, and proposes solutions by analyzing the factors that affect PH, in order to improve the physical quality of students.

### 3.2 Research Methods

Mathematical statistics method: Collect the PF measurement data of students in the school in the past five years, use the Internet of Things DAT to mine effective data, and then use Excel and other analysis software to make graphs to intuitively analyze the physical changes of ordinary college students (CS).

Comparative analysis method: This study selects the school's sports and PH test data from 2017 to 2021, and longitudinally compares the change trend of students' PH in the past five years.

## 4 Analysis of College Students' PF and Health Status Based on IoT Data Mining Technology

### 4.1 Current Status of Physical Fitness

(1) Comparative analysis of student BMI index.

**Table 1.** BMI index situation

		2017	2018	2019	2020	2021
Male	Height (cm)	176.28	176.54	176.62	176.93	177.09
	Weight (kg)	67.73	68.17	68.85	68.64	69.12
	BMI	21.80	21.87	22.07	21.93	22.04
Female	Height (cm)	163.87	164.05	164.36	164.19	164.26
	Weight (kg)	55.62	55.54	55.67	55.71	55.63
	BMI	20.71	20.63	20.61	20.67	20.62

As shown in Table 1, the overall body shape of the school's students has changed in the past five years from 2017 to 2021. Through comparative analysis, we can see that for boys, from 2017 to 2021, the overall height of male CS has increased. The trend has increased from 176.28 cm per capita in 2017 to 177.09 cm in 2021; at the same time, the overall weight has also shown an upward trend, increasing from 67.73 kg per capita in 2017 to 69.12 kg in 2019. For girls, from 2017 to 2021, the height of female college students showed a wavy change, with the highest average height in 2019 being 164.36cm. Body weight also changed in waves overall.

To sum up, the students of this school are in a normal state in terms of body shape from 2017 to 2021, and the BMI index of boys is on the rise as a whole, which is currently within the normal range. The BMI index of girls showed irregular changes as a whole. From 2017 to 2021, the body shape of the students in the school was generally within the normal range (normal value range: 17.9 ~ 23.9), and there was no obvious abnormality. Due to the rapid economic development, people's living standards and quality have been improved, and people's dietary structure and nutritional conditions have undergone tremendous changes. Although the average body mass index of BMI is in the normal range, and the students' physical development is at a normal level, some students are overweight or underweight, which is related to the students' own physical exercise and eating habits to a certain extent. Thin and obese students should pay attention. Obesity is not only related to body shape, but also to personal health.

(2) Student's vital capacity analysis.

Figure 1 shows the physical and healthy vital capacity of the school's students. Through comparison, it can be seen that from 2017 to 2021, the vital capacity of the male CS in the school showed a trend of first decreasing and then increasing. The lung capacity of male CS has decreased from 4563.73 ml per capita in 2017 to 4467.47 ml in 2019, and has been on a decreasing trend in the past three years; it has increased

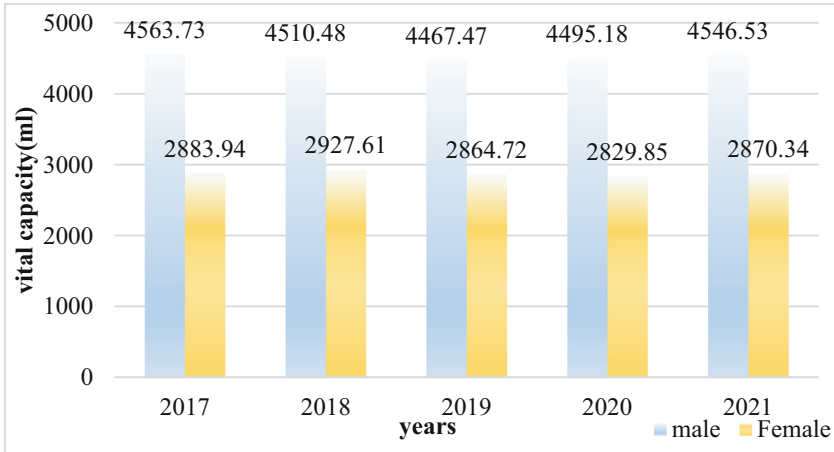


Fig. 1. Lung capacity (ml)

from 4467.47 ml in 2019 to 4546.53 ml in 2021. In the past three years, male CS have The lung capacity is gradually increasing. From 2017 to 2021, the changes in the lung capacity of female CS in the school were in the form of waves, first increasing, then decreasing, and then increasing. In 2017, the per capita vital capacity of female CS was 2883.94 ml, and in 2018, the per capita vital capacity of female CS was 2927.61 ml, which has been increasing in the past two years.

The average lung capacity of boys was significantly higher than that of girls, and the trend of boys was to increase first and then decrease, while girls fluctuated more and decreased. It may be because boys participate in more physical activity than girls, and most girls do not like vigorous exercise and prefer quieter activities, so they have less exercise in lung capacity. Aerobic exercise has a great impact on the improvement of lung capacity. In the teaching process, schools should focus on cultivating students' interest in sports, carry out various forms of exercise, make full use of time to strengthen exercise, and improve cardiopulmonary function.

(3) 50 m and standing long jump.

Table 2. Statistics of 50 m and standing long jump

		2017	2018	2019	2020	2021
50 m	Male	7.3 s	7.8 s	7.5 s	7.0 s	7.2 s
	Female	9.4 s	8.9 s	9.1 s	9.2 s	9.5 s
Standing long jump	Male	2.36 m	2.41 m	2.43 m	2.46 m	2.60 m
	Female	1.72 m	1.68 m	1.74 m	1.75 m	1.71 m

As shown in Table 2, the 50-m and standing long jump scores of the school's students from 2017 to 2021. In the 50-m score, the average time for boys is about 7.5 s, and the

boys' scores show a downward trend; the average time for girls is about 9.0 s, And girls' performance showed a trend of first decreasing and then increasing. In the standing long jump, the average performance of boys is on the rise, while the performance of girls is constantly fluctuating.

(4) General condition of physical health.

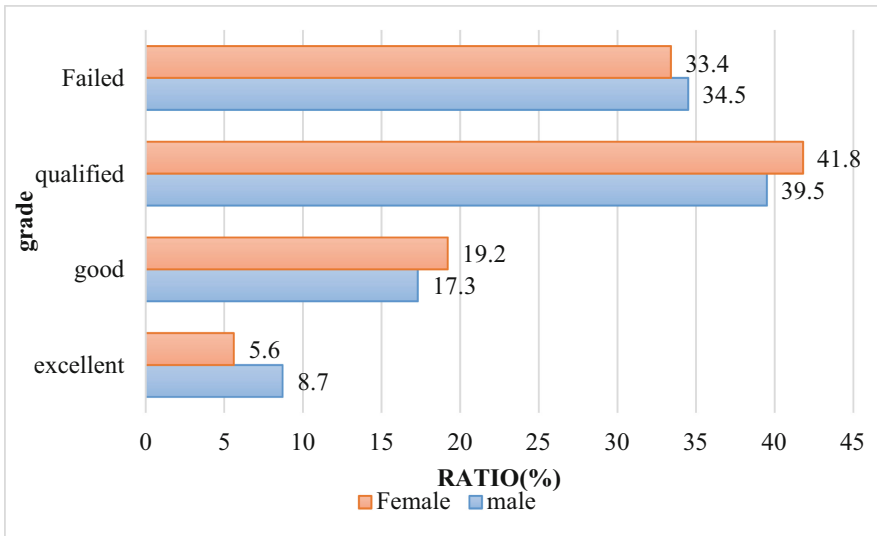


Fig. 2. Physical health pass rate (%)

From Fig. 2, it can be seen intuitively the grade distribution of the PH measurement results of the school's students. The proportion of qualified and good students exceeds half of the total number of students, but the proportion of students with excellent PF measurement results is less than 10%. As a result, more than one-third of the total number of students failed, and the physical health of students is worrying. It shows that the PH of the students in this school is relatively poor, and it needs to be paid full attention by the school and parents. Teachers and parents should pay more attention to strengthening sports and improving students' physical and health quality while educating students to learn cultural lessons.

## 4.2 Reasons Affecting Students' Physical Fitness

(1) Diet and rest problems.

After CS entered the relatively free university environment from the collective unified and regular work and rest environment of high schools, some students could not achieve regular work and rest, skip breakfast, eat irregularly, overeating, picky eaters and partial eclipses, etc. Eating habits and nutritional intake are problematic. Unable to guarantee a reasonable diet makes the university appear fat, thin, malnutrition and other problems. Nowadays, many students are picky eaters or their

parents have weak awareness, and their diet is not nutritious. Many students even never eat breakfast, and their nutritional intake is unbalanced, which has a negative impact on the growth and development of students [8, 9].

The academic burden of CS is relatively light, and the time at their disposal will increase. Some students with low self-control will develop bad habits such as smoking, drinking, staying up late to play games, etc., which are the living habits of college students. Smoking and drinking are important factors affecting human health, and staying up late accelerates damage to human health. Staying up late for a long time can lead to problems such as low concentration, lethargy, and weakened immunity. The bad work and rest habits of college students lead to the decline of students' PH [10].

- (2) Lack of sports venues and fitness equipment.

Hardware facilities such as sports venues and fitness equipment are essential in physical education courses. It is related to whether students can have a venue for sports activities and whether they can have various physical exercise projects on the venue. For teachers of physical education courses, sports venues and fitness equipment are necessary conditions for physical education, and for students, they are an important material basis for participating in sports [11, 12]. During the holidays, students will not be able to find suitable sports venues because of the closed school sports venues, the long distance of social sports venues, and the high fees, so they cannot exercise.

- (3) Lack of students' awareness of physical exercise.

Students have deviations in their understanding of sports and physical education. Exercise should reach a certain intensity and time. However, students think that physical education and sports are after-school pastimes for cultural learning, and subconsciously equate physical education with rest and sports. In terms of free activities, they subjectively resist the physical education teacher's skill teaching or intensity training in class. Aerobic endurance training has a very good effect on the improvement of students' physical fitness. Students who have undergone aerobic endurance exercise prescription exercise will significantly improve their PF.

### 4.3 Ways to Improve Students' Physical Health

- (1) Develop good living habits.

Students' healthy eating and rest habits are closely related to their physical health. Adequate intake of nutrition every day can provide middle school students with necessary energy, maintain normal growth and development, and provide them with sufficient physical energy for physical exercise. Schools and families should strengthen students' physical health education and supervision, improve children's awareness of physical health and participation in physical exercise, and help students formulate exercise plans and supervise their implementation. In daily life, we should not be picky about eating and exercise, keep balanced nutrition, and comprehensively develop various bodily functions and qualities.

- (2) Increase the number of full-time teachers and equip the venue with equipment and facilities.

Lack of full-time physical education teachers is one of the conditions that affect students' participation in activities. Venue and activity equipment are the most basic material conditions for junior high school students to participate in physical fitness activities. The shortage of venue equipment will indirectly affect the physical health of students. Relevant departments should increase investment in school sports funds, and school funds must be allocated an appropriate proportion for school physical education work. In addition, according to the actual conditions of the school, the equipment and facilities related to physical exercise are assembled and the sports venues are perfected, so that the needs of campus sports activities and physical education classes can be met.

(3) Strengthen students' ideas of physical exercise.

In order to effectively improve the PH of students, we must first start with the subject, and we need to change the students' ideology. Mobilize their minds during class meetings, physical education classes and inter-class activities, let students combine their preferences and make full use of their own advantages, and improve their interest in sports. With the interest in physical exercise, students will consciously take the initiative to play sports.

## 5 Conclusions

This paper focuses on the physical fitness and health status of CS in a CAU, and uses the Internet of Things DAT to calculate the BMI index, lung capacity, sports performance and PF pass rate of the students in the school, but the excellent rate of students' PH is low, which is caused by factors such as students' diet, irregular work and rest, lack of luck equipment, and weak concept of exercise. Therefore, this paper puts forward targeted suggestions for these three reasons, hoping to effectively help students improve their PF level.

**Acknowledgments.** This work was supported by Philosophy and Social Science research project of Hubei Provincial Department of Education (21G116).

## References

1. Ismail, W.N., Hassan, M.M., Alsalamah, H.A.: Mining of productive periodic-frequent patterns for IoT data analytics. *Future Gener. Comput. Syst.* **88**, 512–523 (2018)
2. Piccialli, F., Cuomo, S., Jeon, G.: Parallel approaches for data mining in the internet of things realm. *Int. J. Parallel Prog.* **46**(5), 807–811 (2018). <https://doi.org/10.1007/s10766-018-0565-y>
3. Lammatha, K.K.: Data mining on 5G technology IOT. *Int. J. Sci. Res. Manag. Stud.* **8**(5), 24655–24660 (2019)
4. Kongkaew, C.,L., et al.: The effects of Thai yoga on physical fitness: a meta-analysis of randomized control trials. *J. Altern. Complement. Med. Res. Paradigm Practice Policy* **24**(6), 541–551 (2018)
5. Espada, M., Figueiredo, T., Ferreira, C., et al.: Body composition and physical fitness analysis in different field position U-15 soccer players. *J. Phys. Educ. Sport* **20**(4), 1917–1924 (2020)

6. Park, M.H., Kang, M.S., Choi, D.S.: Comparison analysis of isokinetic muscle function, aerobic and anaerobic exercise ability, and basic physical fitness according to the gender of middle School Taekwondo Athletes. *J. Korean Soc. Living Environ. Syst.* **28**(3), 261–270 (2021)
7. Gh, A., Jec, B., Jo, C., et al.: Health-related physical fitness and weight status in 13–to 15–year–old Latino adolescents. A pooled analysis. *Jornal de Pediatria (Versão em Português)* **95**(4), 435–442 (2019)
8. Choi, D.S.: Comparison analysis among the event groups in physical fitness factors of elite track and field jumpers. *Korean J. Sports Sci.* **29**(5), 1373–1383 (2020)
9. Kim, S., Kim, J.: A comparative analysis of physical fitness in Korean police officers: focus on results between 2014 to 2019. *Exer. Sci.* **28**(4), 396–400 (2019)
10. Kim, K.J., Song, H.S., Chun, B.O., et al.: An analysis of physique and physical fitness according to skeletal age and chronological age in youth baseball players. *Korean J. Growth Dev.* **26**(4), 357–365 (2018)
11. Kim, H.K., Chai, J.H.: Physical fitness and characteristic analysis of Korean national prospective badminton team members stratified by gender and game type. *Korean J. Sports Med.* **38**(2), 95–100 (2020)
12. Hart, P.D.: Multivariate analysis of vertical jump predicting health-related physical fitness performance. *Amer. J. Sports Sci. Med.* **6**(4), 99–105 (2018)



# Application of Artificial Intelligence Technology in Big Data Mining

Xueyun Zhou<sup>1</sup>, Yongjun Qi<sup>2,3</sup>(✉), and Hailin Tang<sup>2,3</sup>

<sup>1</sup> Department of Computer Science and Engineering, Guangzhou College of Technology and Business, Guangdong 510850, China

<sup>2</sup> Faculty of Metadata and Computing, Guangdong Baiyun University, Guangdong 510450, China

qyj200702022@baiyunu.edu.cn

<sup>3</sup> Mongolian University of Science and Technology, Bayanzurkh District, Ulaanbaatar 13341, Mongolia

**Abstract.** Data mining is the process of expressing knowledge through data collection and processing. This is an important technique for efficiently extracting batch data. Artificial intelligence (AI) is a simulation technology of the human brain. It makes full use of computer technology and other subject technologies for intelligent learning and machine control. Both are advanced data management technologies, and both have broad application prospects. This article mainly introduces the clustering algorithm and TextRank algorithm. This paper uses AI technology to conduct research and application in big data mining (BDM), and establishes a mathematical model of the underlying TextRank algorithm. The model is solved by the TextRank algorithm, and the status of research and application of AI technology in BDM is assessed. Experimental results show that the algorithm improves the research and application of embedded big data intelligent technology by 55% and reduces false positive rate and false positive percentage. Finally, by comparing the research and application analysis of artificial intelligence technology in large data storage.

**Keywords:** Artificial Intelligence Technology · Big Data Mining · TextRank Algorithm · Clustering Algorithm

## 1 Introduction

### 1.1 Background and Significance

With the continuous development of computer information technology, the amount of data in all walks of life is growing explosively [1]. Big data has penetrated into all aspects of life, but there is still a huge amount of unknown information and value hidden behind the data, and its quantifiable value is almost equivalent to natural resources [2]. Therefore, such huge data resources urge various industries to carry out data-based reform and establish standard data storage, analysis and utilization specifications. As an important means to promote digital construction, data mining technology mainly



uses clustering, classification, prediction and correlation methods to analyze massive raw data, mine potential value information, and guide decision-making [3]. At present, it is widely used in many fields such as medical treatment, finance, retail and so on. The algorithm analyzes the internal relationship between parameters and indicators by mining the potential association rules between data [4]. With the acceleration of industrial Internet of things, how to extract useful rules from multi-dimensional, heterogeneous and massive industrial control processes for regulating and controlling the entire industrial production is a difficult problem in process control [5].

## 1.2 Related Work

Li s provides a method of evaluating stakeholder participatory innovation in a complex multi-stakeholder environment for solving key problems [6, 7]. Based on the principle of shared value creation, it proposed an evaluation framework, which describes the process of social interaction and integrates their own resources and capabilities for the development of innovative products and services [8, 9]. In order to evaluate this evaluation framework, a large amount of data was collected in the study. This case represents the environment of many stakeholders involved in the research and application of artificial intelligence technology in large data mining [10, 11]. However, due to the complexity of the message collection process, the data results are not very accurate.

## 1.3 Main Content

The innovation of this work is in the proposed clustering algorithm and TextRank algorithm. Based on the research and application of AI technologies in BDM, evaluating the research and application of AI technologies in BDM. Create a calculation method for the TextRank algorithm in combination with a clustering algorithm to provide guidance for researching and using AI technology in large data mining.

# 2 Research and Application Methods of AI Technology in BDM

## 2.1 Clustering Algorithm

As a multidisciplinary field, data mining can be defined in many ways. Even the definition of the term “data mining” does not fully cover the rich content it contains. Strictly speaking, “data mining” in the industry is a broad concept, which should be more accurately called “knowledge mining from data” or “knowledge discovery in data”. The narrow sense of “data mining” is only a basic step in the whole process of knowledge discovery.

After years of development, data mining algorithms have been divided into many types, among which semi-supervised learning includes the clustering algorithm studied in this paper. Clustering divides unlabeled data into multiple clusters or classes composed of several similar objects.

The clustering process divides the data objects of the data set into several different clusters, which meets the clustering requirements as if the similarity of a cluster is the

largest. If it is regarded as a global optimization problem, it means that the optimization method can be used for clustering. Different data sets and different clustering rules may lead to large deviation of clustering results. In order to analyze and evaluate the clustering results, people have proposed clustering evaluation indicators with different characteristics. These evaluation indexes can be used as the optimization objectives of optimization method clustering.

The problem of clustering analysis is also included in the problem of data classification. Its classification model has the attribute of unknownness. If it is not clear how to classify, the data can be re-analyzed by clustering algorithm according to the difference in the intrinsic attributes of the data. Appropriate grouping and classification, and finally get the generated objects. For each cluster type, the similarity of the internal data is relatively high. On the contrary, for the cluster type and the cluster type, the data is basically not connected.

## 2.2 TextRank Algorithm

The TextRank algorithm is mainly used to generate text and abstract keywords, and the page ranking algorithm is used to calculate the importance of web pages. If the entire network is regarded as an LED graph, each web page is a node in the LED graph, and there is a link between each web page, and the link can be expressed from one web page to another web page in the corner. Several directed edges constitute a complete network. Then the importance of web pages can be calculated by formula (1):

$$|A(N_i) = (1 - d) + d^* \sum_{j \in \ln(N_i)} 1/Out(N_j)^s(N_j) \quad (1)$$

Among them,  $s$  is the frequency of words appearing in the text,  $S$  is the total number of texts, and  $N$  is the number of texts where the word  $i$  appears.

$$E = B_j \log\left(\frac{n}{m_i}\right) \quad (2)$$

The calculation of the TextRank algorithm fully considers the independence of a single word from the overall document, but in actual situations, if a word appears more frequently in the text, then the word should well reflect the text characteristics and should be given a higher weight, this also leads to the lack of TextRank algorithm.

## 3 Research and Application Experiment of AI Technology in BDM

### 3.1 State Design of AI Technology Applied to Big Data Mining

It can be seen development data processing and technical analysis involves integration, network and complexity. Integration mainly the continuous integration of many technical tools with open interdisciplinary and interdisciplinary phenomena. The network gives full play to the main role of the network, which can expand the processing capacity of terminal equipment indefinitely and form powerful management and control capabilities.

In many ways, complexity means that different technical solutions are not limited to the computer field, but can also be effectively used in business models, industry control and financial decision-making. Big data design is perfectly driven. Literally, it is based on data as a design reference and foundation. Before data-driven design, you need to understand two different types of data. For data research, researchers can analyze data through various research methods. Producers’ motivations and behaviors to study their hidden needs, and then check the results of data analysis and provide appropriate solutions as soon as possible. Research data should start with qualitative data and quantitative data.

### 3.2 AI Technology Applied to Data Collection in BDM

Today, data mining technology has been applied to many fields such as banking, telecommunications, insurance and data mining, and can solve many common business problems. The steps to process the above information using AI are as follows. Firstly process the original data, fill in the remaining, eliminate abnormal data, etc., then use algorithms to summarize and abstract the original data, and finally use data mining. The basic content is AI technology applied to BDM data. The data is shown in the Table 1.

**Table 1.** Data Table of AI technology applied to BDM

AI	Mean standard	Deviation	Variance value	Error rate
Data collection	97	24	12	40%
Data mining	87	48	22	38%
Data Classification	95	25	5	29%

From Table 1, we can get the relationship between data mining and AI. Knowledge representation generally refers to a description of knowledge by a computer, and is a description structure of data that can be accepted by a computer. Although there are still many problems in the construction of data knowledge system, there are still some specific knowledge representation methods in the process of researching AI: symbolic representation and connection mechanism representation. These two representation methods use data mining technology to varying degrees. As far as AI technology itself is concerned, its development should also be combined with people’s realities in life, and then continue to improve existing technologies. As a kind of big data technology, data mining technology has certain limitations, but it can still provide the necessary impetus for the current AI development.

## 4 Research and Application AI Technology in BDM

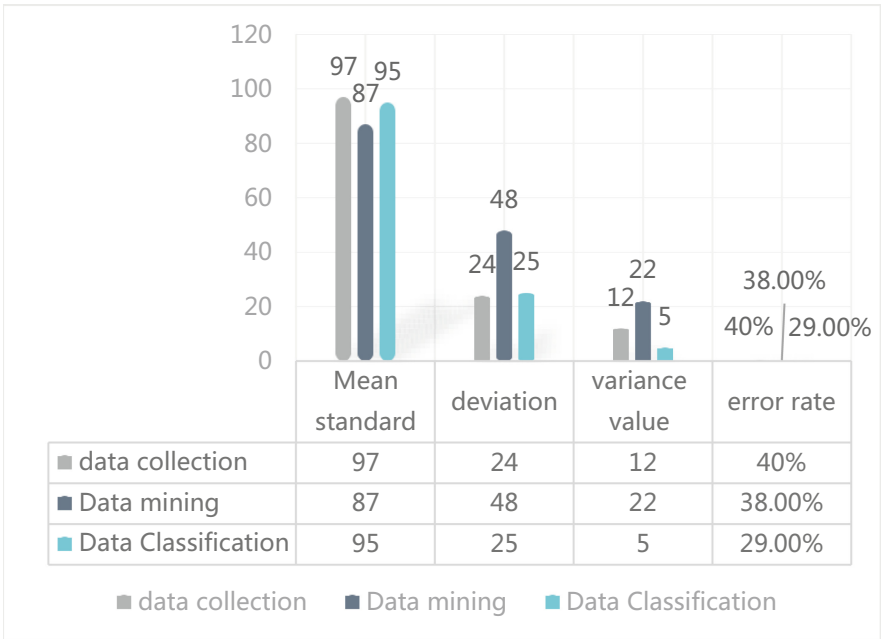
### 4.1 AI Technology in BDM

AI sample recognition means using computers to replace people or make people feel like patterns. This is a simulation of people’s perception of the outside world. Research

is a computer pattern recognition system, which is a technology that forces a computer system to imitate people to receive external information and recognize and understand the environment through perception. The first major achievement of AI was the development of chess programs that could solve problems.

Technology has made some progress in this field, including machine learning, neural networks, computational consciousness and evolutionary computational data. Intelligent learning is the basic way for computers to become intelligent. Starting from the study of the problems related to the robot arm, and then reaching the best planning technology to obtain the best sequence of robot motion as the goal, and finally successfully creating artificial life. In the future, the successful development of intelligent AI life will surely be a sign of breakthroughs in AI technology.

The internal relationship between the data is clearer. The results of the analysis are presented in Fig. 1:



**Fig. 1.** The inner relationship between the data more clearly and the analysis results

As shown in the above figure, the increase in the number of nodes in the complex will lead to a linear increase in the acceleration speed of the algorithm. Compare the performance of data collection, data mining, and data classification through data. Set support to 0.23 for frequent data mining through data collection, data mining and data sorting.

### 4.2 Experimental Data AI Technology in BDM

Data mining is closely related to AI technology, and even many key technologies are consistent with each other. In particular, there is a high degree of consistency between reasoning and data exploration. Whether it is traditional reasoning, non-logical reasoning, internal reasoning or model reasoning, it has high value. The basic principle is the accuracy of reasoning. The efficiency and processing of manual data is of great significance. Search engines are fully reflected in the data mining process. It will continuously search for existing paths according to the needs of users, and set up a process for calculating cheaper conclusions. The efficiency of data retrieval directly determines the speed of data retrieval. It can be seen that there are many connections between data production technology and AI technology, especially at the intersection of technology. Therefore, experiments are conducted on BDM. The experimental data are shown in Table 2.

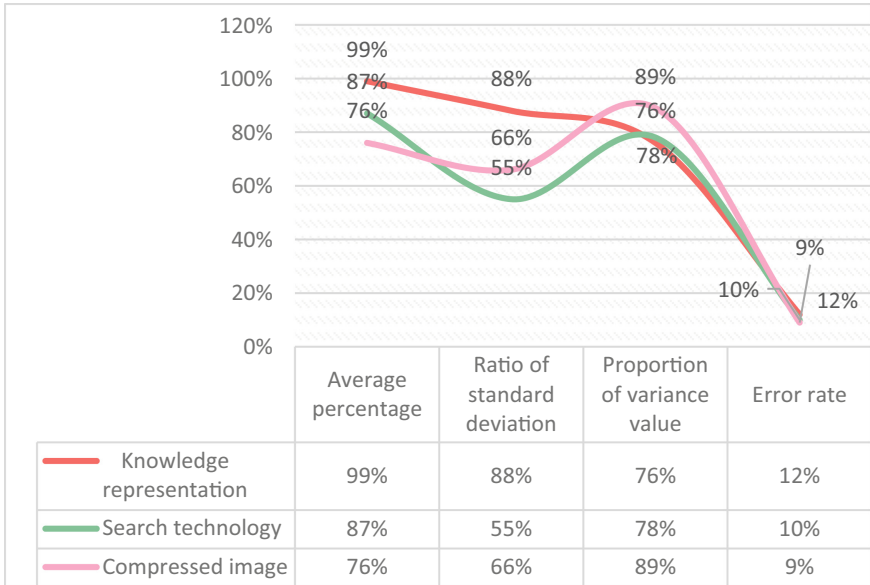
**Table 2.** Experimental data Table of AI technology in BDM

experimental project	Average percentage	Ratio of standard deviation	Proportion of variance value	Error rate
Knowledge representation	99%	88%	76%	12%
Search technology	87%	55%	78%	10%
Compressed image	76%	66%	89%	9%

Through the analysis of Table 2, the data mining in this work has formed a new data table, which is based on the development of education reform,

The data analysis results in the original database are shown in Fig. 2:

As shown in the figure, large data have high security, stability and portability. Knowledge representation, search technology and compressed image used in the logistics chain provide the requirement for decentralisation and non-deception, and use strong function and distributed storage function to solve problems in data mining.



**Fig. 2.** Experimental data diagram of AI technology in BDM

Through the analysis of data mining and AI technology, it can be seen that its future development direction is towards integration, networking and complexity. Integration mainly refers to the continuous integration of several technical means. The phenomenon of interdisciplinary and interdisciplinary is very obvious. Networking can make full use of the core role of the network, unlimitedly expand the processing capabilities of terminal equipment, and form powerful management and control capabilities. Complexity usually means that different technical solutions are not limited to the computer field, but can also be effectively used in business design, industrial management and financial decision-making.

## 5 Conclusions

Although this work has done some research on the grouping algorithm and the text algorithm. The research method and application of artificial intelligence technology to BDM still have a lot of content worth in-depth study. Due to the influence of personal ability and time, there are many steps in the decision-making process in this experiment that do not involve more detailed parts.

**Acknowledgements.** Special project in key fields of colleges and universities in Guangdong province, No. 2020ZDZX3009.

Key Research Platform of Guangdong Province: Low Code Platform Engineering Technology Development Centre, No. 2022GCZX009.

Key Research Project of Guangdong Baiyun College: Research on Text Detection and Recognition in Complex Scenes Based on Double Attention Mechanism.

## References

1. Bist, A.S., Febriani, W., Lukita, C., et al.: Design of face recognition AttendX for recording student attendance data based on artificial intelligence technology. *Solid State Technol.* **63**(2s), 4505–4518 (2020)
2. Kim, K.Y., Jin, H.J., Yoon, Y.A., et al.: Designing a performance certification test for automatic detection equipment based on artificial intelligence technology. *J. Appl. Reliab.* **20**(1), 43–51 (2020)
3. Kim, D., Cho, M.K., Shin, H.: The application of artificial intelligence technology in counseling and psychotherapy: recent foreign cases. *The Korean J. Couns. Psychother.* **32**(2), 821–847 (2020)
4. Aghaeipoor, F., Javidi, M.M., Fernandez, A.: IFC-BD: an InterpreTable fuzzy classifier for boosting explainable AI in big data. *IEEE Trans. Fuzzy Syst.* **PP**(99), 1 (2021)
5. Kotsenas, A.L., Balthazar, P., Andrews, D., et al.: Rethinking patient consent in the Era of AI and big data. *J. Am. Coll. Radiol.* **18**(1), 180–184 (2021)
6. Yunpeng, L., Ziqiang, X.U., Gang, L.I., et al.: Review on applications of ai driven data analysis technology in condition based maintenance of power transformers. *Gaodianya Jishu/High Voltage Eng.* **45**(2), 337–348 (2019)
7. Neumann, K., Waight, N.: Call for Papers: science teaching, learning, and assessment with 21st century, cutting-edge digital ecologies. *J. Res. Sci. Teach.* **56**(2), 115–117 (2019)
8. Maulina, E., Purnomo, M., Wicaksono, A.R., et al.: Analysis of the use of artificial intelligence technology on digital startups in Indonesia. *Int. J. Adv. Sci. Technol.* **29**(3), 750–758 (2020)
9. Kobayashi, K.: (7)Artificial intelligence technology and medical image processing. *No shinkei geka. Neurological Surgery* **48**(7), 654–664 (2020)
10. Japkowicz, N., Stefanowski, J.: An overview of concept drift applications, 91–114 (2016). <https://doi.org/10.1007/978-3-319-26989-4>. (Chapter 4)
11. Jha, S., Mishra, P.: A survey on the use of artificial intelligence technology and google tools for multilingual students sitting and studying in a common classroom. *Int. J. Comput. Sci. Eng.* **7**(6), 765–768 (2019)



# Enterprise Credit Rating Prediction Model Based on Data Mining Algorithm

Yingying Song<sup>(✉)</sup>

China West Normal University, Nanchong, Sichuan, China  
296551026@qq.com

**Abstract.** The role of corporate credit rating in my country's financial market cannot be ignored more and more. Its significance lies in ensuring the safety of investors' funds and reducing information asymmetry between companies, financial institutions and individuals. The purpose of this paper is to use the data mining algorithm to create a prediction model of enterprise credit, and to rate the credit of unrated enterprises and enterprises whose rating information has not been updated in time, so as to effectively improve the prediction accuracy. This document uses powerful data mining technology to extract implicit, unknown and potentially useful information from data to establish a credit rating model for private enterprises. In the empirical process, this paper selects 150 listed private enterprises, including 75 non-ST enterprises with good operation and 75 ST enterprises with abnormal financial situation. Comparing the model classification accuracy, it is found that the model classification accuracy is increased from 80% to 87.5% after the selection of indicators, the operation speed is accelerated, the indicators are reduced, and the explanatory ability of the model is improved.

**Keywords:** Data Mining · Credit Rating · Predictive Models · CART Algorithms

## 1 Introduction

Personal credit has become one of the important cornerstones of the operation of China's modern market economy, and the personal credit market in developed countries has also developed earlier, and a relatively complete personal credit evaluation system, evaluation organization, and personal credit information release website have been formed [1]. In contrast, China's credit evaluation business has just started for more than 30 years, and the development and growth of China's credit evaluation companies are still in the stage of continuous learning. It plays an important role in promoting the standardized development of China Credit Rating Corporation, regulating the communication channels between individuals and enterprises, improving the efficiency of market economic activities, and promoting the healthy development of China's market economy. Can not be ignored [2].

Many scholars at home and abroad have carried out research on the analysis of enterprise credit rating prediction model based on data mining algorithm. Cash D looks at whether credit rating agencies (CRAs) could play a greater role in ECs in achieving



these specific goals, both given their influence in the market and their approach to rating governance. Based on a normative methodology, on which to contextualize the issue and make recommendations for methodological changes that the CRA could develop. Findings found that CRA may play a greater role in achieving EC goals [3]. Liswaniso S is designed to evaluate the performance of Chi-Square Automatic Interaction Detector (CHAID) and Classification and Regression Tree (CART) data mining algorithms in predicting egg weight based on different egg characteristics such as egg length (EL), egg width, shell weight, etc. (SW), shell thickness (ST), egg white weight (AW), yolk height (YH), yolk width (YD) and yolk weight (YW). For this purpose, 364 domestic free-range eggs were used. A good fit test was performed to compare the predictive performance of these algorithms [4]. Tayfor NB primarily aims to use data mining techniques to classify cancer datasets into blood cancers and in vitro cancers based on predefined and meta-defined information obtained after blood and CT scans. Use WEKA data mining tools to perform 10-fold cross-validation, evaluate and compare different classification algorithms, extract important information from datasets, and accurately determine the most suitable predictive model [5]. Therefore, it is imperative to establish a corporate credit rating model and reasonably measure risks.

On the basis of domestic enterprise credit risk assessment research results, according to the characteristics of Chinese enterprises, using the powerful data mining method based on neural network, this paper conducts in-depth research on the theoretical model and technical issues of Chinese enterprise credit assessment, and proposes several meaningful results and provide targeted guidance. From the perspective of the actual use of enterprises, it is possible to make reasonable predictions on the credit scores of some enterprises, and individual investors can also use them easily and conveniently, which reduces the cost and reduces the trouble caused by information asymmetry, thereby reducing the possibility of individual investors. Economic losses. For the company, it can be implemented by checking the deficiencies of its own information in the model to improve the enterprise credit score.

## **2 Research on Enterprise Credit Rating Prediction Model Based on Data Mining Algorithm**

### **2.1 Meaning of Credit Rating**

- (1) Credit rating is a comprehensive reflection of an enterprise's performance and solvency.

According to the generation and development, credit evaluation can be divided into broad sense and narrow sense. In a narrow sense, credit evaluation is an evaluation of the company's repayment situation, contract performance and integrity level; the so-called social credit rating includes subjects in various fields; in a broad sense, credit rating refers to various market participants and various financial instruments. The issuer's ability and credibility to fulfill various financial commitments [6, 7].

- (2) The key content of credit rating.

Credit evaluation is divided into two aspects: debt repayment evaluation and debt repayment evaluation. Repayment ability refers to the ability of enterprises and

related parties to form capital flows in business activities, the ability of enterprises to realize debt activities and form capital flows, as well as the pressure of enterprises to repay debts and the pressure of capital payment necessary for normal operations. Compare [8]. Debt repayment ability mainly refers to whether the creditor is willing to repay the debt in time and whether the debt repayment history has a history of arrears [9, 10].

(3) Specific symbols for credit ratings.

The credit rating symbol indicates that the rating result is a simple and easy-to-understand method, and the credit rating level is strictly divided, such as AAA, AA, A, BBB, BB, B, CCC, CC, C, D and other symbols are generally used. Represents ten levels. Once investors understand the basic meaning of each symbol, they can understand the risk situation of the investment object [11].

## 2.2 Data Mining

Data mining, also known as knowledge discovery, data mining, etc., refers to the process of extracting correct, new, potentially useful and understandable knowledge from large amounts of data [12, 13]. Data mining can perform classification prediction, cluster analysis, association rules and sequence patterns, association analysis, anomaly monitoring and trend analysis [14].

From a business point of view, the role of data mining is to find “gold nuggets” in the “data mine” to help companies reduce unnecessary investments and improve return on capital. Innovative companies around the world have begun to use data mining techniques to identify who their most valuable customers are and redesign their product promotion strategies to achieve best sales at the lowest cost [15, 16].

## 2.3 Rating Model

The establishment algorithm of the rating model:

The first step is the comprehensive preprocessing of the data;

The second step selects important variables, establishes a CART decision tree, and extracts classification rules at the same time;

The third step is to determine the degree of influence of each attribute on the decision tree, that is, the weight of each attribute, according to the contribution of each attribute to the tree construction and pruning process;

The fourth step is to differentiate each attribute and determine the relative importance and rank in the specific attribute according to the list of attributes related to the decision attribute and the credit rating of the decision category.

# 3 Investigation and Research on the Prediction Model of Enterprise Credit Rating Based on Data Mining Algorithm

## 3.1 Selection of Samples

The samples selected in this paper are all listed companies in 2020, with a total of 150 samples, and all data are from the wind database. The selected listed companies are

divided into two categories, one is: ST companies, the financial status of such companies is abnormal, and the selected samples are 75; the other is: non-ST companies, with excellent operating performance and normal financial status, the selected samples The same is 75.

### 3.2 Selection of Indicators

For the selected 150 listed private enterprises, this paper selects 20 indicators from the six aspects of capital structure, profitability, operating ability, solvency, growth ability and cash flow, which can comprehensively and accurately reflect the risk status of enterprises.

The measurement method of the attributes of the CART algorithm is based on the GINI coefficient, the smaller the coefficient, the more reasonable the division. The categorical attribute C of the dataset S has m different discrete attribute values  $c_1, c_2, \dots, c_m$ ,

“That is, there are m categories of records in S, then its GINI coefficient is

$$Gini = 1 - \sum_{i=1}^m P_i^2 \quad (1)$$

Where  $P_i$  is the frequency of occurrence of category  $c_i$ . If attribute A is used to divide dataset S into two parts  $S_1, S_2$ . Then the GINI coefficient of this division is:

$$Gini(S) = (S_1/S) \times Gini(S_1) + (S_2/S) \times Gini(S_2) \quad (2)$$

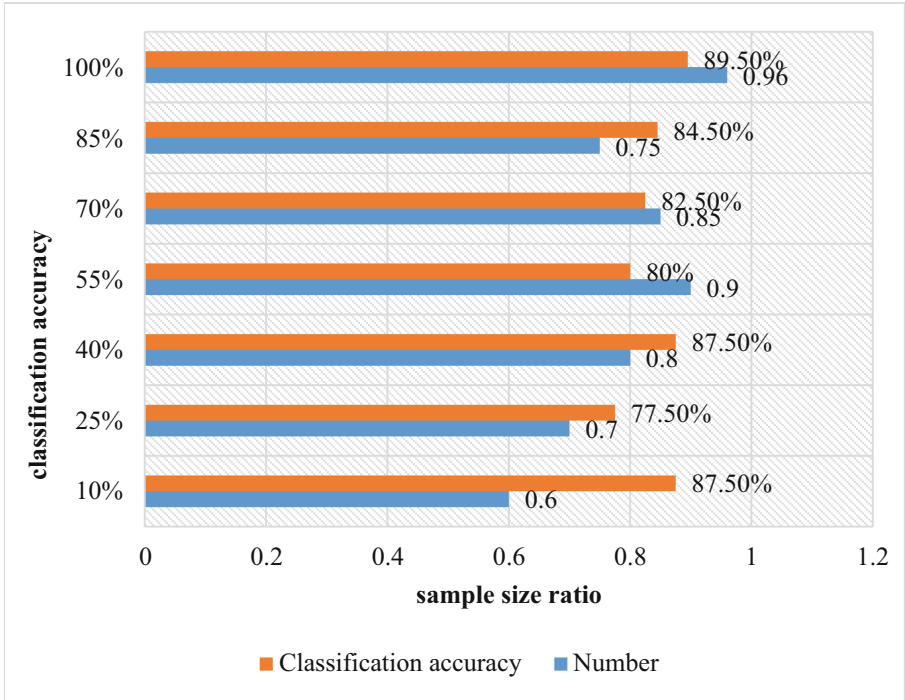
Among them,  $S_1$  and  $S_2$  are the number of samples of these two parts, respectively.

## 4 Analysis and Research of Enterprise Credit Rating Prediction Model Based on Data Mining Algorithm

### 4.1 Model Sample Analysis

There are 150 samples selected in this paper, and the training set contains 110 samples, including 55 samples in ST and non-ST classes; 55 samples in test set, and 26 samples in ST and non-ST classes. In the above-mentioned empirical analysis stage, a model is established based on such a sample situation and the classification accuracy of the model is analyzed. The following empirical process will test whether the model is robust under the above-mentioned sample conditions, that is, what is the minimum sample size, and the model can better classify the credit of enterprises.

As shown in Fig. 1, when the number of samples is small (20 samples), although the sample classification accuracy is greater than 85%, the model is over-fitting, and the model is still in an unstable state. As the number increases, the classification accuracy begins to decline. When the number of samples reaches 60, the classification accuracy fluctuates in a small range, and the average accuracy reaches more than 75%. The model is robust and can better rate enterprises. Therefore, this paper selects 150 samples to establish the model for the two types of credit rating of private enterprises is robust, and the result analysis is effective.



**Fig. 1.** Sample Increase Model Accuracy Analysis

## 4.2 Cost Sensitivity Analysis

In addition to the accuracy analysis of the model, it is also necessary to consider the cost of wrong decisions and misclassification. The actual credit rating of private enterprises is more concerned with the classification accuracy of ST enterprises, because the loss cost of misjudging ST enterprises with abnormal financial conditions as non-ST enterprises with excellent operating performance is far less than that caused by misjudging non-ST enterprises as ST enterprises. Loss, as shown in Table 1.

**Table 1.** Rating Accuracy Comparison

	Non-ST enterprise	ST enterprise	Sensitivity
Training samples	46	4	90%
Test sample	24	3	85%

As can be seen from Fig. 2, the classification accuracy of the model for training samples is 87.5%, of which the non-ST classification accuracy is 90%, and the ST business classification accuracy is 85%. In the test sample, the model's prediction accuracy for both ST companies and non-ST companies declined, with an overall sampling accuracy

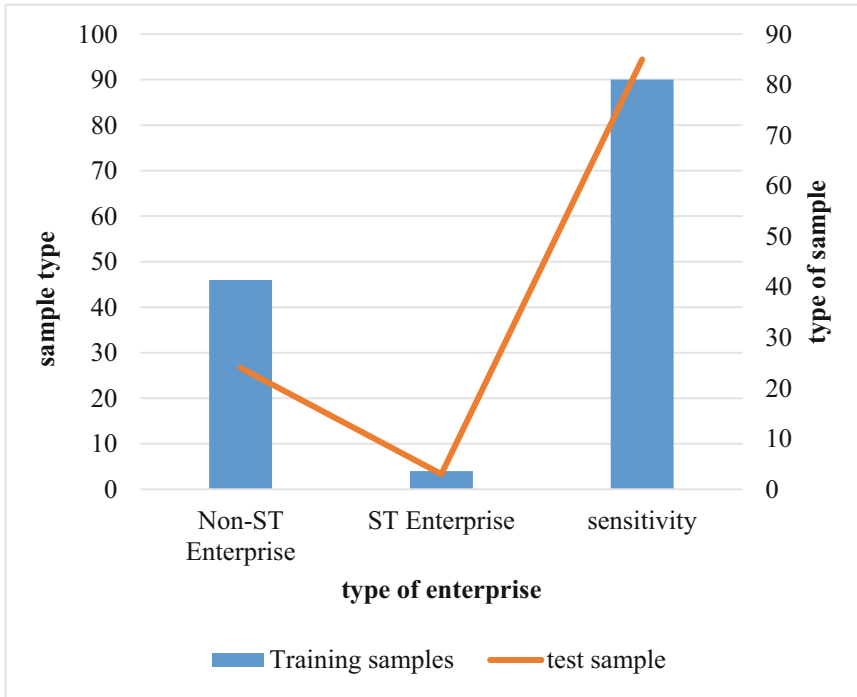


Fig. 2. Rating Accuracy Comparison Chart

of 87.5%. Assuming that the average classification accuracy of the model is greater than 85%, a certain range of accuracy drop is acceptable. Therefore, the model in this paper has high prediction accuracy and true sensitivity in ST and non-ST companies, and is relatively stable, non-ST companies and ST companies are classified.

## 5 Conclusions

With the gradual progress of our economic system reform, the enterprise economy is showing a rapid upward trend. With the expansion of business scale and lack of funds, their demand and desire for funds continues to increase, and internal financing methods such as profit retention and capital accumulation are increasingly unable to meet business development and external needs. Financing has gradually become a prerequisite for further industrial upgrading of private enterprises. However, due to the large adverse selection risk and moral hazard of the enterprise itself, the indirect financing channel of the enterprise is narrow. The difficulty of financing has seriously hindered its growth and has become a narrow point of business development. Therefore, enterprise credit rating index system and evaluation method, formulation of scientific and realistic private enterprise credit rating prediction model, and objective evaluation of enterprise credit rating are particularly important.

## References

1. Marjorie, C.: S&P lowers Northrop's corporate credit rating following orbital ATK deal. *Inside Pentagon* **34**(24), 13 (2018)
2. Silvederio, J.: S&P raises Warrior Met Coal's ratings on strong outlook for metallurgical coal. *SNL Energy Coal Rep.* **14**(3), 9–10 (2018)
3. Cash, D.: Can credit rating agencies play a greater role in corporate governance disclosure? *Corp. Govern. Int. J. Bus. Soc.* **18**(5), 954–964 (2018)
4. Liswaniso, S., Qin, N., Tyasi, T.L., et al.: Use of data mining algorithms Chaid and cart in predicting egg weight from egg quality traits of indigenous free-range Chickens in Zambia. *Adv. Animal Vet. Sci.* **9**(2), 215–220 (2021)
5. Tayfor, N.B., Mohammed, S.: A comparison study of data mining algorithms for Blood Cancer prediction. *J. Garmian Univ.* **3**(2), 174–179 (2021)
6. Alagrash, Y., Drebee, A., Zirjawi, N.: Comparing the area of data mining algorithms in network intrusion detection. *J. Inf. Secur.* **11**(1), 1–18 (2020)
7. Obiedat, R.: A comparative study of different data mining algorithms with different oversampling techniques in predicting online shopper behavior. *Int. J. Adv. Trends Comput. Sci. Eng.* **9**(3), 3575–3583 (2020)
8. Lewis, K.P.: Image processing techniques and data mining algorithms for coffee plant's leaves classification. *Int. J. Adv. Trends Comput. Sci. Eng.* **9**(2), 1101–1106 (2020)
9. Shin, C., Cho, Y., Park, J., et al.: An energy consumption prediction model for smart factory using data mining algorithms. *KIPS Trans. Softw. Data Eng.* **9**(5), 153–160 (2020)
10. Uzut, G., Buyrukoglu, S.: Hyperparameter optimization of data mining algorithms on car evaluation dataset. *Euroasia J. Math. Eng. Natural Med. Sci.* **7**(9), 70–76 (2020)
11. Chaudhari, P.: Redundant data normalization using the novel data mining algorithms. *Int. J. Adv. Trends Comput. Sci. Eng.* **9**(4), 6675–6683 (2020)
12. Uzut, G., Buyrukoglu, S.: Prediction of real estate prices with data mining algorithms. *Euroasia J. Math. Eng. Nat. Med. Sci.* **7**(9), 77–84 (2020)
13. Alam, M.I., Bharti, A.: Interpretation of data mining algorithms on healthcare decision support system. *J. Inf. Comput. Sci.* **10**(1), 46–53 (2020)
14. Mansoor, M., Ellahi, N., Malik, Q.A.: Corporate board attributes, shariah board attributes and credit rating: evidence from islamic banks of Pakistan. *Turkish J. Islamic Econ.* **8**(1), 19–34 (2021)
15. Kisgen, D.J.: The impact of credit ratings on corporate behavior: evidence from Moody's adjustments - ScienceDirect. *J. Corp. Finan.* **58**, 567–582 (2019)
16. Prayagsing, C.C.: Modeling the behaviour of corporate demand for credit in mauritius before and after the global financial crisis under conditions of excess liquidity by banks. *Arch. Bus. Res.* **8**(5), 63–79 (2020)



# Application of Machine Learning Algorithm in Financial Market Risk Prediction

Juyi Wang<sup>(✉)</sup>

Shanghai Maritime University, Shanghai, China  
lyndsey0727@163.com

**Abstract.** Due to the complexity and openness of financial market system, there are complex relationships among its internal economic variables. Understanding the fluctuation law of the financial market is of great value for improving people's ability to prevent financial risks. This paper mainly studies the application of machine learning algorithm in financial market risk prediction. In this paper, LSTM algorithm is analyzed firstly, and the forgetting gate, input gate and output gate of repetitive structure are analyzed in detail. LSTM xgboost hybrid forecasting method is designed by combining LSTM and xgboost methods. The traditional LSTM algorithm is compared with the improved LSTM algorithm by using the collected data of Shanghai Stock Exchange Index and Shenzhen Stock Exchange Index. The results show that the improved LSTM algorithm is superior to the traditional machine learning algorithm.

**Keywords:** Machine Learning · Financial Market · Risk Prediction · Lstm Algorithm

## 1 Introduction

With the rapid rise of economic globalization and digital economy, it has promoted the rapid growth of the world economy. As a key link in the financial industry, the stock market will affect economic fluctuations. How to effectively predict the stock price trend has become the focus of many scholars and investors [1]. In the process of forecasting stock prices, accurate, stable and effective models are very necessary. Because the predictability of stock prices has been verified, many researchers turn to predict stock price fluctuations. Then, some scholars use a single model to predict stock prices. However, because the trend of stock changes is nonlinear and there are many factors that affect the fluctuation of stock prices, it is unilateral to rely on a single model to predict. With the emergence of this problem, researchers applied portfolio model to stock forecasting and achieved good forecasting results [2]. This paper uses artificial intelligence technology to continuously explore the combination model with better financial time series prediction effect, which meets the application requirements under the current environment, and has certain theoretical significance and application value. How to predict the change trend of stock price has become the focus of many scholars and investors [3]. The rise and fall of stock prices are affected by many complex factors. Due to the nonlinearity and

irregularity of the stock time series, how to accurately predict the stock time series is still a prominent problem faced by modern social economy and social organizations. Therefore, accurate prediction of stock time series is generally considered as one of the most challenging world problems. With the continuous development of the field of artificial intelligence, many researchers use modern technology to predict financial trends.

Stock market is one of the key variables in the field of financial market. The fluctuation and trend of stock price has always been the focus of scholars and financial participants at home and abroad. In order to study stock price fluctuations more reasonably and improve the accuracy of prediction, scholars have put forward many prediction models [4]. Such as the linear regression model and the gray prediction model, the self-reflective model under heterocentricity conditions (arcs), the generalized self-reflective model under heterocentricity conditions (GARCH), the Self-Reflective Moving Mean Model (ARMA) and the Differential Self-Reflective Moving Mean Model (ARMA) commonly used in econometrics, Vector Support Machine (SVM), the neural network (NN) and other algorithms commonly used in the field of computer science, as well as fractal theory and methods [5]. Some scholars use the machine learning model to calculate the stock risk premium and find that the optimal model is regression tree and neural network [6]. In some cases, compared with linear regression method, the return of investment strategy constructed by machine learning model is twice that of linear method. Other scholars used the support vector machine model to model the high-frequency limit order book data to predict the trend of the middle price of the stock in the limit order book, and found that the model has a good prediction effect on the short-term price trend [7].

By systematically testing the performance of various models, this paper improves the empirical research on machine learning models in the field of financial risk prediction, enriches the research results for the application of financial risk prediction and machine learning, and provides ideas for more in-depth research on financial risk prediction and the theory and application of machine learning models.

## 2 Financial Forecasting Model based on LSTM

### 2.1 LSTM Algorithm

Machine learning is outstanding in predicting multidimensional and heterogeneous data, which makes the application of machine learning algorithm in financial stock prediction become a research hotspot at this stage. When predicting non time series data, neural network ANN and other model algorithms are usually used for prediction [8]. However, RNN, LSTM, and CNN are commonly used to predict time series data.

RNN cannot learn the information that the connection steps are too far apart, which is a long-term dependency problem [9]. In the later stage, many studies improved and popularized it, so that it could be popularized and applied. In the standard RNNs, there is a module chain form of neural network repetition, and this part often has such a simple structure as a single tanh layer.

Similarly, as an improved RNN, LSTM also has this chain structure. However, the structure of the repeating module is interacted by four neural networks in a special way [10].



There are three types of gates in the repeating module structure of LSTM: forgetting the gate, the entrance gate, the exit gate. The well-designed gate structure can be used to delete or add information to cells. It is a method that allows selective transmission of information. It consists of a neural network layer and point-to-point function. The sigmoid layer produces a number between 0 and 1, which describes how much information each component can carry, and maps the memory attenuation factor. 0 means no quantity is allowed to pass, 1 means any quantity can pass. Three different types of gate can effectively protect and control its cell condition.

The function of the forgetting gate is to control the memory unit and determine the extent to which the state of the previous moment reaches the present moment according to the memory attenuation coefficient. Forgetting gate is the guarantee that LSTM can remember long-term memory. The detailed calculation formula is as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

The input gate determines how many inputs at the current time are reserved to the current status, and is mainly responsible for updating the cell status. Forget some information and remember some new information. The detailed calculation formula is as follows:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

After forgetting, entering and selecting, the information needs to be updated and saved to the next cell status. The specific operations are shown in formula (3):

$$C_t = f_t * C_{t-1} + i_t * C_t \quad (3)$$

The last step in structure design is to determine what value will come out? The value is filtered based on the state of the cell. First, an output gate is used to determine what part of the cell state is the production. The function is then used to process the cell state, multiplied by the output gate to determine the output portion of the information.

## 2.2 Construction of LSTM Prediction Model

In this paper, LSTM xgboost hybrid prediction model is designed by combining LSTM and xgboost. First, interpolate and sort the data set to avoid possible confusion and missing values in the original data table. Next, take the converted time information and stock price rise and fall of the data samples in the stock data list as the prediction data set, input the xgboost model, complete the training and save the training results. Then, input attribute values such as "open" and "close" in the prediction model for prediction, reconstruct the training data and time data of different characteristics, use the reconstructed data set as the test set for prediction after training, and set different parameters according to the actual prediction results. Finally, the model has the best training effect.

In this paper, the workflow of building LSTM xgboost model is as follows:

Obtain the historical attribute data of the stock index as the initial experimental data set of this experiment, and then deal with the missing values and sorting of the data set;

The time data and “date” attribute in the test set are decomposed into three attributes, referred to as time component;

The technology in the antenna package is used to build the neural network model. The project characteristics of stock prediction in the stock history data set are trained respectively, and the prediction model is established. In model LSTM\_Size = 128, unit = 256, batch processing, forgetting rate is set to 0.4.

Recombine the prediction results with the “time characteristics” to build a “stock price” test set;

The xgboost algorithm is implemented by using the xgbregressor method in the sklearn package to train the attributes of the stock history dataset, such as the adjustment threshold and “time component”, and build a prediction model;

The stock price rise and fall prediction model then predicts the reformulated “text” data samples after LSTM prediction. Compared with other models, the algorithm uses the original default parameters.

Finally, by comparing the actual value with the predicted value, the performance of the model designed in this paper in stock price prediction is evaluated.

### 3 Simulation Experiment of Financial Forecasting Model based on LSTM

#### 3.1 Simulation Experiment Platform

The experimental environment parameters of this chapter are as follows: the computer processor is Intel i7, the CPU is dual core and four threads, the windows10 operating system is adopted, and pycharm is used as the experimental simulation platform. The compiling language of this experiment is python.

#### 3.2 Experimental Data

The financial time series used in the experiment are two financial data sets, including Shanghai Stock Exchange Index and Shenzhen index. The data index time series is from January 2016 to February 2020.

##### (1) Data characteristics and labels

The dataset uses the same feature set, which contains 13 features: opening price (open), closing price (close), highest price (HI), lowest price (10W), trading volume (volume), p\_change, 5-day average (Ma5), 10 day average (mal0), 20 day average (mA20), 5-day variation average ( $v\setminus u$  Ma5), 10 day variation average ( $v\setminus u$  mal0), 20 day variation average ( $v\setminus u$  mA20), turnover rate (turnover). V\_Ma variation average is the average calculated by adding the opening price, closing price, highest price and lowest price of each trading day and dividing by 4.

##### (2) (2) Data preprocessing

First, take the 20th day as the sliding window to change the two-dimensional data set into the three-dimensional data set. According to the time series, the data of every 20 trading days corresponds to a tag, and the tag value is the tag of the 20th day. Data sets are generated circularly from the first trading day.

Secondly, each data is normalized so that the value range of each feature is between 0 and L.

Finally, the three-dimensional and normalized data sets are divided into training set, verification set and test set according to the ratio of 6:2:2.

### 3.3 Evaluation Function

The following five indicators are used to evaluate the classification performance of the classification algorithm: accuracy, precision, recall, FL and AUC.

AUC is the area below the ROC curve. The curve is the characteristic working curve of the subject. The horizontal axis of the curve is the vertical axis, equal to the retraction rate. In the binary classification problem, ROC curve generates different TPR and fpr values from the prediction results according to different thresholds, and connects TPR and fpr into a curve. The AUC value range is between 0 and 1. The closer the AUC value is to 1, the better the classification effect of the algorithm is. If AUC value is lower than 0.5, random guess is better than this algorithm.

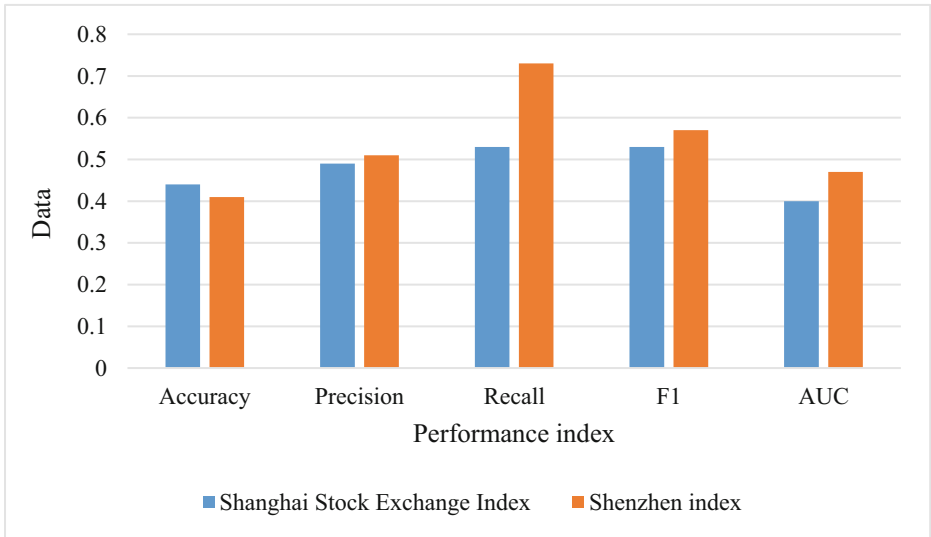
## 4 Analysis of Simulation Experiment Results

### 4.1 Analysis Results of Traditional LSTM Model

As shown in Fig. 1 and table 1, the accuracy, precision, recall, F1 and AUC of Shanghai composite index are 0.44, 0.49, 0.53, 0.53 and 0.40 respectively; The accuracy of Shenzhen index is 0.41, precision is 0.51, recall is 0.73, F1 value is 0.57 and AUC is 0.47 From the above experimental results, it can be shown that the prediction performance of the traditional LSTM is equivalent to that of a weak learner.

**Table1.** Predicted performance table of LSTM

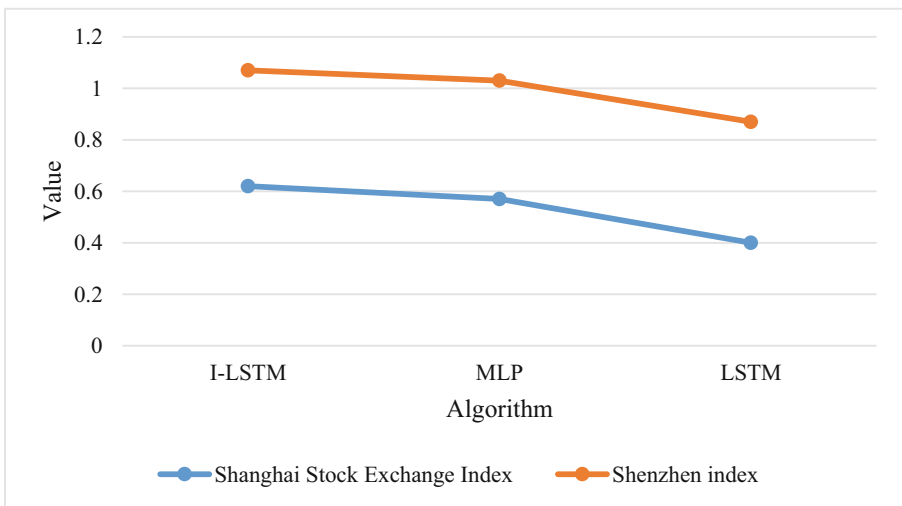
	Accuracy	Precision	Recall	F1	AUC
Shanghai Stock Exchange Index	0.44	0.49	0.53	0.53	0.40
Shenzhen index	0.41	0.51	0.73	0.57	0.47



**Fig. 1.** Statistics of traditional LSTM prediction results

#### 4.2 Comparison between Improved LSTM Model and other Algorithms

As shown in Fig. 2, this is a comparison between the accuracy of the improved LSTM algorithm and the integrated MLP and traditional LSTM algorithms. The prediction effect of the improved LSTM algorithm is better than that of the integrated MLP and traditional LSTM algorithm. The results show that the improved LSTM algorithm is superior to the integrated MLP algorithm and the traditional LSTM algorithm.



**Fig. 2.** Comparison of improved LSTM algorithm with other algorithms

## 5 Conclusions

Using machine learning algorithm to predict financial stocks has become a hot spot and trend of existing research, and relevant experiments have confirmed that investors benefit from the prediction results of this method. The multidimensional and heterogeneous input characteristics comprehensively analyze the stock from the technical and basic aspects. The complex and diverse input features of stock forecasting make it necessary to consider the combination of various types of input features in feature selection. This paper uses the combination forecasting method of LSTM and xgboost to construct a financial risk forecasting model based on machine learning. The experimental results show that the performance of the financial risk forecasting model constructed in this paper is higher than that of the traditional machine learning model. However, there are many factors that affect the price fluctuation of the stock market, including investor sentiment, major natural disasters, economic crisis and other uncertainties and force majeure factors. Therefore, if you want to predict the stock price fluctuation more accurately, you need to fully consider the main influencing factors in all aspects. These issues need to be further studied.

## References

1. Shah, A.A., Karhade, A.V., Bono, C.M., et al.: Development of a machine learning algorithm for prediction of failure of non-operative management in spinal epidural abscess. *Spine J.* **19**(10) (2019)
2. Sidey-Gibbons, C., Pfob, A., Asaad, M., et al.: Development of machine learning algorithms for the prediction of financial toxicity in localized breast cancer following surgical treatment. *JCO Clinical Cancer Inform.* **5**, 338–347 (2021)
3. Damiati, S.A., Martini, L.G., Smith, N.W., et al.: Application of machine learning in prediction of hydrotrope-enhanced solubilisation of indomethacin. *Int. J. Pharm.* **530**(1–2), 99–106 (2017)
4. Tadepalli, S.K., Lakshmi, P.V.: application of machine learning and artificial intelligence techniques for IVF analysis and prediction. *Int. J. Big Data Anal. Healthc.* **4**(2), 21–33 (2019)
5. Karimanzira, D., Rauschenbach, T.: Performance prediction of a reverse osmosis desalination system using machine learning. *J. Geosci. Environ. Prot.* **9**(7), 16 (2021)
6. Lopes, R.G., Ladeira, M., Carvalho, R.N.: Use of machine learning techniques in the prediction of credit recovery. *Adv. Sci. Technol. Eng. Syst. J.* **2**(3), 1432–1442 (2017)
7. Oosterhoff, J.H.F., Karhade, A.V., Oberai, T., et al.: Prediction of postoperative delirium in geriatric hip fracture patients: a clinical prediction model using machine learning algorithms. *Geriatr. Orthop. Surg. Rehabil.* **12**(2), 435–446 (2021)
8. Le, T., Bondarev, A.V., Kozlova, E.V.: Application of machine learning algorithms in predicting pyrolytic analysis result. *IOP Conf. Ser.: Earth Environ. Sci.* **931**(1), 012013 (10pp) (2021)
9. Masud, F.A., Rejaul, M., Royel, I., et al.: Smart risk prediction tools of appendicitis patients: a machine learning approach. *Biointerface Res. Appl. Chem.* **11**(1), 7804–7813 (2021)
10. Balasubramanian, K., Ananthamoorthy, N.P., Ramya, K.: Prediction of neuro-degenerative disorders using sunflower optimisation algorithm and Kernel extreme learning machine: a case-study with Parkinson's and Alzheimer's disease. In: *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 236, no. 3, pp. 438–453 (2022)



# Credit Risk Assessment Model of Supply Chain Finance Under the Background of Big Data

Jun Ma<sup>(✉)</sup>

Xi'an Eurasia University, Xi'an 710065, Shaanxi, China  
majun80031@163.com

**Abstract.** Financial institutions must clarify the logical relationship of supply chain finance operations, objectively evaluate the qualifications of credit entities, and strengthen credit risk early warning and prevention in order to survive in the fierce competition of supply chain finance. The purpose of this paper is to study the credit risk assessment model of supply chain finance based on the background of big data. Combined with the characteristics of the automobile industry and supply chain business, the selection principle and selection process of indicators are introduced, and the evaluation system of the financial credit risk of the automobile industry supply chain is constructed through the combination of qualitative and quantitative. Then, the model analysis of the credit risk of supply chain finance is carried out, and an empirical study is carried out on the supply chain finance of the automobile industry through four parts: data source and description, sample selection, principal component analysis method and logistic model regression analysis. Finally, the comprehensive accuracy rate of model prediction is 92.5%, and it is concluded that the credit risk assessment system constructed in this paper and the robust model adopted in this paper have high accuracy and applicability in assessing credit risk.

**Keywords:** Big Data Technology · Financial Credit · Risk Assessment · Assessment Model

## 1 Introduction

At present, with the wide application of my country's supply chain model and the innovation of the supply chain financial order, a series of problems have arisen, which are directly reflected in the entire social life [1]. Credit risk assessment is the premise of financing activities. Compared with traditional financing models, supply chain finance makes traditional risk credit assessment methods no longer applicable due to numerous participants, complex business models and different application industries [2]. So how to formulate an effective credit risk assessment method under the supply chain financial model according to the industry characteristics of new energy vehicles is the key difficulty in expanding market financing. If it can be solved, it will become a powerful engine to promote the development of the new energy industry [3, 4].

Credit risk assessment is a key step before financing activities. The parallel social spider algorithm proposed by Shukla UP designs a credit card lending model. Appropriate changes have been made to coding concepts and matching processes to effectively address credit rating issues. Tests are performed against many available standard credit card data, such as German, Australian, and Japanese credit card data. This algorithm performs better than the algorithms used by K-mean, Parallel Real Genetic Algorithm, and Parallel Particle Swarm Optimization (PPSO). The profile metrics obtained from different specific algorithms on the German dataset are: 0.56% for K-mean and 0.86% for real-coded genetic algorithms [5]. Kalimashi helps users understand the operational aspects of financial transaction processes and review processes in order to properly assess a bank's credit risk. Explore scientific considerations related to the importance and development of financial management information for local and international financial institutions. Describe the relationship between reviewing financial information and managing customer credit risk associated with bank spending [6]. Therefore, it is necessary to investigate the credit risk of the financial supply chain in order to promote the healthy improvement of the supply chain in my country [7].

The main objective of this paper is the financial risk of the supply chain industry chain. After introducing the concept and method of traditional supply chain management, the concept, service structure and method of financial supply chain supply are deeply studied. This paper analyzes the unique credit risk characteristics of the automotive supply chain in detail. According to the characteristics of the supply chain and the cost of automation, develop the financial risk index system of automated chain supply supply and conduct capability analysis.

## **2 Research on the Credit Risk Assessment Model of Supply Chain Finance under the Background of Big Data**

### **2.1 Supply Chain Finance**

At present, there is no precise definition of supply chain spending, and there are certain differences in our understanding of supply chain spending. Supply chain budgets are very different from traditional spending. Its performance depends on the financial needs of supply chain participants [8, 9]. On the one hand, due to the shortage of fixed assets and deficits in spending plans, it is difficult for supply chain participants to obtain funds from financial institutions through traditional channels, and expenditure costs are uneven; institutional benefits are limited. Compared with other financial services, my country's financial institutions are more willing to participate in the main credit, as well as risk prevention methods and technologies [10, 11].

### **2.2 Logistic Model**

Logistic model is a generalized linear regression analysis model, which is mainly used to predict the influence of various factors on events and the probability of occurrence of events. It is widely used in various fields [12]. The dependent variable of the logistic model usually needs to obey the binomial distribution, so instead of using the least

squares method, the maximum likelihood method is used for regression. The value of the model dependent variable can only take “0” or “1”. When the value is “1”, it means that the event occurs, and when the value is “0”, it means that the event does not occur [13, 14].

### 2.3 Build an Indicator System

(1) The principle of comprehensiveness

There is more than one factor that affects the credit risk of supply chain finance, so we should pay attention to the comprehensiveness of the indicators when selecting the corresponding evaluation indicators [15].

(2) Scientific principles

There are both quantitative and qualitative factors that affect the credit risk of supply chain finance. In addition, when considering the scientific nature of the indicators, attention should also be paid to the relationship between the indicators and the practical significance of each indicator, which are contradictory, repetitive or unrealistic. Significance indicators should not be selected [16].

(3) The principle of operability

In the face of a large number of indicators, necessary trade-offs should be made, and those indicators that are easy to collect and quantify should be selected as much as possible. For those indicators that are repetitive, meaningless, and difficult to collect and quantify, without affecting the comprehensiveness and scientificity of the evaluation system. Can be appropriately discarded [17].

### 2.4 Constructing Risk Indicators

**Macroeconomic environment:** No matter what nature or type of enterprise, the specific production and operation are inseparable from the influence of the specific economic environment in which it is located. Therefore, changes in the macroeconomic environment will directly bring about the development of financing enterprises. Influence, further affects the repayment ability of financing enterprises [18].

**Industry development potential:** The investigation of industry development potential is mainly carried out from two aspects: the life cycle stage of the industry and the intensity of market competition. Through the inspection of these two indicators, we can judge the future development trend of the industry and the management capabilities of enterprises in the industry.

**Operational Capability:** Operational capability is a reflection of an enterprise’s operating level and capital utilization efficiency, and reveals how the enterprise manages various assets it owns.

**Profitability:** Profitability reveals the efficiency with which a business creates value. If the company has strong profitability, it can achieve stable growth in operating income, be able to repay bank loans on time, and have low credit risk. The profitability of a company can be measured by indicators such as operating profit rate and return on equity.

**Solvency:** Solvency reveals the ability of an enterprise to use its assets to repay debts owed, and is one of the key indicators for judging the financial status of an enterprise.



Pledge status: Pledge status is an important concern for banks in the credit review process, because commercial banks first evaluate the value of the pledge before granting credit to the financing enterprise, and determine the loan amount based on the final result, once the trustee defaults, the value of the pledge is the guarantee for the bank to make up its own losses.

### 3 Investigation and Research on the Credit Risk Assessment Model of Supply Chain Finance under the Background of Big Data

#### 3.1 Sample Collection

This paper selects the data of 50 auto dealers from 2019 to 2021 as the research sample data, a total of 100 samples, and the data comes from a domestic bank’s public credit database. The selected car dealership brands include luxury brands such as Mercedes-Benz, Audi and BMW, as well as mid-end brands such as FAW-Volkswagen. The regional distribution of the sample data is relatively scattered, with a total of 15 cities, including municipalities, provincial capitals and other cities, fully considering the diversity of the samples. Since most of the manufacturers of automobile brands selected in the sample are Sino-foreign joint ventures, traditional financial indicators such as macroeconomic environment and profitability within the sample range are selected. The guarantee effectiveness index is defined by whether a tripartite agreement is signed, and the industry status index is The annual sales growth rate is used as a comprehensive indicator to measure the core vehicle manufacturers, and the data sources are corporate announcements and internal bank databases.

#### 3.2 Regression Analysis Based on Logistic Model

This paper finally adopts the Logistic model to measure the risk level of auto dealers. Suppose  $P(Z = 1|K) = P_i$  is the conditional probability of the dealer default, where  $Z$  represents the default risk of the financing company,  $K = 1$  represents the company defaults, and  $K = 0$  represents the company repays the principal and interest on time.  $K = (K_1, K_2, \dots, K_p)^T$  is a  $P$ -dimensional random variable, and  $Y$  represents the influencing factors of corporate default risk.  $\beta = (\beta_1, \beta_2, \dots, \beta_k)^T$  is the Logistic coefficient of the independent variable  $Y$ ,  $\beta_0$  is a constant term, and the model is:

$$P_i = \frac{1}{1 + e^{-z_i}} \tag{1}$$

$$z_i = \beta_0 + \beta_1\beta_{1i} + \beta_2\beta_{2i} + \dots + \beta_m\beta_{mi} = \beta_0 + \sum_{k=1}^m \beta_k K_{ki} \tag{2}$$

This function is a (0, 1) increasing function. The closer the  $P$  value is to 1, the lower the corporate credit level and the higher the probability of default. After principal component analysis, 8 new variables were selected as independent variables for regression analysis ( $K_1, K_2, K_3, K_4, K_5, K_6, K_7, K_8$ ), and Logistic binomial regression was performed by SPSS22 software. The five independent variables  $K_1, K_2, K_4, K_6$  and  $K_7$  were all more

than 95% significant and were retained in the equation. Therefore, the logistic regression equation is:

$$P_i = \frac{1}{1 + e^{-(2.463 - 0.254K_1 + 1.56K_2 - 1.836K_4 - 0.254K_6 - 1.356K_7)}} \quad (3)$$

In the equation, P represents the default probability of auto dealers. If the value of P is closer to 1, it means that the company is more likely to fail to repay the principal and interest; the closer the value of P is to 0, the lower the possibility of the company defaulting. According to the absolute value of the regression coefficient, the degree of influence of this factor on the default risk probability of auto dealers can be judged. The larger the absolute value, the deeper the degree of influence. In addition, the positive or negative sign of the coefficient represents the correlation between its default influencing factors and the default probability, and the positive sign represents a positive correlation with the P value. The larger the factor, the greater the dealer default probability.

## 4 Analysis and Research on the Credit Risk Assessment Model of Supply Chain Finance under the Background of Big Data

### 4.1 Omnibus Test of Model Coefficients

The Omnibus test of the model coefficients is shown in Table 1, and the sig. Are all less than 0.02, as shown in Fig. 1. It shows that the results of the Logistic regression equation are significant, and the factors such as macroeconomic environment factors, industry development potential factors, operating factors, profitability factors, debt repayment factors and pledge conditions extracted by principal component analysis in this paper are all consistent with the research object of this paper, namely online There is a significant relationship between the credit risk assessment of supply chain finance, so the model measurement used in this paper has practical operational significance.

**Table 1.** Omnibus test for model coefficients

step	Bangla	degrees of freedom	salience
step	63.5	8	0.01
piece	63.5	10	0.001
Model	63.5	9	0.004

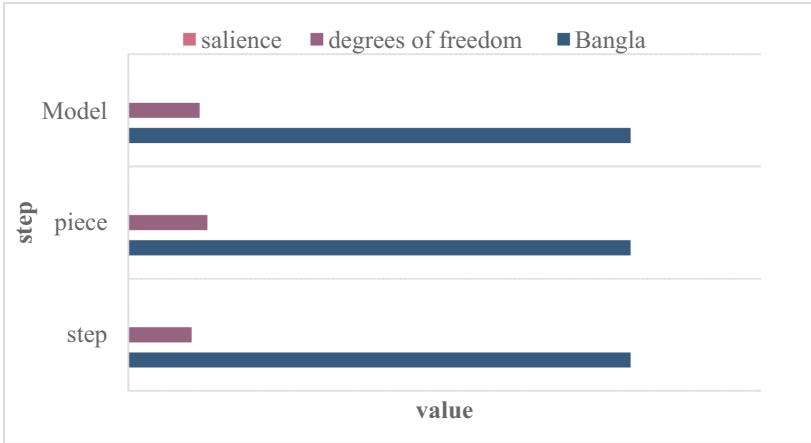


Fig. 1. Test results

### 4.2 Empirical Results

Logistic model test will have two results. One result is that the model judges companies that do not have default risk as having default risk; and the other is that companies with default risk that were originally set as risk groups are judged to have no default risk. The model makes predictions among 50 companies in the auto industry. Among the 40 companies that do not have default risks, 38 predictions are correct, with an accuracy rate of 95%. Among the 10 companies with default risks, 9 predictions are correct. The correct rate is 90%, and the comprehensive correct rate is 92.5%. It can be seen that the model has a good effect on corporate default risk prediction, as shown in Fig. 2.

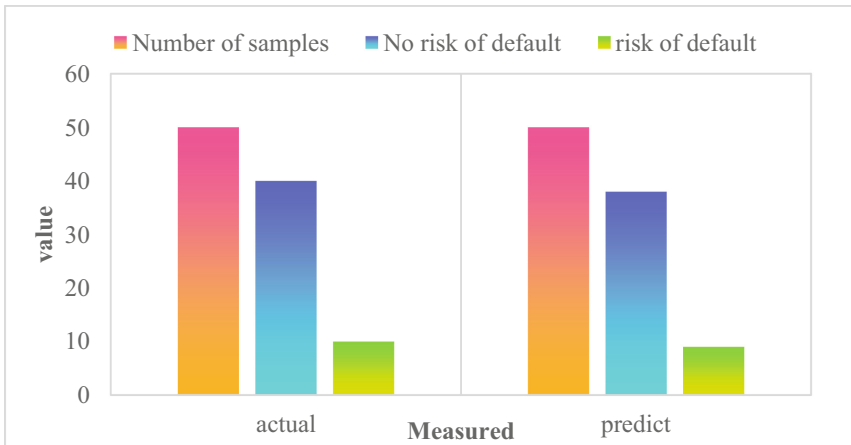


Fig. 2. Empirical Results

This paper uses a logistic model to assess the credit risk of automotive supply chain financing. It can be seen that the model in this paper can effectively evaluate the credit risk indicators of the automotive supply chain, such as the solvency and profitability of financing companies, and has significant corresponding effects. Facts have proved that the index system and robust model measurement method constructed in this paper can be widely used in risk assessment.

## 5 Conclusions

This paper conducts an in-depth study on the supply chain financial business of my country's automobile industry, and draws some relatively novel conclusions. A comprehensive measurement of risk can be used as a follow-up research direction. In my country, the number of small and medium-sized enterprises is huge and the degree of information disclosure is low, and it is difficult to collect data. Therefore, this paper only selects a domestic bank to conduct research on supply chain financial credit risk of small and medium-sized enterprises in the public credit database, and fails to cover all small and medium-sized enterprises. In the study, the sample size can be expanded to obtain more accurate results. In terms of industry selection, not all industries are involved in this financial business, and only well-developed industries can be selected for research, which is underrepresented. In the future research, the supply chain financial credit risk of different industries can be compared and studied to obtain more accurate and comprehensive research conclusions.

## References

1. Chakrabarti, D., Sethi, P., Bhattacharjee, S.: Directed credit, financial development and financial structure: theory and evidence. *Appl. Econ.* **51**(16–18), 1711–1729 (2019)
2. Kibicho, N.K., Mungai, J.: Mobile banking adoption and financial credit accessibility in Wote Sub-County, Makueni County, Kenya. *Int. J. Curr. Aspects*, **3**(IV), 65–79 (2019)
3. Смирнов, Е., Smirnov, E.: Changes in Financial Legislation on Credit. *Auditor* **5**(8), 3–8 (2019)
4. Sadchenko, O.V., Nikola, S.O.: Transformation in economy: financial and credit system and marketing approaches. *Econ. Innov.* **21**(4(73)), 150–160 (2019)
5. Shukla, U.P., Nanda, S.J.: Designing of a risk assessment model for issuing credit card using parallel social spider algorithm. *Appl. Artif. Intell.* **33**(1–4), 191–207 (2019)
6. Kalimashi, A., Ahmeti, Y., Ahmeti, A.: The impact of audited financial statements on credit risk assessment in the banking system of kosovo (Period 2014–2018). *Baltic J. Real Estate Econ. Constr. Manage.* **8**(1), 119–128 (2020)
7. Lee, K.K., Schantl, S.F.: Financial reporting and credit ratings: on the effects of competition in the rating industry and rating agencies' gatekeeper role. *J. Account. Res.* **57**(2), 545–600 (2019)
8. Kwilinski, A., Shteingauz, D., Maslov, V.: Financial and credit instruments for ensuring effective functioning of the residential real estate market. *Financ. Credit Act. Probl. Theory Pract.*, **3**(34), 131–138 (2020)
9. Okafor, L.E.: Bank credit, public financial incentives, tax financial incentives and export performance during the global financial crisis: a review. *Shanlax Int. J. Econ.* **8**(2), 1–4 (2020)

10. Prof, M., Abdul, A., Shwaish, A., et al.: The role of some financial innovation instruments in hedging credit risks an empirical study for a sample of international commercial banks. *J. Adm. Educ. Manage. (ALIGNMENT)* **15**(45), 137–158 (2020)
11. Rozhko, O.D., Marenych, T.H., Onegina, V.M., et al.: Bank credit in financial provision of agricultural enterprises. *Financ. Credit Act. Probl. Theory Pract.* **4**(31), 41–51 (2019)
12. Steblyi, H.Y., Tuvakova, N.V.: Keynesianism and monetarism as theoretical and methodological constructions of financial and credit system. *Financ. Credit Act. Probl. Theory Pract.* **2**(29), 412–418 (2019)
13. Clark, E., Mitra, S., Jokung, O.: Post global financial crisis modelling: credit risk for firms that are too big to fail. *Int. J. Financ. Markets Deriv.* **7**(1), 15–39 (2019)
14. Cherkashyna, E.: The effectiveness of state regulation and supervision of non-banking financial and credit institutions. *Mod. Econ.* **11**(1), 154–159 (2018)
15. Zvorych, R.Y.: Financial and credit imperatives of alter globalization the emerging markets. *Financ. Credit Act. Probl. Theory Pract.* **1**(24), 360–367 (2018)
16. Agasha, E., Monamesti, G., Feela, T.: Loan portfolio quality of microfinance institutions in uganda: a qualitative assessment. *J. Financ. Risk Manage.* **09**(2), 155–177 (2020)
17. Soui, M., Gasmi, I., Smiti, S., et al.: Rule-based credit risk assessment model using multi-objective evolutionary algorithms. *Expert Syst. Appl.* **126**(JUL.), 144–157 (2019)
18. Nasir, Z., Ahmed, Z., Lal, C.: Assessment of default risk factors in the disbursement of home loans. *Int. J. Adv. Trends Comput. Sci. Eng.* **10**(3), 2408–2420 (2021)



# Large Screen for 3D Data Visualization Based on RFG-SVM Algorithm

Guojuan Wang<sup>1</sup>, Lijun Wang<sup>1</sup>, Heqi Zhao<sup>2</sup>, Qi Mu<sup>3</sup>(✉), and Xiangchun Ji<sup>1</sup>

<sup>1</sup> Beijing Guodiantong Network Technology Co., Ltd., Beijing 100070, China

<sup>2</sup> Beijing China-Power Information Technology Co., Ltd., Beijing 100085, China

<sup>3</sup> Beijing Creative Cloud Technology Co., Ltd., Beijing 100022, China

muqi8520@163.com

**Abstract.** With the rapid development of big data today, the huge amount of data has made many people see its value. In order to display data more intuitively and friendly, data visualization technology is constantly developing. The purpose of this paper is to study the design and implementation of large-screen 3D data visualization based on RFG-SVM algorithm, and propose a feature-weighted support vector machine RFG-SVM algorithm, which is improved and modified on the traditional support vector machine algorithm. Select Gaussian weighted kernel function, appropriate penalty coefficient and kernel function parameters to construct and solve the optimization problem. Memory usage optimization is achieved by calling the cleanup memory API in ECharts - the dispose() method. After tuning, the memory occupied by the browser is basically kept at 300MB, which achieves the effect of reducing the memory usage of the browser and reducing the freeze. Real-time large-screen performance optimization is achieved through throttling design. After optimization, the average response time of real-time views is 274ms, and the response time of views is significantly improved.

**Keywords:** RFG-SVM Algorithm · 3D Data · Large-screen Visualization · Design and Implementation

## 1 Introduction

With the gradual expansion of the application scope of data visualization technology, its influence is also increasing year by year, various institutions including universities, governments, enterprises, etc. are conducting in-depth research on visualization technology, and have established research groups for different fields. Realize the rapid development of visualization. In the research of data visualization, people have not only paid attention to how to visualize complex multi-dimensional data, but have begun to pay more attention to how visualization can better demonstrate the real-time and interactive analysis capabilities of data [1]. As one of the intuitive display forms of data visualization technology, the 3D data visualization large screen can visualize complex multi-source data in a limited window, and present the hidden business relationships behind large-scale complex data in real time. It is widely used in finance, It can be used in various fields such as medical treatment and urban transportation [2].

Large visual screens are widely used in many fields due to their superior performance. Al-Kharusi I analyzes the spatial data coverage and data reuse properties of higher-order, Hilbert, and Morton series and their impact on molecular dynamics simulation performance. A simple caching model is presented, and the results are consistent with the timing results obtained for particle shock calculations on NVidia GeForce GTX960 and Tesla P100 GPUs. Further analysis of the observed memory usage can further explain the execution behavior of different rules based on the number of cache hits and memory transactions. To our knowledge, this is the first study of memory analysis and data coverage problems for molecular dynamics simulations of Lennard-Jones pumps on NVidia's Maxwell and Tesla architectures [3]. Simple TL provides a comprehensive workflow for manipulating and viewing 3D data, with basic and advanced options for creating images, videos and interactive 3D-PDFs from volume settings and surface mesh adjustments. The importance of visualization for quantitative analysis of invertebrate morphology from 3D data is discussed, and example designs are provided to demonstrate different options for creating 3D designs for publication. As more biology journals adopt 3D-PDF as a standard option, research on microscopic invertebrates and other organisms can be presented in high-resolution 3D images, enhancing scientific communication [4]. Angela used drone imagery to assess the impact of the periodic appearance of cicadas in mixed forests. The aim was to assess the potential of this technology as a forest health monitoring tool. We classified the cicada results using two possible classification methods, one using the highest spectrum (MLC 1) extracted from leaf images, while in the second method we added the leaf canopy model from leaf height (CHM) The data, Digital Terrain Model (DTM) is extracted from Digital Terrain Model (DSM) - Information During Classification (MLC 2) [5]. Therefore, the theoretical significance and display significance of the design based on the RFG-SVM algorithm in this paper are obvious.

The main research content of this paper is to design and implement a large screen for 3D data visualization based on the RFG-SVM algorithm. It defines the construction process of large screen visualization aiming at the application of large screen functions, and realizes the creation of 3D graphics from reading a variety of data to 3D graphics. Large-screen manufacturing, large-screen operation applications and other complete sets of services. It solves the problems of complex use, long learning curve, and the need for professionals. A variety of simple and easy-to-understand interactive methods designed in this system enable users to use it immediately, reducing the user learning curve and use cost.

## **2 Research on Large Screen of 3D Data Visualization Based on RFG-SVM Algorithm**

### **2.1 Support Vector Machine (SVM)**

As a machine learning method, support vector machine is helpful to learn and understand the principle of support vector machine better. Learning can summarize and analyze the historical information of existing things, so as to discover its inherent laws, and then predict future behaviors. Machine learning refers to the method of artificial intelligence,

by establishing a specific mathematical model, and then analyzing, training and learning existing historical data and results, and searching for a specific internal relationship between these data and results. After the rules are established, certain regression predictions or classification judgments can be made for future data and results [6, 7].

## 2.2 Feature Weighted Support Vector Machine (RFG-SVM)

Let the known training sample set  $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)\}$ ,  $\chi_i = (\chi_i^1, \chi_i^2, \dots, \chi_i^n)$  is an  $n$ -dimensional vector, and let the feature vector weights be  $\{w_1, w_2, \dots, w_n\}$ , the eigenvector weight is a constant calculated according to the algorithm introduced in the previous section. For nonlinear classification problems [8, 9]. It is defined as follows.

Then the feature weighted kernel function is:

$$K_p(\chi_i, \chi_j) = K(P\chi_i, P\chi_j) \quad (1)$$

Apply feature weighted kernel function Gaussian feature weighted kernel function to support vector machine:

$$K_p(\chi_i, \chi_j) = \exp(-\gamma \|P\chi_i - P\chi_j\|^2) \quad (2)$$

## 2.3 Large-Screen Visualization

### (1) Visualization chart

The built-in graphics of the large-screen designer are the result of independent research and development, which can avoid the inconvenience of intellectual property, services and documents caused by the use of third-party plug-ins. Users can select charts according to their data visualization needs or use the chart type automatic prompt function to quickly create charts [10, 11].

### (2) Data warning

Users can define special effects for conditional display on the chart. When the graphic data meets certain conditions, a series of graphic animations or flashing effects can be displayed. Combined with monitoring and refresh, the application of the large-screen front panel can alarm the data on the same panel, and the alarm effect can also be set according to the alarm situation [12, 13].

### (3) Automatic rotation

The indicator supports switching multiple chart styles at the same location, and the data source of each chart can also be different. It also supports auto-rotation of multiple chart and table elements using tab block elements, and even auto-rotation of multiple templates. The switching time is set by the user, and the user can also change the automatic switching to manual switching [14, 15]. Additionally, maps such as scene maps, logic trees, and carousel objects have their own carousel capabilities. By moving parameters, an image can be rotated while connecting local elements to display multiple scenes on a large screen model. And data size. Autorotate is not limited to charts. For some projects, it is normal to use automatic scrolling or display scrolling, such as automatic scrolling of large lists (ie, marquee effect), automatic scrolling of KPI cards, etc. [16].



### 3 Design and Development of Large-screen 3D Data Visualization Based on RFG-SVM Algorithm

#### 3.1 Development Environment

(1) Software environment

Mac OS: v10.13.5

Webstorm: v2020

Google Chrome: v87

MongoDB: v6.14.10

Vue: v2.9.3

@vue/cli:v4.5.10

Node.js: v14.15.4

(2) Hardware environment

Processor: 2.5 GHz Intel Core i7

Memory: 16 GB 1600 MHz DDR3

Hard Disk: 500GB

Display: 15.4 inches, 2880 x 1800 resolution

#### 3.2 System Architecture Design

(1) System physical architecture design

The C/S (Client/Server, client server) mode is superior to the B/S mode in terms of graphics performance and running speed, but considering that the main usage scenario of this platform is that users can The management, construction and viewing of large-screen applications can be conveniently carried out through the browser through the Internet in all kinds of industry occasions. At the same time, the platform needs to optimize and expand platform functions in a timely and convenient manner according to user feedback, and users can update synchronously without re-downloading and installing specific client applications after the function is upgraded.

(2) System logical architecture design

From top to bottom, the B/S architecture pattern is the interface layer, the business logic layer and the data access layer.

**Presentation layer:** It mainly realizes the recognition of user operation instructions by the system, the transmission of background requests, and the display of output results. For web applications, the form is a browser web page.

**Business logic layer:** It mainly accepts user instructions in the presentation layer, and converts them into data operations in the data access layer. Finally, the information returned by the data access layer is processed and processed according to the platform business rules and returned to the presentation layer, which plays a linking role.

**Data access layer:** It mainly realizes receiving the database operation instructions transmitted from the business logic layer, executing the corresponding library operation and feeding back the encapsulated operation result information. The direct interaction with the underlying database determines the system performance to a certain extent.

### 3.3 Design of Core Functional Modules

#### (1) Multi-scene large-screen management module

The multi-scene large-screen management module is divided into four main functions: scene creation, large-screen editing, large-screen preview and large-screen publishing. A brand-new editable large-screen is created by the user-defined business scene field. The screen project provides four quick function guides for editing, deleting, previewing and publishing. Users can view all created scene field large screens in the entire large screen management subsystem, and directly perform corresponding editing operations on the selected large screen according to the shortcut function guidance of each large screen.

#### (2) Public visualization chart library

Important materials for creating large data visualizations for data analysis and tracking are charts, maps, graphs, and other visual elements. Due to the complexity and complexity of the data source, using different visualization charts for the same data will show different information behind the data, and the types of data structures supported by different views are also different. In order to fully mine data information in different business scenarios in a limited view, the platform provides a rich type of public chart component library for building large data visualization screens on the basis of componentized design.

#### (3) Graphical drag-and-drop editing module

Based on the rich and diverse public visualization chart library, the platform provides users with a list of visual views that can be dragged and dropped. The user selects a view and places it on the large-screen canvas by dragging and dropping it with the mouse, and it can be freely dragged and placed on the canvas. Determine the specific position of the visual view in the large screen and change the view size by free stretching, and you can right-click to choose to configure the cascading relationship between the views. In addition to the functional components developed by the platform itself, the specific implementation of the above functions also needs to be improved by external plug-ins.

## 4 Optimization Test of Large-screen 3D Data Visualization Based on RFG-SVM Algorithm

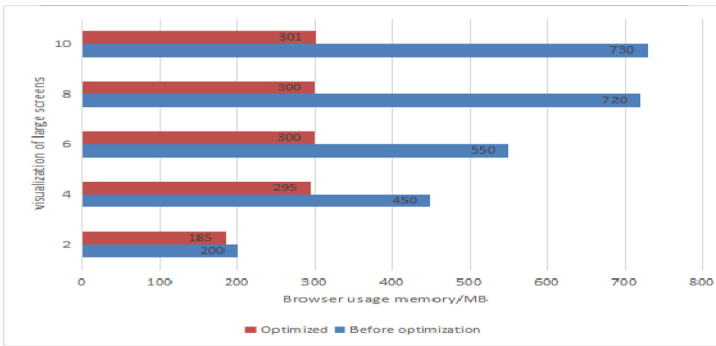
### 4.1 Occupied Memory Optimization

As the number of large screens edited by users increases, even if the large visual screen is closed immediately after editing, the browser memory continues to increase until the browser memory reaches the critical value of garbage collection. After analysis, it is found that the system can improve the rendering speed and ensure that the editing can be returned at any time after the preview. The system has been adjusted in two aspects. One is to set the browser to save only the information of the three newly opened large-screen visualizations, which satisfies the first-in-first-out principle of the caching algorithm. The second is to call the cleanup memory API in ECharts - the dispose() method. When a new visual large screen is opened, the Dom element of the visual view in the previous large screen is automatically cleaned and released. In order to reduce the memory usage

of the browser and reduce the effect of freezing. After the relevant optimization of the system, and the test results obtained are shown in Table 1:

**Table 1.** Performance comparison before and after tuning

Visualization of large screens	2	4	6	8	10
Before optimization	200	450	550	720	730
Optimized	185	295	300	300	301

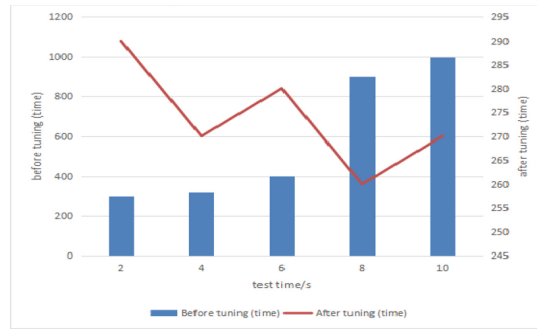


**Fig. 1.** Comparison before and after tuning

It can be found from Fig. 1 that the memory occupied by the browser has been improved to a certain extent when the first three large-scale visualization screens are opened. After more than three large screens, since the system releases the large screens previously cached, the memory occupied by the browser is basically kept at 300MB, so the memory consumption is basically stable. It can be seen that the optimization scheme implemented in this system reduces the impact of the increase in the number of large screens on the memory consumption of the browser, and improves the performance of the system.

### 4.2 Real-Time Large Screen Performance Optimization

This paper has carried out the throttling design. Throttling refers to setting a time period in advance. When the time of the action to be executed is greater than or equal to this period, the command is executed. That is to say, within the 3s interval set by the system, if the real-time data table is updated multiple times, only the last operation will be published to update the visual view. In this paper, the performance of the optimized real-time large screen is tested again, and the test results are shown in Fig. 2:



**Fig. 2.** Latency comparison before and after tuning

It can be seen from the figure that when the system is designed for throttling, the response time of the view is significantly improved. The average response time for the live view is both 274ms. According to the principle of front-end page response time, when users visit a page, if they get a response within 3 s, they will feel that the system responds quickly. It achieves the purpose of increasing the robustness of the real-time large screen and improving the user experience.

It can be seen from the figure that when the system is designed for throttling, the response time of the view is significantly improved. The average response time for the live view is both 274ms. According to the principle of front-end page response time, when users visit a page, if they get a response within 3 s, they will feel that the system responds quickly. It achieves the purpose of increasing the robustness of the real-time large screen and improving the user experience.

## 5 Conclusions

This paper optimizes the traditional support vector machine, proposes a feature weighted support vector machine (RFG-SVM) algorithm, and applies it to the design and implementation of the 3D big data visualization platform. The big data visualization platform can use rich functions and capabilities to create a visual index interface and analysis process, shorten the use and service cycle, and reduce enterprise costs. Through the powerful display of big data, use the Internet to conduct intelligent analysis and real-time monitoring of data, and effectively interpret the display results in front of you. Lightweight solution for flexible big data visualization, data visualization, and data fusion to meet needs easily and efficiently.

With the advancement in networking and multimedia technologies enables the distribution and sharing of multimedia content widely. In the meantime, piracy becomes increasingly rampant as the customers can easily duplicate and redistribute the received multimedia content to a large audience [3].

## References

1. AlibertiPinder, L.F., Parham, G.P., Basu, P., et al.: Thermal ablation versus cryotherapy or loop excision to treat women positive for cervical precancer on visual inspection with acetic acid test: pilot phase of a randomised controlled trial. *Lancet Oncol.* **21**(1), 175–184 (2020)
2. Behrisch, M., Schreck, T., Pfister, H.: GUIRO: User-guided matrix reordering. *IEEE Trans. Visual Comput. Graph.* **26**(1), 184–194 (2020)
3. Al-Kharusi, I., Walker, D.W.: Locality properties of 3D data orderings with application to parallel molecular dynamics simulations. *Int. J. High Perform. Comput. Appl.* **33**(5), 998–1018 (2019)
4. Semple, T.L., Peakall, R., Tatarnic, N.J.: A comprehensive and user-friendly framework for 3D-data visualisation in invertebrates and other organisms. *J. Morphol.* **280**(2), 223–231 (2019)
5. Maria Klein Hentz, A., Strager, M.P.: Cicada (Magicicada) Tree Damage Detection Based on UAV Spectral and 3D Data. *Nat. Sci.* **10**(1), 31–44 (2018)
6. Zinzi, A., Ciarniello, M., Corte, V.D., et al.: The SSDC contribution to the improvement of knowledge by means of 3D data projections of minor bodies. *Adv. Space Res.* **62**(8), 2306–2316 (2018)
7. RNA-Primed Amplification for Noise-Suppressed Visualization of Single-Cell Splice Variants. *Anal. Chem.* **92**(13), 9356–9361 (2020)
8. High-Speed SICM for the visualization of nanoscale dynamic structural changes in hippocampal neurons. *Anal. Chem.* **92**(2), 2159–2167 (2020)
9. Selvaraju, R.R., Cogswell, M., Das, A., et al.: Grad-CAM: visual explanations from deep networks via gradient-based localization. *Int. J. Comput. Vis.* **128**(2), 336–359 (2020)
10. Oliveira, S.M.D., Bernardes, A.T., Sá Martins, J.S.: Self-organisation of female menopause in populations with child-care and reproductive risk. *Eur. Phys. J. B* **7**(3), 501–504 (2021)
11. Ramlawi, N., Bharadwaj, N.A., Ewoldt, R.H.: The weakly nonlinear response and nonaffine interpretation of the Johnson–Segalman/Gordon–Schowalter model. *J. Rheol.* **64**(6), 1409–1424 (2020)
12. Bensch, K., Then, C., Postigo, A., et al.: Exploring 3D data of dwarf galaxies. In: *Proceedings of the International Astronomical Union*, vol. 14(S344), pp. 359–362 (2018)
13. Infanger, A.: The use of 3D data in lineside component applications. *Signal + Draht*, **110**(12), 17–20 (2018)
14. James, B., Palmer, A., et al.: A Definitive Pipeline to Display Multivariate Geotechnical Data in a 3D Environment. *J. Nuclear Mater. Manage.* **46**(3), 81–96 (2018)
15. Hundera, H., Mpandeli, S., Bantider, A.: Spatiotemporal analysis of land-use and land-cover dynamics of Adama district, Ethiopia and its implication to greenhouse gas emissions. *Integr. Environ. Assess. Manag.* **16**(1), 90–102 (2020)
16. Atzori, M., Koutrika, G., Pes, B., et al.: Special issue on “Data Exploration in the Web 3.0 Age”. *Future Gener. Comput. Syst.* **112**(1), 1177–1179 (2020)



# Construction of Inspection Work Management System Based on Data Mining Technology

Hongchang Wen<sup>1</sup>(✉), Huiwen Peng<sup>2</sup>, Ying Zhu<sup>1</sup>, Yameng Wan<sup>1</sup>, and Hao Niu<sup>1</sup>

<sup>1</sup> Beijing Guodiantong Network Technology Co., Ltd., Beijing 100070, China  
18911522110@163.com

<sup>2</sup> State Grid Corporation of China, Beijing 100031, China

**Abstract.** Intra-party inspection system is an important measure to strengthen intra-party supervision and promote the construction of party style and clean government. It is also an effective way to build an anti-corruption work system and achieve institutional anti-corruption. Therefore, only with a scientific and efficient inspection system can inspections and inspections have a wide-ranging and long-lasting deterrent effect, and then promote the quality and efficiency of grassroots inspections. The purpose of this paper is to build a more complete system for inspection work based on data mining technology. In the experiment, a questionnaire survey was used to find many problems in the inspection work. Therefore, in order to play the role of inspection to a greater extent, it is necessary to self-examine the problem, analyze the cause of the problem, and solve the problem fundamentally, in order to live up to the party committee, the government and the masses' trust.

**Keywords:** Data Mining Technology · Patrol Inspection · Work Management · System Construction

## 1 Introduction

In order to improve the supervision system, the Party Central Committee attaches great importance to inspections and inspections, and has deployed inspections and inspections many times. The inspection team did not fail to be entrusted by the central government and the people, and achieved good results, laying a foundation for implementing the requirements of comprehensively and strictly governing the party, and was widely praised by party committees at all levels, governments at all levels and the masses. Although inspections and inspections are usually mentioned at the same time, inspections and inspections are still quite different. The inspections carried out at the central and provincial levels are called inspections, and those carried out at the county level are called inspections [1].

Carrying out inspection work is helpful for weaving the grassroots supervision network, building a management system for inspection work based on data mining, and helping to deeply solve the micro-corruption problem and the “four winds” around the masses. Data mining techniques are used in many fields. Marozzo F Extracting useful

information from data is often a complex process that can be easily modeled as a data analysis process. When very large data sets must be analyzed and/or complex data mining algorithms must be executed, data analysis processes can take a long time to complete their execution. Therefore, there is a need for efficient systems to perform data analysis workflows in a scalable manner using the computing services of cloud platforms that store data. The goal is to show how cloud software technologies can be integrated to achieve an efficient environment for designing and executing quantitative data analysis workflows. Describes the design and implementation of a data mining cloud framework, a data analysis system that integrates a visual workflow language and parallel runtime with a service-as-a-service model [2]. Sattarian M believes that IoT is becoming an important part of daily life, and in the future of modern cities, there will be a large number of IoT-based scenarios, from small indoor areas to large outdoor areas. In this day and age, mobility is still an important factor in both outdoor and indoor environments, and in both cases there are many solutions. On the other hand, recent smart objects have generated huge amounts of data, requiring sophisticated data mining solutions to process them. Research on the application of data mining technology in indoor navigation system in the Internet of Things scenario. The goal is to understand the types of navigation problems that exist in different IoT scenarios, focusing on indoor environments, and then we study how data mining solutions can provide solutions to these challenges [3]. Data mining technology is widely used in various fields, inspection work is still a “new thing”, and the inspection work management system based on data mining technology is not very perfect.

In the introduction of data mining, this paper expounds the brief introduction of data mining and the overview of data mining technology, and studies the connotation and difference of inspection, as well as the theoretical basis of inspection work. In the experiment, taking the inspection work of city M as an example, the questionnaire survey was used to investigate and study the investigation results of the inspection work, and the investigation results of the inspection work.

## **2 Research on the Construction of Inspection Work Management System**

### **2.1 Introduction to Data Mining**

#### **(1) Brief introduction to data mining**

The concept of data mining originated from the annual computer conference in the United States. Subsequently, data mining technology has been continuously improved and extended under the efforts of many experts and scholars, and has received extensive attention [4]. Up to now, the definition of data mining has different interpretations in different literatures. Through the research and understanding of data mining technology, the author believes that “data mining is a process from a large number of, incomplete, noisy, fuzzy, In random system or application data, the process of mining meaningful and understandable data through analysis, cleaning, etc.” This definition is more in line with the data mining technology itself. The definition highlights two points. First, the data used by data mining technology is

non-standard data with different forms in the real system. Second, the output of data mining technology must be meaningful and valuable results. The ultimate purpose of data mining technology is to provide in-depth decision support and behavioral guidance for data producers and data maintainers, so data mining technology has rich application scenarios and is highly attractive to users.

(2) Overview of data mining technology

The concept of data mining technology was proposed late, but the research on its predecessor technology was very early [5]. It is a comprehensive discipline, covering many technologies such as machine learning, pattern recognition, intelligent database, expert system, high-performance computer, etc. With the advent of the era of information explosion, this technology has been difficult to meet people's needs. Without the formation of database and data warehouse technology, the data mining process is difficult to carry out. People are freed from complicated and massive business data, allowing people to save time and energy, and devote more time to data analysis to find data models that are very valuable for optimizing business. This factor also promotes data mining. Improvement of technology. The competition within the industry and even between different industries is becoming increasingly fierce. People need to understand the current situation in the field and the prediction of the future through a large amount of historical data. These contents are difficult to find through simple manual analysis, but are very necessary for decision-making, it can find the root cause of many problems in the field [6].

## 2.2 The Connotation of Inspection

(1) Inspection in city

According to the relevant explanations, "patrol" means "patrol, visiting and inspecting", and "shi" means "inspecting, watching", so the general concept of "patrol" can be understood as "travel and inspection, while walking and watching, while observing on the other hand, inspection. Inspection work is a "political physical examination" carried out on party organizations and party members and cadres. The purpose is to find and solve problems, promote reform and improvement, and carry out inspection work for strengthening the supervision and management of party organizations and party members and cadres. It has extremely important practical significance [7].

(2) Cruising

The basic meaning of "Cruising", as literally understood, is the meaning of inspection. It checks the actual situation according to the target, records and organizes reports in a timely manner, with irregularity and flexibility [8]. Inspection is an extension of inspection at the city and county level. It is an effective supplement to the insufficient coverage of inspections. It is of great significance to strengthen inner-party supervision and improve the inner-party supervision system. The inspections are dispatched by the central and provincial party committees at the first level. They are highly authoritative and pay more attention to the macro and overall situation. The inspections focus on "Cruising", that is, to understand the situation. Compared with the inspection, the inspection level is relatively low,



and the authority is not as good as the inspection, but its role should not be underestimated. In-depth practice of the people's thinking, he attaches great importance to inspections and inspections. His important speeches and expositions greatly enrich the theory and connotation of inspections and inspections, and provide fundamental guidelines for doing a good job in inspections and inspections.

(3) The difference between inspection

Inspections, like inspections, are both important means of implementing intra-party supervision. During one term, inspections of the party organizations under their jurisdiction must be completed. Inspections and inspections are the needs of comprehensively and strictly governing the party. The party must manage the party without leaving blank areas and blind spots, and must not miss any omissions.

One of the differences is that the dispatching subjects are different, and the dispatching level of the inspection team is higher than that of the inspection team. Inspection teams are dispatched by the central and provincial levels, and inspection teams are dispatched by cities, counties, and districts [9].

The second difference is that the objects of supervision are different. The inspection objects are the leading cadres at the provincial, ministerial and bureau level, and the inspection objects are the cadres at the county, division and township level and below.

The third point of difference is that the focus of supervision is different. The focus of inspections is on "viewing". From top to bottom, it often focuses on the macro and the overall situation, while the focus of inspections is on "observation", that is, to understand the specific situation and keep close to the grassroots. The point is to spot the unhealthy trends around the masses.

### 2.3 Theoretical Basis for Research on Inspection Work

Inspection is the extension of inspection to cities and counties. It can be said that research inspection work is the study of the extension of inspection to the grassroots level. The inspection system has certain similarities with the ancient supervision system, which is the inheritance and improvement of the ancient Chinese supervision system [10].

(1) The theory of inner-party supervision.

For western political life, party politics is the main manifestation of western politics, and the supervision of political parties is also an important part of the improvement of political parties. Since the western capitalist countries mainly implement the competition mechanism of political parties, they are more inclined to research and improve the supervision of the ruling party. The theory of intra-party supervision of the Communist Party of China has undergone historical exploration and has Chinese characteristics. It was established and improved in the process of applying Marxism-Leninism from the beginning of the party's establishment, combining with China's national conditions, guiding the Chinese revolution to victory, and strengthening the party's construction and exploration and practice. With the improvement and growth of the party, it keeps up with China's improvement and changes, and constantly innovates and improves.

## (2) Principal-agent theory.

As one of the main theories of economics, the principal-agent theory mainly discusses the formation of “agent costs”, which are often formed in modern corporate enterprises with separation of enterprise ownership and management rights. Asymmetric, principal-agent relationship often has the problem of agency cost. The field of principal agent is also widely used in the field of political science. The principal-agent system is also often used in the field of political science. The principal is played by people from all over the country, and the government becomes the agent.

## (3) The theory of power constraints.

Having power can serve the people better, but too much power can lead to corruption. So to avoid the occurrence of absolute power. The Western research on rights restriction has certain drawbacks and is not suitable for my country’s national conditions and value system, but there are still some experiences and explorations worthy of our reference. The theory of China’s control of power is based on China’s national conditions and combined with Marxist thought.

### 3 Investigation and Research on Inspection Work

#### 3.1 Sample Selection

Taking the inspection work of M city as an example, in the process of inspection work, new problems are constantly discovered, solved, and finally gained experience, so as to continuously improve work ability and achieve better work results. In order to have a clearer understanding of the progress and existing problems of the inspection work in M city, we used a questionnaire to understand the basic situation of the inspection work in M city, and analyze the effect of the work.

#### 3.2 Research Methods

The content of the questionnaire is divided into two parts, the first part is the basic situation of the sample, and the second part is the effect of specific questions. During the investigation, a total of 150 questionnaires were distributed, half of which were filled out by the staff of the inspection unit and half by the inspectors. Finally, 140 questionnaires were recovered, with an effective rate of 90%. Through the feedback of the inspected units and inspectors, the problems existing in the inspection work are analyzed from two aspects, so as to ensure the relative objectivity and accuracy of the investigation results. The t-test formula used in this paper is as follows:

$$t = \frac{\bar{X} - \mu}{\frac{\sigma_X}{\sqrt{n}}} \quad (1)$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad (2)$$

Among them, formula (1) is the single population test, which  $\bar{x}$  is the sample mean,  $s$  is the sample standard deviation, and  $n$  is the number of samples. Equation (2) is a double population test.

## 4 Problems and Result Analysis of Investigation in Inspection Work Management System

### 4.1 Investigation Results of Inspection Work on Major and Difficult Rectification Problems

In view of the existing work of inspection in M City, the key and difficult rectification work after the inspection is analyzed, as shown in Table 1 and Fig. 1:

**Table 1.** M Work contents and effect of work City

Survey content	Satisfied	Basically satisfied	Unsatisfied
Committee special meeting discipline assigned problem management effect	19%	51%	30%
The governance effect of the problems left over from history	29%	31%	40%
The governance effect of collective and individual bad behavior	6%	51%	43%

It can be seen from the above table that in the inspection work of M City, the overall satisfaction with the effect of the management of major and difficult points is still at a low level. The dissatisfaction with the governance of the problems assigned by the municipal party committee's special meeting is as high as 30%, and the dissatisfaction with the governance effect of historical problems and collective and individual bad behaviors is above 43%. For government agencies, further reforms are still needed.

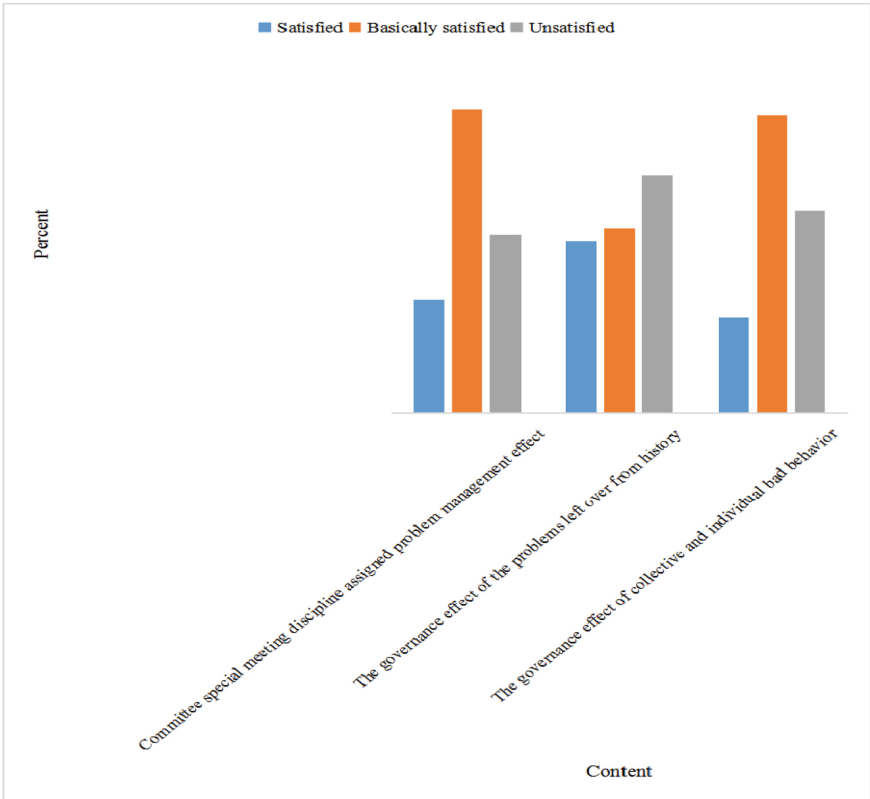


Fig. 1. Comparison of the effect of inspection in M City

#### 4.2 Investigation on the Effect of Inspections to Solve Both the Symptoms and Root Causes

In view of the current work of return visit and inspection in M City, this paper analyzes the overall rectification effect after its inspection, that is, the effect of treating both the symptoms and the root causes, as shown in Table 2 and Fig. 2:

Table 2. Statistical Table of the effect of treating both the symptoms and root causes in M City

Survey content	Satisfied	Basically satisfied	Unsatisfied
Check the reform ratio of prisoners involving problems	26%	36%	38%
The inspection involved the recurrence ratio of the problem	35%	45%	20%
The construction of party conduct and government conduct	23%	27%	50%



Fig. 2. Comparison of the effect of treating both symptoms and root causes in M City

It can be seen from the above table that in the inspection work of M City, the overall satisfaction with the effect of treating both the symptoms and root causes is still at a low level. The phenomenon of committing crimes while making corrections is relatively prominent, and the overall satisfaction rate is only 26%. In addition, only 35% of people are satisfied with the governance satisfaction of the construction of party style and government style. It is only 23%, indicating that such problems still need attention.

### 5 Conclusions

After all, inspection is a new thing, and it is still in the exploratory stage. Whether some practices are appropriate still needs time and practice to test, and some existing problems cannot be solved immediately, and need to be carried out step by step. Only through continuous optimization of inspection work and strengthening of institutional mechanisms and other constructions can we provide a guarantee for the rapid and high-quality improvement of inspection work, make inspection work effective, and further promote the formation of a clean and upright political ecology, so as to protect the party and the country. The various businesses are progressing smoothly. Research Prospects Due to the limited time and my level, further research is still needed, and more in-depth inspections are needed. The spirit of the central government's documents on inspections and inspections needs to be studied in depth. It is necessary to actively observe the inspections carried out in other nearby counties and cities. Summarize more experience, do more exchanges, and actively make suggestions for improving inspections..

### References

1. Benitez, P., Rocha, E., Talukdar, S., et al.: Efficiency analysis of optimal inspection management for reinforced concrete structures under carbonation-induced corrosion risk. *Constr. Build. Mater.* **211**(6), 1000–1012 (2019)
2. Marozzo, F., Talia, D., Trunfio, P.: A workflow management system for scalable data mining on clouds. *IEEE Trans. Serv. Comput.* **11**(3), 480–492 (2018)

3. Sattarian, M., Rezazadeh, J., Farahbakhsh, R., Bagheri, A.: Indoor navigation systems based on data mining techniques in internet of things: a survey. *Wirel. Netw.* **25**(3), 1385–1402 (2018). <https://doi.org/10.1007/s11276-018-1766-4>
4. Deutsch, Y., Boaz, G.: The effect of risk aversion on the outcomes of inspection games. *J. Oper. Res. Soc.* **69**(5), 645–660 (2018)
5. Azeem, M., Mataruna, L.: Identifying factor measuring collective leadership at academic workplaces. *Int. J. Educ. Manag.* **33**(6), 1316–1335 (2019)
6. Maeda, C., Kaneko, R., et al.: A practical data verification process for introducing of road inspection equipment. *J. Jpn. Soc. Civil Eng. Ser. F5 (Prof. Pract. Civil Eng.)* **75**(1), 1–9 (2019)
7. Paolo, T., Valentino, P., Giuseppe, C.: [The AUDIT tool in security management at work in companies operating in the port area of Trieste. An interinstitutional prevention project]. *Giornale italiano di medicina del lavoro ed ergonomia* **414**(4), 316–319 (2020)
8. Rafa, L.W., Marc, A., Maggie, G., et al.: Preventing psychosocial risks at work: an evaluation study of labour inspectorate interventions. *Saf. Sci.* **110**(1), 355–362 (2018)
9. Yamaura, J., Muench, S.T.: Assessing the impacts of mobile technology on public transportation project inspection. *Autom. Constr.* **96**(12), 55–64 (2018)
10. Pointon, A., Hamilton, D., Kiermeier, A.: Assessment of the post-mortem inspection of beef, sheep, goats and pigs in Australia: approach and qualitative risk-based results. *Food Control* **90**(1), 222–232 (2018)



# Constructing and Evaluating Service Quality Evaluation Index System of Airline Using PSM and Entropy Weight

Liyang Wang<sup>1</sup> (✉), Hongtao Li<sup>1</sup>, Jing Bai<sup>1</sup>, Xingming Lu<sup>2</sup>, Pengpeng Zhou<sup>1</sup>, and Zhijie Liu<sup>1</sup>

<sup>1</sup> China Academy of Civil Aviation Science and Technology, Beijing, China  
wangliy@mail.castc.org.cn

<sup>2</sup> All China Federation of Supply and Marketing Cooperatives, Beijing, China

**Abstract.** The service quality evaluation is of great significance for airlines to improve their service quality and thus optimize management decisions. In order to scientifically and reasonably evaluate the current situation of airline service quality, this paper first constructs the airline service quality evaluation system based on the SERVQUAL model. Secondly, after considering the advantages of subjective weighting methods and objective weighting methods, this paper proposes a combined weight evaluation method, i.e., PSM-Entropy weight method. Through the investigation of 650 participants, result showed that PSM-Entropy weight method can scientifically and effectively reflect the contribution of each index to the airline service quality evaluation system, providing a new idea for the airline service quality evaluation method.

**Keywords:** Service Quality Evaluation Index System · Airline · Psm · Entropy Weight

## 1 Introduction

Service quality is an important dimension to measure the service level of airlines, a magic weapon to maintain their competitive advantages, and plays an important guiding role in Airlines' management decisions and future development [1, 2]. Nowadays, as service quality is paid more and more attention by the market and passengers [3, 4], an objective, scientific and comprehensive evaluation of airline service quality is the key to effectively improving service levels and optimizing management decisions.

Airline service quality evaluation is a multi-dimensional and multi-index value evaluation mechanism that comprises three processes: the one is to build an evaluation index system according to the characters of the airline service; the second is to determine the weight measurement method according to the characteristics of each index; the third is to determine the final evaluation results according to the index system and weight. Particularly, the key is to determine the weight measurement method [5–7]. The existing research on evaluation methods mainly discusses subjective weighting methods (such as

analytic hierarchy process [8], product scale method [9], and objective weighting methods (such as entropy weight method [8], rough set theory [10]). Subjective weighting method is to judge the importance of each evaluation index based on the experience and knowledge of industry experts. It can well reflect the will of decision makers but susceptible to the subjective will of decision makers. And objective weighting method is to determine the weight based on the relationship between the original data. It can objectively reflect the real situation but easy to ignore the will of the decision makers.

This paper first builds an airline service evaluation system based on SERVQUAL model after analyzing the key factors affecting passengers’ perception of airline service quality. Secondly, this paper uses PSM-Entropy weight method to determine the index weight after considering the multi-dimensional complexity characteristics of airline service evaluation and the advantages of subjective weighting methods and objective weighting methods. Through the investigation of 650 participants, it is found that PSM-Entropy weight method can scientifically and effectively reflect the contribution of each index to the airline service quality evaluation system.

## 2 Construction of Service Quality Evaluation Model Based on SERVQUAL

### 2.1 Construction of Service Quality Evaluation Index System Based on SERVQUAL

First, we sort out the rules, industry standards and existing research. Then, we analyze the key processes of airline service delivery and the key factors affecting passengers’ perception of airline service quality. Finally, we construct an index system of airline service quality evaluation based on SERVQUAL (Fig. 1).

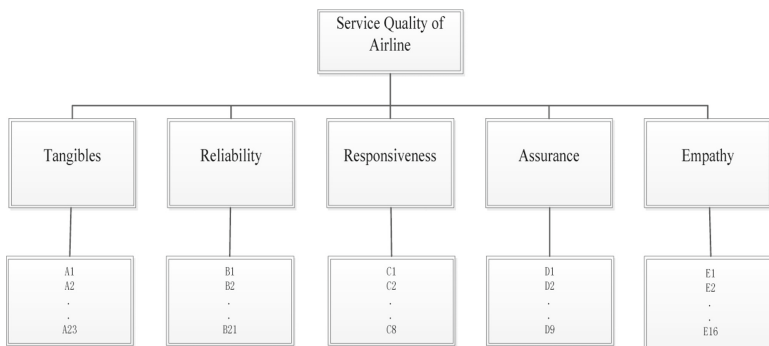


Fig. 1. Service quality evaluation index system

### 2.2 Determination of Weight Measurement Method

Product Scale method (PSM) is a commonly used subjective fuzzy evaluation method. Its basic idea is to set only two grades instead of dividing too many grades when comparing



the importance of index in pairs. That is, we should first find the cardinality in order of importance, and then compare all indexes with that cardinality for progressive product analysis. It is found that this method has the characteristic of greater flexibility. Entropy weight theory is an objective weighting method. Its basic principle is to calculate the entropy weight according to the information contained in the entropy value of each evaluation index, and then modify the weight of each index through the entropy weight to obtain a more objective weight.

Although the product scale method can reflect the decision maker’s intention well, it is easily affected by evaluator’s subjective will. Entropy weight method mainly relies on mathematical theories and methods to calculate weight, which is not easily affected by the subjective will of decision-makers and differences of indicators themselves, but may ignore the real situation. In short, this paper takes into account the advantages of subjective weighting method and objective weighting method, and selects PSM-Entropy weight method to determine the weight, so as to avoid the shortcomings of the two methods.

### 3 Methods and Results

#### 3.1 Questionnaire and Survey Analysis

In order to objectively evaluate the service quality of airlines, this paper conducted a two-month survey among experts and passengers in the field of civil aviation. After eliminating invalid questionnaires, a total of 650 valid questionnaires were received.

The paper firstly conducts exploratory factor analysis and correlation analysis with SPSS. The results showed that the KMO value was 0.963, and the correlation coefficient between variables was not more than 0.700, indicating that the questionnaire had good validity. Then, the paper conducts reliability analysis. It was showed that the  $\alpha$  of each dimension is greater than 0.800, indicating that the scale we used is reliable.

#### 3.2 Evaluation Index Weight Values of Each Dimension

The steps for calculating evaluation index weight values of each dimension are as follows:

- (1) Index weight calculation by product scale method

According to the principle of product scale method and formula (1), the weights of evaluation indexes affecting airline service quality in all dimensions are calculated. The results are shown in Tables 1 and 2.

$$(\omega A : \omega B) = \left( \frac{1.35^k}{1 + 1.35^k}, \frac{1}{1 + 1.35^k} \right) \tag{1}$$

- (2) Index weight calculation by entropy weight method

According to the principle of entropy weight method and formula (2), (3), the weights of evaluation indexes affecting airline service quality in all dimensions are calculated. The results are shown in Tables 1 and 2.

$$e_j = -\frac{1}{\ln(m)} \sum_{i=1}^m z_{ij} \ln(z_{ij}), j = 1, 2, \dots, n \tag{2}$$

$$w_j = \frac{d_j}{\sum_{j=1}^n d_j}, j = 1, 2, \dots, n \tag{3}$$

(3) Final index weight calculation

Combined with the index weights values by product scale method and the index weight values by entropy weight method, the comprehensive weights of evaluation indexes at all levels are calculated according to formula (4). The results are shown in Tables 1 and 2.

$$\varpi = \left\{ \frac{(\omega_1 w_1)^{0.5}}{\sum_{j=1}^n (\omega_j w_j)^{0.5}}, \frac{(\omega_2 w_2)^{0.5}}{\sum_{j=1}^n (\omega_j w_j)^{0.5}}, \dots, \frac{(\omega_n w_n)^{0.5}}{\sum_{j=1}^n (\omega_j w_j)^{0.5}} \right\} = (\varpi_1, \varpi_2, \dots, \varpi_n) \tag{4}$$

**Table 1.** Weight values of first-level evaluation indexes

dimension	weight by PSM	Entropy weight	combination weight
Tangibles	0.0997	0.2147	0.1496
Reliability	0.2475	0.1891	0.1634
Responsiveness	0.3351	0.1934	0.1923
Assurance	0.1828	0.1924	0.2232
Empathy	0.1350	0.2103	0.2715

**Table 2.** Weight values of second-level evaluation indexes

First-level index	Item	weight by PSM	Entropy weight	combination weight
Tangibles	A1	0.0009	0.0322	0.0069
	A2	0.0611	0.0662	0.0828
	A3	0.0029	0.0271	0.0116
	A4	0.1516	0.0471	0.1100
	A5	0.0246	0.0225	0.0307
	A6	0.0040	0.0324	0.0148
	A7	0.0005	0.0341	0.0053
	A8	0.1120	0.0430	0.0904
	A9	0.0022	0.0314	0.0108
	A10	0.2779	0.0443	0.1446
	A11	0.0016	0.0312	0.0092

(continued)

**Table 2.** (continued)

First-level index	Item	weight by PSM	Entropy weight	combination weight
	A12	0.0054	0.0468	0.0207
	A13	0.2052	0.0478	0.1290
	A14	0.0099	0.0458	0.0277
	A15	0.0134	0.0479	0.0330
	A16	0.0333	0.0503	0.0533
	A17	0.0827	0.0436	0.0782
	A18	0.0451	0.0315	0.0491
	A19	0.0004	0.0309	0.0043
	A20	0.0006	0.0727	0.0089
	A21	0.0012	0.0584	0.0108
	A22	0.0182	0.0538	0.0407
	A23	0.0073	0.0589	0.0270
Reliability	B1	0.0985	0.0470	0.0772
	B2	0.1334	0.0452	0.0880
	B3	0.0035	0.0531	0.0155
	B4	0.0728	0.0544	0.0714
	B5	0.0985	0.0423	0.0732
	B6	0.1806	0.0473	0.1048
	B7	0.1806	0.0407	0.0972
	B8	0.0397	0.0503	0.0507
	B9	0.0537	0.0564	0.0624
	B10	0.0293	0.0466	0.0419
	B11	0.0397	0.0468	0.0489
	B12	0.0026	0.0348	0.0108
	B13	0.0118	0.0583	0.0297
	B14	0.0087	0.0468	0.0229
	B15	0.0160	0.0427	0.0296
	B16	0.0216	0.0440	0.0350
	B17	0.0064	0.0550	0.0213
	B18	0.0019	0.0502	0.0111
	B19	0.0728	0.0315	0.0543
	B20	0.0216	0.0425	0.0344

(continued)

**Table 2.** (continued)

First-level index	Item	weight by PSM	Entropy weight	combination weight
Responsiveness	B21	0.0048	0.0639	0.0198
	C1	0.0670	0.1191	0.0817
	C2	0.2250	0.1064	0.1416
	C3	0.2250	0.1350	0.1596
	C4	0.1227	0.1186	0.1105
	C5	0.1662	0.1207	0.1296
	C6	0.3047	0.1089	0.1667
	C7	0.0907	0.1584	0.1097
Assurance	C8	0.0907	0.1330	0.1005
	D1	0.3221	0.1229	0.1570
	D2	0.4361	0.1061	0.1697
	D3	0.0708	0.1174	0.0719
	D4	0.0523	0.1029	0.0579
	D5	0.1297	0.1317	0.1031
	D6	0.1757	0.0988	0.1040
	D7	0.2379	0.1148	0.1304
	D8	0.2379	0.1327	0.1402
Empathy	D9	0.0958	0.0725	0.0658
	E1	0.1232	0.0817	0.0994
	E2	0.0200	0.0541	0.0326
	E3	0.0496	0.0553	0.0519
	E4	0.0672	0.0540	0.0597
	E5	0.0271	0.0532	0.0376
	E6	0.0496	0.0593	0.0538
	E7	0.0148	0.0511	0.0272
	E8	0.0672	0.0984	0.0806
	E9	0.0910	0.0648	0.0761
	E10	0.0496	0.0671	0.0572
	E11	0.2258	0.0626	0.1178
	E12	0.1668	0.0654	0.1035
	E13	0.0910	0.0594	0.0728
E14	0.0271	0.0654	0.0417	

(continued)

**Table 2.** (continued)

First-level index	Item	weight by PSM	Entropy weight	combination weight
	E15	0.0366	0.0598	0.0464
	E16	0.0366	0.0483	0.0417

## 4 Conclusions

This paper constructs the airline service quality evaluation system based on SERVQUAL model. Through the analysis of subjective weighting methods and objective weighting methods, this paper puts forward a new weight evaluation method, i.e., PSM-Entropy weight method. Specifically, this paper uses PSM-Entropy weight method to measure the weights of each primary and secondary index. Among all the first-level indexes, empathy (0.2715) has the highest weight, and tangibles (0.1496) has the lowest weight. Empirical research shows that this method takes into account the advantages of subjective weighting methods and objective weighting methods at the same time. In addition, the method can also comprehensively, systematically and effectively reflect the contribution of each index to the service quality evaluation system, and better reflect the characteristics of airline service quality.

**Acknowledgements.** This research was financially supported by Key technologies on Civil aviation service quality monitoring and risk early warning based on big data (Grant No.: x222060302544).

## References

1. Muhammad, A., Sadaf, Z., Asif, Y.: Impact of service quality and trust on repurchase intentions – the case of Pakistan airline industry. *Asia Pac. J. Mark. Logist.* **29**(5), 1136–1159 (2017)
2. Chen, C.: Investigating structural relationships between service quality, perceived value, satisfaction, and behavioral intentions for air passengers: evidence from Taiwan. *Transp. Res. Part A* **42**(4), 709–717 (2008)
3. Ana, B., Paulo, R., Cristina, O., Fernando, O.: Airline passengers' perceptions of service quality: themes in online reviews. *Int. J. Contemp. Hosp. Manag.* **31**(2), 855–873 (2019)
4. Lim, J., Lee, H.: Comparisons of service quality perceptions between full service carriers and low cost carriers in airline travel. *Curr. Issue Tour.* **23**(10), 1261–1276 (2020)
5. Wang, T., Pham, Y.: An application of cluster analysis method to determine Vietnam airlines' ground handling service quality benchmarks. *J. Adv. Transp.* **2020**(1), 1–13 (2020)
6. Vaidya, O., Kumar, S.: Analytic hierarchy process: an overview of applications. *Eur. J. Oper. Res.* **169**(1), 1–29 (2006)
7. Chou, C., Liu, L., Huang, S., Yih, J., Han, T.: An evaluation of airline service quality using the fuzzy weighted SERVQUAL method. *Appl. Soft Comput. J.* **11**(2), 2117–2128 (2011)
8. Wu, J.: Study on library characteristic services quality evaluation method based on LibQual + combining with Ridit analysis and its practice analysis. *J. Mod. Inf.* **35**(8), 75–107 (2015)

9. Wang, Q., Wu, C., Sun, Y.: Evaluating corporate social responsibility of airlines using entropy weight and grey relation analysis. *J. Air Transp. Manag.* **42**, 55–62 (2015)
10. Karimi, T., Hojati, A., Forrest, J.: A new methodology for sustainability measurement of banks based on rough set theory. *CEJOR* **30**(1), 415–431 (2022)



# Improving Service Quality Measurement with Servqual: Evidence from Airport Industry

Liying Wang<sup>1</sup>, Hongtao Li<sup>1</sup>(✉), Xingming Lu<sup>2</sup>, Chenxi Zheng<sup>1</sup>, and Licheng Xu<sup>1</sup>

<sup>1</sup> China Academy of Civil Aviation Science and Technology, Beijing, China  
1070665884@qq.com

<sup>2</sup> All China Federation of Supply and Marketing Cooperatives, Beijing, China

**Abstract.** This study aims to conduct a systematic examination of airport Service Quality. Firstly, this paper constructs the airport service quality evaluation system using interview method, literature method and experience summary method. Secondly, this paper uses PSM-Entropy weight method to measure the index weight of airport service quality evaluation system. Finally, this paper grades the level of airport service quality. Specifically, this paper divides the airport service quality evaluation system into 5 first-level indicators and 80 s-level indicators. Then this paper uses PSM-Entropy weight method to determine the weight of primary and secondary indicators, in which the responsiveness weight is the highest (0.2465) and the empathy weight is the lowest (0.1626). Finally, through the service quality rating, it is found that the airport service quality is at the four-star service level and needs to be further improved.

**Keywords:** Service Quality Measurement · Servqual · PSM-Entropy

## 1 Introduction

The increasing popularization and generalization of air travel brings more diversified needs of passengers, which makes passengers pay more attention to the overall service quality of air. This change has brought new opportunities and challenges to the development of the airport [1, 2]. It has been shown that the improvement of service quality is an important magic weapon for the airport [3–5]. Therefore, establishing and improving the airport service quality evaluation system and accurately identifying the current situation, situation and weaknesses of airport service quality has become a major theoretical and practical topic that must be overcome in airport development.

Although many studies have constructed and measured the service quality in the aviation field based on SERVQUAL model, they have not considered all the factors that may affect the service quality, and mainly measured service quality in the context of airline rather than airport. For example, [6] uses 17 factors to evaluate service quality. [7] identify 27 factors affecting service quality when studying travel behavior intention. [8] uses 22 indicators representing service quality to measure the relationship between service quality and passenger safety perception. [9] uses 26 indicators to measure the service quality of the Turkish aviation industry. [10] uses 24 indicators to measure the

service quality of airline. [11] uses 13 indicators to evaluate the service quality of the air baggage handling system.

This paper takes airport as the research object, and comprehensively considers all service links, including airport traffic, check-in and security, baggage services, environment and facilities, departure and arrival services, flight delay services, special passenger services, complaint handling, etc. Combined with the whole process service characteristics of the airport and the key elements affecting passengers' perception of airport service quality, this paper constructs the airport service quality evaluation system based on SERVQUAL model. Through the investigation of 618 subjects, it is found that passengers gave the airport a four-star service rating of 84.40. That is, the overall airport service is above medium level and needs to be further improved.

## **2 Methods**

### **2.1 Construction of Airport Service Quality Evaluation System**

This paper uses the methods of interview, literature and experience summary to analyze the key factors affecting passengers' perception of airport service quality. Then, this paper constructs an airport service quality evaluation system based on SERVQUAL according to the industry norms and the characteristics of the airport providing services for passengers. Specially, the system has 5 first-level indexes and 80 s-level indexes.

### **2.2 PSM-Entropy Weight Method**

PSM-Entropy weight method is a combination method consisting of product scale method (PSM) and entropy weight method. Generally, the product scale method is used to calculate the weight of each index from the subjective level. In other words, when adopting this method, we should first invite experts to evaluate the importance of each index, and then calculate the weight according to the principle of the product scale method to scale differences in importance. Secondly, the entropy weight method is used to calculate the weight of each index from the objective level. That is, we invite passengers to score the importance of the index and calculate the weight according to the information contained in the entropy value of each evaluation index. This method can obtain objective evaluation results, as well as consider the opinions of decision makers, thus enhancing the scientific nature of the results.

### **2.3 Questionnaire Survey**

We used a 5-point Likert scale to survey experts and passengers in the field of civil aviation. A total of 618 valid questionnaires were collected. Then we conduct exploratory factor analysis and correlation analysis with SPSS. The results showed that the KMO value was 0.950, and the correlation coefficient between variables was not more than 0.700, indicating that the questionnaire had good validity. Then, the paper conducts reliability analysis. It was showed that the  $\alpha$  of each dimension is greater than 0.700, indicating that the scale we used is reliable.



### 3 Results and Analysis

#### 3.1 Index Weight Calculation

Firstly, according to the principle of product scale method and formula (3.1), this paper calculates the weights of evaluation indexes affecting airport service quality in all dimensions. The results are shown in Tables 2 and 3.

$$(\omega A : \omega B) = \left( \frac{1.35^k}{1 + 1.35^k}, \frac{1}{1 + 1.35^k} \right) \tag{3.1}$$

$$e_j = -\frac{1}{\ln(m)} \sum_{i=1}^m z_{ij} \ln(z_{ij}), j = 1, 2, \dots, n$$

Secondly, according to the principle of entropy weight method and formula (3.2) and (3.3) this paper calculates the entropy weights of evaluation indexes in all dimensions. The results are shown in Tables 2 and 3.

$$e_j = -\frac{1}{\ln(m)} \sum_{i=1}^m z_{ij} \ln(z_{ij}), j = 1, 2, \dots, n \tag{3.2}$$

$$w_j = \frac{d_j}{\sum_{j=1}^n d_j}, j = 1, 2, \dots, n \tag{3.3}$$

Finally, combined with the index weights values by product scale method and the index weight values by entropy weight method, the comprehensive weights of evaluation indexes at all levels are calculated according to formula (3.4). The results are shown in Tables 2 and 3.

$$\varpi = \left\{ \frac{(\omega_1 w_1)^{0.5}}{\sum_{j=1}^n (\omega_j w_j)^{0.5}}, \frac{(\omega_2 w_2)^{0.5}}{\sum_{j=1}^n (\omega_j w_j)^{0.5}}, \dots, \frac{(\omega_n w_n)^{0.5}}{\sum_{j=1}^n (\omega_j w_j)^{0.5}} \right\} = (\varpi_1, \varpi_2, \dots, \varpi_n) \tag{3.4}$$

#### 3.2 Comprehensive Evaluation Results

According to formula (3.5), this paper calculates the service quality of each dimension. The results are shown in Tables 3.

$$R_i = \sum_{i=1}^n \varpi_i * f_i \tag{3.5}$$

- Note:  $R_i$  is the total score of the service quality evaluation of the first-level indexes.
- $\varpi_i$  is the comprehensive weight of second-level indexes.
- $f_i$  is the average satisfaction score of second-level indexes.
- $n$  is the total number of bottom-level indexes.

The paper converts the R in formula (3.5) into percentile system, and then evaluate the service quality level with reference to Table 1.

**Table 1.** Service quality rating criteria

Quality level	Five-star service level	Four-star service level	Three-star service level	Two-star service level	One-star service level
Score	100–86	85–75	74–60	59–30	Less than 30

### 3.3 Overall Evaluation Results

Combined with the comprehensive weights and comprehensive evaluation scores of first-level indexes, we calculate the passenger’s overall evaluation of airport service quality through formula.  $Q = \sum_{i=1}^n \varpi_i R_i$

Note:  $Q$  is the total score of the service quality evaluation of the first-level indexes.

$\varpi_i'$  is the comprehensive weight of first-level indexes.

$R_i$  is the average satisfaction score of first-level indexes.

$n$  is the total number of bottom-level indexes.

$$Q=(0.1991, 0.2118, 0.2465, 0.1799, 0.1626) * (4.34, 4.25, 3.80, 4.46, 4.43)T=4.22 \cdot$$

### 3.4 Analysis

**Table 2.** Evaluation results of first-level indexes

dimension	weight by PSM	Entropy weight	Combination weight	Result	Percentile Score	Service level	Overall Score and Service Level
Tangibles	0.1688	0.2277	0.1991	4.34	87	Five-star level	4.22 (84)
Reliability	0.2285	0.1903	0.2118	4.25	85	Four-star level	
Responsiveness	0.3094	0.1903	0.2465	3.80	76	Four-star level	
Assurance	0.1688	0.1859	0.1799	4.46	89	Five-star level	
Empathy	0.1246	0.2057	0.1626	4.43	89	Five-star level	

According to Table 2 and Table 3, the passenger’s comprehensive score for the airport’s service quality is 4.22 point, and the corresponding percentage score is 84.40 point on the percentile scale, which is rated as a four-star service. The results show that the airport’s service quality is at an upper-middle level, which has not effectively met the needs of passengers, and needs to be further improved and improved. According

to Table 2 and Table 3, it can be seen that the three dimensions of service quality, including tangibles, assurance and empathy has been recognized by passengers and rated as five-star service. Reliability and responsiveness were rated as four-star service; especially responsiveness only scored 76 point. This shows that the airport still has many deficiencies and needs to be improved in terms of relevant information notification, prompts, and remedial measures after flight delays, especially in the response speed of passengers' demands and the handling speed of passengers' complaints, which affect passengers' overall cognition of airport service quality. In order to obtain the continuous satisfaction of passengers, the airport should provide the promised services on time and accurately, care for passengers who encounter difficulties and try their best to help passengers, so as to quickly improve the service level to meet the needs of passengers.

**Table 3.** Evaluation results of second-level indexes

Second index	Weight by PSM	Entropy weight	Combination weight	Average Score
A1	0.0351	0.0126	0.0263	4.48
A2	0.0141	0.0113	0.0158	4.43
A3	0.0042	0.0198	0.0114	4.56
A4	0.0009	0.0142	0.0045	4.47
A5	0.0259	0.0330	0.0366	4.34
A6	0.0077	0.0161	0.0139	4.63
A7	0.0192	0.0267	0.0283	4.45
A8	0.0017	0.0291	0.0088	4.59
A9	0.0644	0.0132	0.0364	4.38
A10	0.0141	0.0231	0.0226	4.50
A11	0.0105	0.0182	0.0173	4.40
A12	0.0031	0.0264	0.0113	4.52
A13	0.0017	0.0158	0.0065	4.36
A14	0.0476	0.0353	0.0512	4.34
A15	0.0013	0.0174	0.0058	4.33
A16	0.0031	0.0426	0.0144	4.68
A17	0.0872	0.0417	0.0754	4.22
A18	0.0007	0.0194	0.0046	4.63
A19	0.0003	0.0234	0.0032	4.55
A20	0.0005	0.0186	0.0038	4.46
A21	0.0057	0.0154	0.0117	4.20

(continued)

**Table 3.** (continued)

Second index	Weight by PSM	Entropy weight	Combination weight	Average Score
A22	0.0031	0.0175	0.0092	4.61
A23	0.0872	0.0158	0.0464	4.55
A24	0.0872	0.0171	0.0483	4.57
A25	0.0351	0.0163	0.0299	4.25
A26	0.0351	0.0156	0.0292	4.58
A27	0.0009	0.0220	0.0056	4.50
A28	0.0017	0.0197	0.0072	4.38
A29	0.0644	0.0127	0.0357	4.57
A30	0.0017	0.0206	0.0074	4.58
A31	0.0476	0.0190	0.0376	4.55
A32	0.0141	0.0168	0.0193	4.55
A33	0.0023	0.0296	0.0103	4.44
A34	0.0105	0.0178	0.0170	4.40
A35	0.0013	0.0197	0.0062	4.45
A36	0.0259	0.0162	0.0256	4.09
A37	0.0192	0.0318	0.0309	4.02
A38	0.0004	0.0184	0.0033	4.42
A39	0.0192	0.0173	0.0227	4.31
A40	0.0023	0.0134	0.0069	4.46
A41	0.0042	0.0197	0.0114	4.89
A42	0.0192	0.0201	0.0245	4.89
A43	0.0476	0.0161	0.0346	3.99
A44	0.0057	0.0164	0.0121	3.84
A45	0.0259	0.0148	0.0245	4.54
A46	0.0644	0.0154	0.0394	4.55
A47	0.0002	0.0080	0.0014	4.33
A48	0.0002	0.0113	0.0019	4.18
A49	0.0105	0.0315	0.0227	2.33
A50	0.0141	0.0160	0.0188	2.66
B1	0.0599	0.1154	0.0879	4.38
B2	0.0812	0.0996	0.0950	4.44
B3	0.1488	0.0915	0.1234	4.39

(continued)

**Table 3.** (continued)

Second index	Weight by PSM	Entropy weight	Combination weight	Average Score
B4	0.1488	0.1175	0.1398	4.30
B5	0.1099	0.1156	0.1192	3.94
B6	0.0241	0.1311	0.0595	4.50
B7	0.1488	0.0861	0.1196	4.55
B8	0.2015	0.0899	0.1423	4.51
B9	0.0443	0.0638	0.0562	3.25
B10	0.0327	0.0893	0.0571	3.49
C1	0.0997	0.2598	0.1652	4.32
C2	0.3351	0.3141	0.3330	4.57
C3	0.1828	0.1378	0.1629	4.45
C4	0.2475	0.1562	0.2018	2.29
C5	0.1350	0.1321	0.1371	2.74
D1	0.0630	0.2262	0.1284	4.47
D2	0.2868	0.1488	0.2221	4.45
D3	0.1565	0.0931	0.1298	4.48
D4	0.0853	0.2195	0.1472	4.44
D5	0.1156	0.0866	0.1075	4.83
D6	0.0344	0.0491	0.0442	3.85
D7	0.0466	0.0762	0.0640	4.53
D8	0.2118	0.1005	0.1568	4.35
E1	0.1730	0.1518	0.1657	4.40
E2	0.1278	0.1345	0.1340	4.57
E3	0.0697	0.1640	0.1093	4.52
E4	0.2343	0.1258	0.1755	4.83
E5	0.1730	0.1566	0.1683	3.88
E6	0.0944	0.1476	0.1207	4.50
E7	0.1278	0.1198	0.1265	4.39

## 4 Conclusions

From the perspective of passenger perception, this paper constructs a relatively comprehensive and systematic service quality evaluation system based on SERVQUAL model, which includes 80 indexes covering all service links. Then, this paper uses PSM-entropy weight method to analyze the contribution of 80 indexes to airport service quality. Finally, this paper evaluates the airport service quality. Specifically, the overall service quality of

the airport is at the four-star service level (84.40 point), in which tangibility, assurance and empathy are at the five-star service level, while reliability and responsiveness are at the four-star service level. All in all, the service quality evaluation system can comprehensively, systematically and effectively reflect the current situation of airport service quality, identify the deficiencies of various services, and provide important practical guidance for optimizing service content and improving service quality.

**Acknowledgements.** This research was financially supported by Key technologies on Civil aviation service quality monitoring and risk early warning based on big data (Grant No.: x222060302544).

## References

1. Barakat, H., Yeniterzi, R., Martín-Domingo, L.: Applying deep learning models to Twitter data to detect airport service quality. *J. Air Transp. Manage.* **91**, 1–8 (2020)
2. Zhou, H., Li, Y., Gu, Y.: Research on evaluation of airport service quality based on improved AHP and Topsis methods. *Inst. Civil Eng.* **174**(4), 267–278 (2021)
3. Prentice, C., Kadan, M.: The role of airport service quality in airport and destination choice. *J. Retail. Consum. Serv.* **47**, 40–48 (2019)
4. Lee, K., Yu, C.: Assessment of airport service quality: a complementary approach to measure perceived service quality based on Google reviews. *J. Air Transp. Manage.* **71**, 28–44 (2018)
5. Adacher, L., Flamini, M.: Optimization of airport check-in service quality focused on operational costs and passengers' satisfaction. *PLoS ONE* **16**(8), 1–20 (2021)
6. Atalay, K., Atalay, B., Feride, B.: FIPIA with information entropy: a new hybrid method to assess airline service quality. *J. Air Transp. Manage.* **76**, 67–77 (2019)
7. Shah, F., Syed, Z., Imam, A., Raza, A.: The impact of airline service quality on passengers' behavioral intentions using passenger satisfaction as a mediator. *J. Air Transp. Manage.* **85**, 2–12 (2020)
8. Shiwakoti, N., Jiang, H., Nguyen, A.D.: Passengers' perception of safety and its relationship with demographics, service quality, satisfaction and loyalty in airlines sector - a case study of Vietnam to Australia route. *Transp. Policy* **124**, 194–202 (2022)
9. Büyüközkan, G., Havle, C., Feyzioğlu, O., Göçer, F., Kahraman, C.: A combined group decision making based IFCM and SERVQUAL approach for strategic analysis of airline service quality. *J. Intell. Fuzzy Syst.* **38**(1), 859–872 (2020)
10. Chou, C., Liu, L., Huang, S., Yih, J., Han, T.: An evaluation of airline service quality using the fuzzy weighted SERVQUAL method. *Appl. Soft Comput. J.* **11**(2), 2117–2128 (2011)
11. Rezaei, J., Kothadiya, O., Tavasszy, L., Kroesen, M.: Quality assessment of airline baggage handling systems using SERVQUAL and BWM. *Tour. Manage.* **66**, 85–93 (2018)



# Research on Visualization of Student Management System Based on Distributed Storage Technology

Fengxia Zhang, Xiaoguang Chen<sup>(✉)</sup>, and Ningning Tong

Dalian Ocean University, Dalian, Liaoning, China  
jackxiaoguang@hotmail.com

**Abstract.** The rapid development of network technology, the rapid expansion of mobile terminals, the massive growth of industry data even the growing student data of colleges will need to be organized systematically. At the same time, the current student management in colleges is lack of uniform, targeted and effective information management methods, the paper aims to help managers adopt efficient management strategies and optimize Management process by improving management through the use of big data technology and innovative ideas so as to build a data-based, informatized and networked management model for student management.

**Keywords:** Colleges · Student Management · Big Data · Data Portrait · Hadoop

## 1 Introduction

In the information age, students' teaching and daily life style have changed drastically, and the requirements of student management and talent training in colleges are also constantly updated. For a long time, students' education management mainly relies on personal experience and written direct management. The in-depth development and application of big data technology can obviously provide help for student management in colleges. A large amount of data generated by students in daily study and life will provide reliable support for student management through certain technical processing and mining [1].

## 2 The Application Significance of Big Data in Student Management

Through the analysis of students' huge amounts of data by using big data technology, it can provide managers with a new means. Compared to traditional management mode, it can be more comprehensive, more easily and more quickly to collect data relevant to the students, and the comprehensive analysis of students' learning, life behavior, emotional state even psychological thoughts can be mastered. Managers can grasp the current situation and tendency of students in time, discover and intervene in abnormal

behaviors earlier so as to guarantee students to complete their studies better, to prevent the occurrence of students' risks and reduce the influence of other factors on school study and life [2]. On the other hand, it improves the management level of managers. So the measures like improving management efficiency, improving the scientificity of students' management, improving the accuracy of management, reducing unnecessary waste of energy and achieving management goals in a more targeted manners have important theoretical significance for improving China's modern student management system in colleges and the establishment of relevant student management theories in colleges [3].

### **3 Problems and Current Situation**

Compared with foreign universities that introduced big data technology earlier, the application of big data technology in student management in colleges in China is still in the initial stage, and the concept of student management is too conservative and lack of innovation. Firstly, the lack of scientific and objective management work often force us to rely on experience to judge things which leads to decisions without basis [4]. Secondly, the management work is lack of forward-looking and preventive goals, the traditional management mode can not timely detect the potential behavior of students, can not advance the effective intervention of bad behavior. Thirdly, the management lacks comprehensiveness and timeliness. For example, due to the influence of subjective impression and emotion, managers have veto power in the evaluation process of students, which results in one-sided selection method and incomplete basis, and cannot conduct comprehensive objective evaluation from individual morality, intelligence, body, beauty and labor. Finally, the lag in governance and lack of talent are serious. Talents are the strategic resources in data era, compared with Internet companies, the salary colleges can provide is restricted, which is unable to attract top talent in the field of big data, in addition, the massive information and irregular character of the big data also increase the difficulty in the application of the big data technology [5].

## **4 The Construction and Innovative Application of Big Data in Student Management**

### **4.1 Establish a Process System for Student Management**

Based on the experience in student management work in colleges, the paper summarizes the main management services like: learning management, practice management, assessment management, recommend grants, archives management, life management, psychological health management and employment management, etc. [6], as shown in Fig. 1:

### **4.2 Distributed Storage and Processing of Data Related to Student Management Work Based on Big Data**

Students' data in this project mainly comes from aspects like students' learning, life and management, which includes student management system, report cards of students in



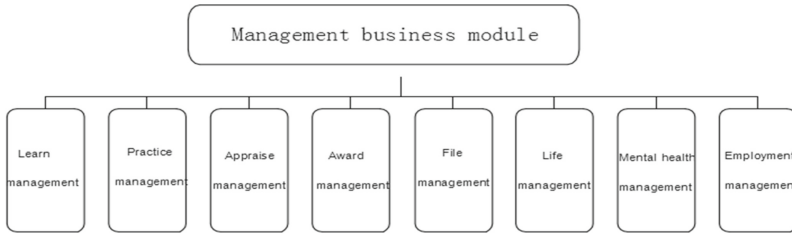


Fig. 1. Student management business module

the second report, educational system, library system, evaluation on students’ mental health, students’ card consumption system and Internet behavior records, etc. they were collected and stored in HDFS mainly by adopting the technologies like ETL, Flume. The Hadoop is a distributed system infrastructure for storing, managing, and analyzing massive amounts of data. HDFS is one of the core modules of Hadoop distributed file system. It has features as followed: high fault tolerance, saving function of multiple copies of data, automatic recovery after being lost, suitable for batch processing of data, millions of files processing scale, the support of a large number of nodes; And it can be built on common devices. It accesses the data of application program by streaming, which greatly improves the data throughput of the whole system, and it is very suitable for the application of large data set. It can analyze, summarize or combine the data in HDFS according to certain requirements and finally output the results to external systems through filtering, conversion and aggregation based on Map Reduce System [7].

### 4.3 Data analysis and Visualization of Student Management Work Based on Big Data

Based on the previous extraction, cleaning, transformation, loading, storage and processing of data, the analysis of these data visualization technology will be used through the comparison of different element value, distribution, composition and relations, and the further use of computer graphics and image processing technology will be taken to convert them to a visual structure. It is convenient for decision makers to understand the data and its potential rules[8], trends and other information from the data for subsequent management applications through the graphic image display screen and image data abstraction, as shown in Fig. 2:

### 4.4 Construct Multidimensional Data Model and Digital Portrait of Students

Digital images of the students and the data model is one of the big data applications, according to data produced by students in study and campus life, on the technical level, it depends on the data, the model and algorithm through the multidimensional data analysis, data label coordinates system of student behavior, which generate student personal image “coordinates board” will be realized [9]. Managers can quickly understand the characteristics and differences of each student according to the image of “coordinate board” with labels, and then carry out individualized teaching, personalized management and accurate teaching. For example, as shown in Fig. 3,

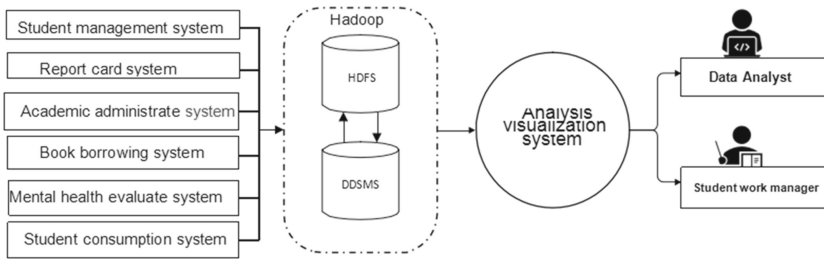


Fig. 2. Date analysis and visualization

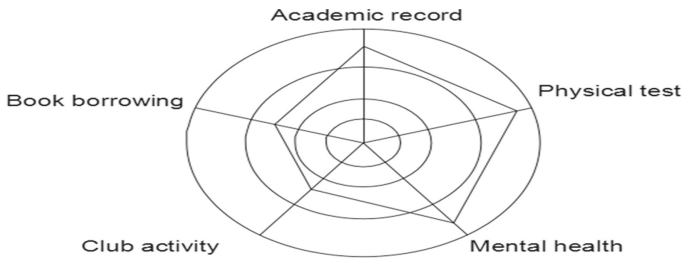


Fig. 3. Digital portrait of a student

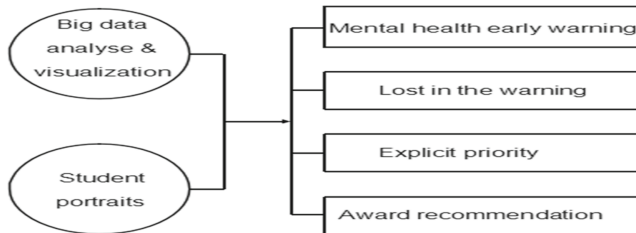


Fig. 4. Implementation of the functions

The digital portrait of a student can intuitively understand the student’s academic performance, physical fitness test, mental health, community activities and book borrowing, as well as his weaknesses, strengths and other potential tendencies.

**4.5 Realize the Functions of Early Warning of Students’ Mental Health, Early Warning of Loss of Contact, Evaluation and Recommendation of Awards and Grants, etc**

The extraction from the student campus life and learning of each part of original data can implement the functions like students’ mental health warning, lost in early warning and merit evaluation and grant recommendation aided by big data [10] after cleaning, conversion, analyzing by using big data and visualization technology even to establish students’ personal digital portrait, as shown in Fig. 4. For example, student management

workers can focus on students with intelligent warning of big data to save most of their energy and take corresponding measures to reduce the occurrence of accidents. With the help of big data analysis results, it can be used as an auxiliary condition for the evaluation of excellence and the recommendation of grant to reduce the influence of personal impression and increase the degree of conviction and fairness [11].

## 5 Conclusions

The innovation and application of big data technology in the student management work mainly depend on the data analysis, data processing, the mining of potential information and the regularity of the hidden data to help managers have a more comprehensive, more accurate understanding of each student, in time and avoid the happening of accidents so as to take more scientific, effective and practical management measures in the implementation of student work, as well to help students learn better and grow faster in the campus, thus achieving high quality education work.

## References

1. Anderson, H.H.: The measurement of domination and of socially integrative behavior in teachers' contacts with children. *Child Dev.* **10**(2), 73–89 (1939)
2. Flanders, N.: Intent, action and feedback: a preparation for teaching. *J. Teach. Educ.* **14**(3), 251–260 (1963)
3. Uskov, V.L., et al.: *Smart Universities: Concepts, Systems, and Technologies*. Springer Press, New York (2018)
4. Uskov, V.L., et al.: Building smart learning analytics system for smart university. In: Uskov, V.L., Howlett, R.J., Jain, L.C. (eds.) *Smart Education and e-Learning 2017*, pp. 191–204. Springer International Publishing, Cham (2018). [https://doi.org/10.1007/978-3-319-59451-4\\_19](https://doi.org/10.1007/978-3-319-59451-4_19)
5. Angeli, C., Howard, S., et al.: Data mining in educational technology classroom research: can it make a contribution. *Comput. Educ.* **113**(5), 226–242 (2017)
6. Holstein, K., Hong, G., et al.: The classroom as a dashboard: co-designing wearable cognitive augmentation for K-12 teachers. AbelardoPardo. In: *Proceedings of the 8th International Conference on Learning Analysis and Knowledge*. ACM Press, New York, pp. 79–88 (2018)
7. Hilles, M.M., Naser, S.S.A.: Knowledge-based intelligent tutoring system for teaching mongo database. *Eur. Acad. Res.* **4**(10), 8783–8794 (2017)
8. Harrington, M., Dennis, S., *Input-driven language learning*. *Studies in Second Language Acquisition*, vol. 24 (2004)
9. Ellis, N.: Frequency - based accounts of SLA. In: Gass, S., Mackey, A. (eds.) *Routledge Handbook of Second Language Acquisition*, p. 196, p. 201. Routledge, London & New York (2012)
10. Zhang, Z.: *A Preliminary Study on Traditional Chinese Education*, p. 1. Shanghai Education Press (1962)
11. Krashen, S.: *Principles and Practice in Second Language Acquisition*. Pergamon Press, Oxford (1982)



# Banking Fintech Capability Evaluation System Based on Data Mining

Jie Fan<sup>(✉)</sup>

Hunan Federation of Rural Credit Cooperatives, Changsha, Hunan, China  
48228272@qq.com

**Abstract.** With the declining macro-economy, the acceleration of financial disintermediation, the prosperity and development of Internet finance, and the tightening of the government's monetary policy, the interest gap between my country's Commercial Bank (CB) has gradually narrowed. Based on this, my country's CBs have begun to transform, committed to aggregating financial resources, organically integrating with science and technology, developing and innovating fintech products and services, and embarking on a distinctive path of fintech development. Therefore, this paper proposes data mining (DM) technology, builds a bank financial technology capability evaluation system, evaluates the financial technology innovation capability of my country's CBs, and then promotes the steady development of my country's CBs; briefly analyzes the impact of financial technology on banks and CBs. Fintech capability; built a Banking Fintech (BF) capability evaluation system, and conducted a Fintech capability evaluation test on my country's listed CBs. The test results verified the accuracy and feasibility of the BF capability evaluation system proposed in this paper.

**Keywords:** Data Mining Technology · Banking Finance · Financial Technology · Capability Assessment System

## 1 Introduction

The environment for the evaluation of financial technology innovation capabilities of CBs is immature. Compared with economically developed countries, China's financial regulation is relatively strict, which greatly affects the innovation and initiative of financial technology business. In addition, financial users do not have a scientific outlook on finance, and the society has not yet established a sound social credit system, which has repeatedly hindered the innovation of CBs' fintech business. Therefore, in order to successfully innovate financial technology, it is necessary to reduce financial controls and improve financial markets. To sum up, the key factors that lead to the evaluation of CBs' financial technology innovation capabilities are that the evaluation indicators for CBs' financial technology innovation capabilities have not been established, the evaluation methods adopted are unscientific, there is a lack of effective evaluation feedback mechanisms, and the evaluation system is not perfect. Through the establishment of a scientific and effective evaluation index system for the financial technology innovation

capability of CBs, it is possible to accurately evaluate and analyze the financial technology innovation capability of CBs, so as to solve the problem of improving the financial technology innovation capability of CBs. Therefore, based on DM technology, this paper constructs an evaluation system for bank financial technology innovation capability.

Compared with the developed western countries, the maturity of China's financial market is not high, and there are few research results on financial technology. Collecting and sorting out the existing research results abroad, it can be seen that experts and scholars pay more attention to the perspective of economy, technology or finance, in-depth analysis of the role of finance in promoting the development of science and technology, and seldom pay attention to and discuss the financial technology itself; domestic experts and scholars start from multiple perspectives. Comprehensively discussing financial technology itself, the research scope is relatively broad, so the results are more fruitful [1]. However, compared with foreign achievements, my country's financial technology research still has a large space. Foreign scientific and technological research and financial research are far earlier than China, especially in the developed Western economies dominated by market economy, their capital markets are more mature and sound, and the direct financing channels for small and medium-sized enterprises are more abundant, which is worthy of reference by relevant Chinese researchers. Learn from [2].

This paper analyzes the development of financial technology in CBs, and evaluates the financial technology capabilities of CBs through principal component analysis, which is innovative to a certain extent. This paper analyzes the status quo of the evaluation of financial technology innovation capabilities of CBs in my country. From the perspective of CBs in my country, this paper proposes DM technology to evaluate the financial technology innovation capabilities of CBs, and proposes an evaluation of banks' financial technology innovation capabilities based on DM. System; Combined with the actual situation of the development of financial technology in my country's CBs, build an empirical analysis model of DM technology to evaluate the technological innovation capabilities of my country's CBs, and obtain specific evaluation results and data. The research method is scientific and has a certain degree of innovation. The research provides reliable data support [3, 4].

## 2 Analysis of BF Capability

### 2.1 Fintech

The connotation of financial technology is shown in Fig. 1. Fintech is the support and application of Fintech-related technologies. The continuous expansion of financial technology will inevitably reconstruct the traditional service model and system structure of banks, and banks will become integrated financial service platforms. At present, some forward-looking bankers have made it clear that in the future, banks will gradually become financial technology companies [5].

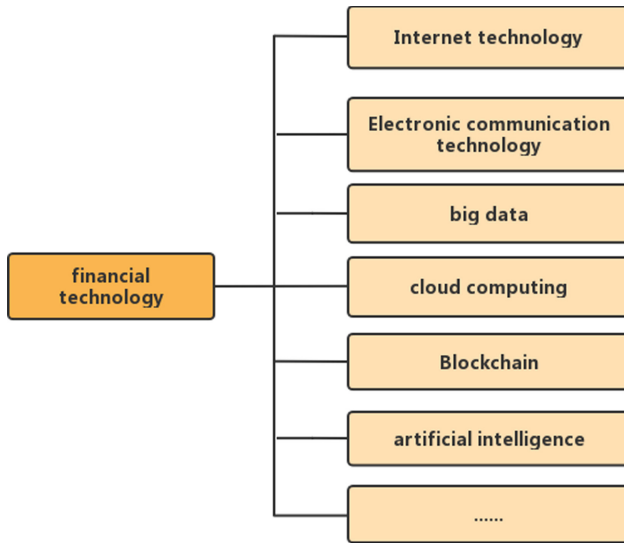


Fig. 1. Connotation of financial science and technology

## 2.2 Analysis of the Influence of Fintech on Banks

### (1) The positive impact of financial technology development on banks

First, financial technology can promote the intelligent construction of bank data. Banks can use the big data and artificial intelligence technology in financial technology to fully mine and analyze the bank's stock data assets, which can support banks to carry out extensive and comprehensive data management and services. According to the needs of operation and management, in terms of customer tiered marketing, outlet performance assessment, retail performance analysis, product effectiveness prediction, operational efficiency optimization, product value analysis, etc., to guide banking business decisions, and promote the construction of bank data intelligence. Second, financial technology will improve the level of banking services [6, 7]. With the maturity of biometric models and computing technology in financial technology, on the one hand, banks can introduce more human-computer interaction scenarios in the channel, such as face recognition and voice recognition in banking business, on the other hand, they can use cloud computing technology to greatly improve The efficiency of business processing makes banking business more efficient and convenient, and greatly improves customer experience and satisfaction. Third, financial technology will promote the continuous innovation of banking business products. With the maturity of financial technology technology, various banks continue to introduce new products and innovative products emerge one after another [8].

### (2) The negative impact of financial technology development on CBs

First, the development of fintech will open up financial markets and increase competition for banks. Fintech has prompted many Internet companies that did not have financial licenses to set foot in the financial field to carry out businesses such as payment and credit. The increase in financial and financial participants has greatly increased the difficulty of

traditional banks' competition. Banks not only need to compete with their peers, but also need Internet companies and financial technology companies compete in the financial field [9].

Second, the development of fintech may reduce the traditional business income of banks. In terms of traditional private financing loans of banks, it mainly involves housing loans, credit card credit business, etc. At present, many Internet companies have carried out various forms of personal loans, such as Huabei, Jibei, etc. Various Internet credit models emerge one after another. Financing through banks has been greatly reduced; in the field of payment, the WeChat payment launched by major Internet companies and Yu'eobao have made the traditional UnionPay payment retreat; Not only is the purchase convenient and there is no initial purchase amount limit required by regulatory requirements, it has received high praise from customers, and the current operating income of banks is shrinking year by year under these business scenarios [10, 11].

Third, fintech makes it difficult for some banks to keep up with the rhythm of the competitive market. With the tightening of regulatory policies, some banks insist on prudent business operations. Banking business products are generally customized and developed using the traditional industrialized waterfall model. The time period is long, the iteration is slow, and under the background of the great development of financial technology, the financial market competition is becoming increasingly fierce., market opportunities are fleeting, bank product development is generally risk averse, and it is difficult to keep up with the current rhythm of financial technology development. In general, the development of financial technology is a double-edged sword for banks. How to make good use of financial technology, stand on the cusp of financial technology development, and promote the development of banking business is worth thinking about by all CBs [12].

### 2.3 Analysis of the Fintech Capability of CBs

(1) The concept of technology management needs to be changed

At present, the concept of domestic financial technology is constantly fermenting, and information technology has gradually moved from the original behind-the-scenes supporter to the "center stage", to the market, and directly creates value by serving more customers and creating new products. Banking technology is gradually changing from the traditional operation service orientation, from the role of auxiliary business processing to the role of business value creator. Only by making full use of and exerting the power of financial technology can promote the real productivity of financial technology, realize technology-led business, and promote business.develop.

(2) The response speed of technology needs to be improved

In recent years, as the state vigorously advocates innovation and encourages and supports the rapid development of small and micro enterprises, various business needs of banks are characterized by high concurrency, periodicity, and regionality. The demand for "fast and good", "fast" refers to the need to match the speed of scientific and technological response with the development of business innovation, effectively improve and strengthen the construction of scientific and technological agility, "good" refers to on the basis of fast, New products or new business models must operate safely, stably

and efficiently. Therefore, the bank's traditional waterfall model software development method can no longer adapt to the new form of the Internet era. Banks need to introduce an agile development model to effectively and quickly carry out iterations to improve their technological responsiveness.

### 3 Fintech DM Technology

With the continuous entry of Internet companies into the financial industry, whether it is Tencent's "Qianqiantong", Alibaba's "Yuebao", as well as the mushrooming online loan companies, coupled with the continuous upgrading of financial technology applications, the integration of finance and technology Continuous integration is changing the model of traditional banking financial services every day. Banking technology must actively innovate in the financial field, and use innovative technologies such as artificial intelligence business scenarios, cloud computing distribution technology, mobile payment mode, and massive big DM applications to play a leading role in business. It will transform the banking business from the traditional offline physical bank branch field to an innovative sunrise industry that supports the online electronicization of the Internet. To this end, this paper proposes DM technology to evaluate and analyze BF.

Suppose there is a constantly reacting change as  $h_i^*$  fintech, which represents the possibility of an event, and its range is  $(-,)fintech$ . When the variable is greater than 0, it means that the event has occurred; the value of the variable is less than or equal to 0, the event does not occur. Which is:

$$\begin{cases} h_i = 1, h_i^* > 0 \\ h_i = 0, h_i^* \leq 0 \end{cases} \quad (1)$$

Here,  $h_i$  is the observed real response variable.  $h_i = 1$  means the event happened;  $h_i = 0$  means the event did not happen. It is assumed that there is a linear relationship between the explained variable and the explanatory variable  $x_i$ , that is:

$$h_i^* = \gamma + \eta x_i + \delta_i \quad (2)$$

Assuming that the random error term  $\delta_i$  obeys the logistic distribution, the logistic function distribution has symmetry, so the probability cumulative distribution function can be obtained:

$$g(h_i = 1|x_i) = g[\delta_i \leq (\gamma + \eta x_i)] = k(\gamma + \eta x_i) \quad (3)$$

For multiple independent variables, assuming there are  $t$  independent variables, the logistic model can be extended to a multivariate logistic model:

$$g = \frac{1}{1 + e^{-(\eta_0 + \eta_1 x_1 + \dots + \eta_n x_n)}} \quad (4)$$

Among them,  $g$  is the probability when the independent variable takes  $(x_1, x_2, \dots, x_n)$ .



## 4 Analysis of BF Capability Evaluation System Based on DM

### 4.1 Evaluation of the Fintech Innovation Capability of CBs

Based on DM technology, this paper introduces SPSS22.0 software to further process the data, and obtains Table 1. According to the tabular data, it can be seen that when the principal components are different, the ranking of listed banks is different. Also, the individual rankings are not identical to the overall rankings. Therefore, it is necessary to identify the evaluation object to clarify the strength of the bank’s innovative financial technology.

**Table 1.** Factor load after rotation

index	1	2	3
Ratio of net income from handling fees and commissions to operating income x1	0.841		
Growth rate of net income from handling fees and commissions x2	0.812		
Per capita net profit x3	0.842		
Bachelor degree or above ratio X4		0.811	
E-banking substitution rate X5		0.804	
Three year average annual growth rate of credit loan x6		0.786	
Non performing loan ratio x7			0.842
Liquidity ratio x8			0.792

Through the analysis of principal component analysis, there are 3 influencing factors whose characteristic root is greater than 1, and the cumulative variance contribution rate is 84.1%. The validity of the data is high, and principal component analysis can be carried out. The above factor loadings are all higher than 0.5, and they are all reserved. After sorting out the three common factors, the results are as follows:

**Table 2.** Common factor sorting result data table

index	Factor load	Common factor	Variance contribution rate	Cumulative variance contribution rate
Ratio of net income from handling fees and commissions to operating income x1	0.841	K1	46.23	46.23

*(continued)*

**Table 2.** (continued)

index	Factor load	Common factor	Variance contribution rate	Cumulative variance contribution rate
Growth rate of net income from handling fees and commissions x2	0.812			
Per capita net profit x3	0.842			
Bachelor degree or above ratio X4	0.811			
E-banking substitution rate X5	0.804	K2	23.22	69.45
Three year average annual growth rate of credit loan x6	0.786			
Non performing loan ratio x7	0.842			
Liquidity ratio x8	0.792	K3	14.65	84.10

It can be seen from the above Table 2 that the final explanatory degree of the three principal components extracted is 84.1%, and three principal components are extracted. Among them, K1 is the innovation factor of fintech business of CBs, including three secondary indicators: the ratio of net fee and commission income to operating income X1, the growth rate of net fee and commission income X2, and the per capita net profit X3; K2 is the commercial BF management innovation factor, including three secondary indicators: the ratio of undergraduate education and above, X4, the replacement rate of e-banking business X5, and the three-year average annual growth rate of credit loans X6; K3 is the financial technology risk innovation factor of CBs, including non-performing loans. There are two secondary indicators, namely, the liquidity ratio X7 and the liquidity ratio X8.

This paper randomly selects 10 listed banks in my country, and evaluates the financial technology capabilities of their main components, as shown in Fig. 2.

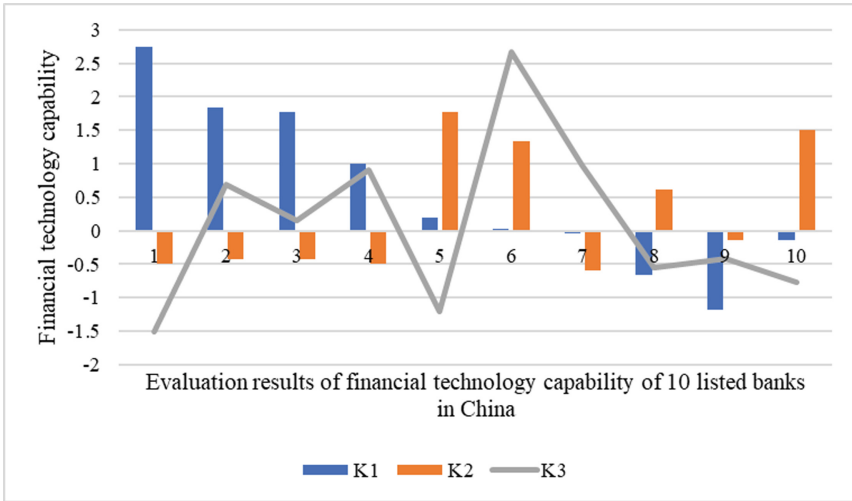


Fig. 2. Evaluation results of financial technology capability of 10 listed banks in China

#### 4.2 Evaluation Results of Fintech Innovation Capability of CBs in My Country

Assessing the financial technology innovation capability of CBs from the principal component K1: K1 focuses on presenting factors that are likely to have an important impact on the financial operation model, so that evaluators can understand the financial innovation status and technological innovation status of the bank. Generally speaking, the higher the K1 value, the higher the level of bank financial technology innovation. In this dimension, state-owned listed banks rank among the top four, and ICBC, ABC, and CCB are relatively stronger than other listed banks. The data shows that the business processing and service fee revenue of the above three major banks is as high as 391.8 billion yuan, which is 68% higher than the related revenue of other listed banks. Among them, cash management activities have the highest revenue. First, with regard to project launch, ICBC pays more attention to structural financing, and advocates improving syndicated allocation and equity financing system, so as to steadily increase related income; Private funds; finally, in terms of asset management, ICBC has gradually expanded its investment funds, trusts, and other bank-enterprise networks to effectively improve customer levels.

From the perspective of technology application, state-owned CBs attach more importance to business innovation: in addition, China Merchants Bank is relatively obvious, and its three-year average annual growth rate of credit loans is 35.47%. Assessing the Fintech Innovation Capability of CBs from the Principal Component K2 From the perspective of Fintech operation and management, state-owned CBs have the largest proportion of the K2 principal component. However, in terms of management model innovation, joint-stock CBs are relatively more innovative than state-owned banks in terms of their operational and management innovation. The reason is that the per capita net profit of major state-owned banks and the total value of employees with a bachelor's degree or

above are higher, and joint-stock CBs can only innovate their management models to gain a larger market share.

Assessing the financial technology innovation capability of CBs from the principal component K3: According to the principal component K3, it can be seen that it can present the financial technology innovation and risk management capabilities of CBs to a certain extent. Ranking the two main components, it is not difficult to find that compared with other listed CBs, China Merchants Bank has a more complete risk management system. Combined with the survey data, it can be seen that the total proportion of non-performing loans of China Merchants Bank is 0.93 percentage points, and the industry average is 92%. The core is still the liquidity ratio. Among all listed CBs, China Merchants Bank has the strongest capital liquidity, with a ratio as high as 66.94%, which is 19.39% higher than the average liquidity ratio of the same industry.

Comprehensive F1, F2, and F3 evaluation of CBs' financial technology innovation capabilities: Comprehensive analysis of financial technology innovation capabilities shows that the Industrial and CB of China ranks first, followed by China Merchants Bank, and China Construction Bank ranked third. Ningbo CB attaches great importance to financial technology innovation, but its innovation ability is average and its performance is not outstanding. It is the bank with the lowest level of financial technology innovation among all CBs. According to the previous research results, it can be seen that the financial technology innovation capability of joint-stock banks and city CBs is relatively weaker than that of state-owned banks.

## 5 Conclusions

This paper conducts research and analysis based on the DM BF capability evaluation system, and sorts out the research results at home and abroad with the key words of innovation capability and fintech. The definition also needs to be further clarified. At present, as domestic and foreign experts and scholars mainly explore the concept of financial innovation from different perspectives, the innovation capability of financial technology is mostly regarded as the operational efficiency of the financial system. Overall, the research on FinTech still needs to be further developed. Due to the limited research time and professional ability of the author, it is impossible to thoroughly explore the issue of financial technology innovation in an all-round way. To sum up, this paper believes that the research results related to financial technology innovation capabilities still need to be improved: in the future, it will closely follow the development trend of financial deficient and national financial strategy, and deeply explore the financial technology innovation capabilities of CBs. Provide reference for deepening its financial technology service capabilities.

## References

1. Muhammadin, A., Ramli, R., Ridjal, S., et al.: Effects of dynamic capability and marketing strategy on the organizational performance of the banking sector in Makassar Indonesia. *Espacios* **41**(24), 26–41 (2020)

2. Salam, M.A., Saha, T., Rahman, M.H., et al.: Challenges to mobile banking adaptation in COVID-19 pandemic. *J. Bus. Manage. Sci.* **9**(3), 101–113 (2021)
3. Colgan, C.: How remote deposit capture produces productive data for customers. *ABA Bank. J.* **111**(3), 26–27 (2019)
4. Luo, G., Li, W., Peng, Y.: Overview of intelligent online banking system base on HERCULES architecture. *IEEE Access* **99**, 1 (2020)
5. Nugroho, K.E., et al.: Dynamic capability and disruptive innovation within perspectives of industry 4.0, research result and innovation on people's prosperity. *IOP Conf. Ser. Mater. Sci. Eng.* **508**(1), 12101–12101 (2019)
6. Okoli, T.T., Tewari, D.D.: Does the adoption process of financial technology in Africa follow an inverted U-shaped hypothesis? an evaluation of rogers diffusion of innovation theory. *Asian Acad. Manage. J. Account. Finan.* **17**(1), 281–305 (2021)
7. Yee, S., Fadell, M.: System-specific evaluation of the dual flip angle MRI technique for quantitative T1 measurement. *Med. Phys.* **48**(6), 2790–2799 (2021)
8. Onge, P., Temme, L., Mcatee, A., et al.: Evaluation of the Commercial, Off-the-Shelf (COTS) king-devick eye tracking system. *Mil. Med.* **184**(Suppl 1), 571–578 (2019)
9. Kalamkar, R., Wadher, S.: Formulation and pharmacokinetic evaluation of phosal based zal-toprofen solid self-nanoemulsifying drug delivery system. *Pharmaceut. Nanotechnol.* **7**(4), 328–338 (2019)
10. Setiawan, I.: The role of islamic banking in the development of economic sectors in Indonesia. *Int. J. Appl. Bus. Res.* **1**(2), 88–99 (2019)
11. Eko, E.U., Adebisi, A.W., Moses, E.J.: Evaluation of forensic accounting techniques in fraud prevention/detection in the banking sector in Nigeria. *Int. J. Finan. Account.* **9**(3), 56–66 (2020)
12. Mustafa, O.: Does economic growth responses to changes in financing policies? signals from sudan economy (1997–2018). *Arch. Bus. Res.* **8**(7), 272–287 (2020)



# Technology Financial Service Platform Based on Data Mining Algorithm

Shuyu Hu<sup>(✉)</sup>

School of Economics and Management, Hunan Open University, Changsha, China  
778666080@qq.com

**Abstract.** The establishment of a technology financial (TF) platform has played an important role in providing investment consulting, information services, match-making, technical guidance, and policy consulting. However, by analyzing the current situation of the development of (TDO) TF in County A, this paper finds that although the number of patent innovations in County A has shown an increasing trend in recent years, County A lacks scientific and technological talent resources, and the degree of development of TF does not match the strength of TF it possesses. In order to reverse the serious lag in TDO TF in County A and improve the level of technological innovation in County A, it is necessary to establish a science and TF service platform (SP) that integrates resources and information from all parties and can serve various types of science and TF supply and demand entities. In this paper, data mining technology is used to mine information resources related to TF, establish a sub-platform of TF information service, and use data mining algorithms to establish a credit rating model to evaluate the level of financial products or rate platform participants.

**Keywords:** Data Mining Algorithm · Technological Finance · Information Resource · Technological Innovation

## 1 Introduction

With the increasingly prominent role of science and technology in promoting economic growth, a series of development models centered on promoting TDO science and technology have become an important way to promote regional economic development. As a new model that brings together the resources of all parties and promotes the efficient development of technological financial activities, the TF SP has been listed as the focus of the provinces to promote TDO scientific and technological finance.

The research on TF SP has achieved good research results. For example, a scholar pointed out the way in which technological innovation and financial innovation interact. The emergence of new technologies in the early stage often brings huge technological changes, and technological changes will lead to huge economic uncertainty and high profits. In order to obtain high profits, venture capitalists continue to explore promising technical fields and realize the deep integration of financial capital and technical capital. The accompanying result is that technical capital and financial capital grow rapidly and

prosper together [1]. A scholar finds substantive laws from the development process. Technology and finance must be integrated with each other. This integration is large-scale and unified, and requires consistent requirements for conditions, personnel, and funds to ensure the coordinated development of technology and finance [2]. Some researchers have analyzed the development status of TF SPs through on-the-spot investigations, made a comparative analysis in terms of functions and operating mechanisms, and found a route suitable for the current financial development [3]. Some scholars believe that the basic organizational form of the scientific and technological innovation public SP includes the platform’s organizational structure and service system, and its operation mode is the independent operation of each service subsystem provided by the platform and the division of labor and cooperation between the systems [4]. Although many researchers have proposed the relationship between science and technology and finance or the operation mode of TF, the technology finance SP currently developed still needs to be improved.

This paper first introduces the concept of data mining, and proposes a WFCM algorithm. The main function of this algorithm is cluster analysis, which can be applied to cluster mining of technological financial information resources. It establishes a TF SP, which includes five sub-platforms, providing different service functions for TDO technology finance in County A.

## 2 Data Mining Algorithms

The essence of data mining is to search for hidden information in data through innovative algorithms to discover the laws behind phenomena, and finally apply them to practice to improve the efficiency of solving problems [5]. Today’s data mining is closely integrated with computer technology, and methods such as information technology, machine learning technology, mathematical statistics technology, and artificial intelligence are used to achieve this goal.

Incremental kernel fuzzy clustering algorithm (wFCM) is one of the most commonly used fuzzy clustering algorithms. It is based on the FCM algorithm and assigns different weights according to the importance of the data points to cluster. Process [6]. For a given dataset  $X = \{x_1, x_2, \dots, x_n\}$ , let  $w_i (i = 1, 2, \dots, n)$  be the weights of the data points  $x_i$ . The wFCM algorithm simply selects different initial cluster centers, calculates its membership degree with each data point according to formula (1), and then updates the cluster centers according to formula (2), and calculates again after obtaining a new cluster. The above process is repeated continuously until the convergence criterion is reached [7].

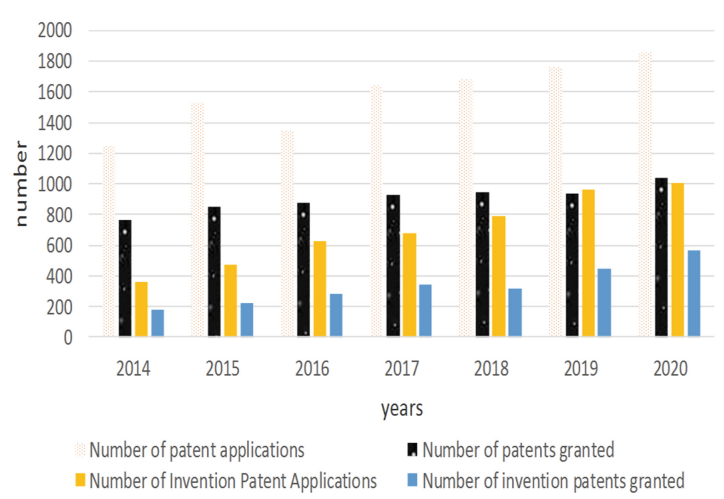
$$p_{ji} = \sum_{h=1}^k \left( \frac{d_{ji}^2}{d_{hi}^2} \right)^{-\frac{1}{r-1}} \tag{1}$$

$$q_j = \frac{\sum_{i=1}^n w_i p_{ji}^m x_i}{\sum_{i=1}^n w_i p_{ji}^m} \tag{2}$$

Among them,  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, k$ ,  $m$  represents the fuzzy coefficient,  $q_j$  represents the cluster center (CC),  $p_{ji}$  represents the degree to which the data point  $x_i$  belongs to the CC  $q_j$ ,  $d_{ji}$  represents the distance from the data point  $x_i$  to the CC  $q_j$ , and its calculation formula can be given by formula (3).

### 3 TDO TF in County A

In terms of patent applications and grants, as shown in Fig. 1, since 2014, the number of patent applications and grants in County A has increased significantly, from 1,245 and 763 in 2014 to 1,855 and 1,043 in 2020, respectively. The number of applications and authorizations for invention patents has increased from 357 and 184 in 2014 to 1,004 and 566 in 2020, showing a good trend of scientific and technological innovation, and the ability of independent innovation has been greatly enhanced.



**Fig. 1.** Number of Patent Applications and Grants in County A

In terms of important scientific and technological achievements, statistics show that from 2014 to 2017, the number of newly registered scientific and technological achievements in County A showed an overall growth trend. In 2018–2019, the number of newly registered achievements was relatively small, but in 2020 it rebounded to 2016. Although it has won the National Science and Technology Award every year, the number is generally declining, as shown in Table 1.



**Table 1.** Scientific and technological achievements in County A

	2014	2015	2016	2017	2018	2019	2020
New registration results	187	235	294	307	215	196	283
National Science and Technology Award	9	7	5	5	4	2	2

**Table 2.** Input of scientific and technological personnel in County A

	2014	2015	2016	2017	2018	2019	2020
Number of people (year/10,000 people)	37.6	41.2	42.8	45.5	46.3	47.2	47.9
growth rate (%)		9.57	3.88	6.30	1.76	1.94	1.48

There is a serious shortage of manpower investment in science and technology activities in County A. According to the statistics in Table 2, there were 428,000 R&D researchers in 2016, an increase of 3.88% compared with the previous year. Although the overall trend of growth, there are still serious shortages of human resources. The problem has led to the slow development of technology finance in County A, which lags behind the average level.

## 4 Construction and Application of TF SP

In order to promote the development level of TF in County A, it is necessary to construct a SP suitable for TDO TF in County A, and apply the functions of the platform to improve County A from all aspects.

### 4.1 Construction of a TF SP in County A

The technology and financial exchange and docking platform improves the efficiency of technology and financial activities by connecting technology companies and investment companies. The investment technology and financial services industry analyzes the operating conditions and financial needs of small and medium-sized enterprises related to technology, and recommends the lowest and most efficient financial tools in combination with products. In order to develop technology-based small and medium-sized enterprises, comprehensive services can be provided to enterprises on the platform through financial service technology. The integrated SP ensures the healthy operation of financial service technology by monitoring and supporting financial service technology, and aims to make money for small and medium-sized enterprises with certain development characteristics of technology [8, 9]. As shown in Fig. 2, a TF SP was established for County A.

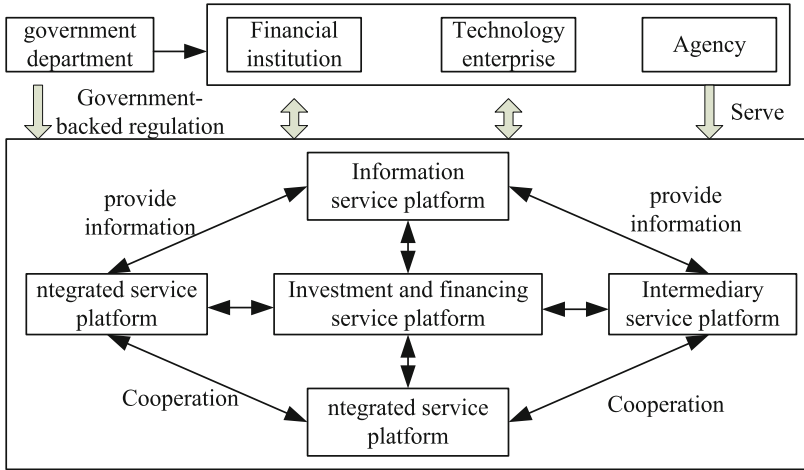


Fig. 2. Structure of the TF SP

(1) Science and TF information SP

Science and TF information SP refers to an information SP that integrates information on various types of science and TF supply and demand entities. Based on various information, the collected information is integrated and reorganized using electronic information technologies such as big data and cloud computing [10].

(2) Technology and finance docking SP

The docking station technology segment is an extension of the TF information platform. Create opportunities for science and technology financing and supply and demand of various services by monitoring and evaluating the need to participate in field technology funding, operational planning and documenting technology implementation and projects, and by organizing collaborative activities to promote high-quality research projects at home and abroad [11, 12].

(3) TF investment and financing SP

According to the financial models of technology companies at different levels, technology investment and financial services combine the characteristics of technology-related financial products such as currency to provide technology companies with the most suitable financing methods [13]. Technology companies and financial institutions that have not participated in the technology-finance docking process can directly carry out investment and financing activities through the technology-finance and financial SP.

(4) TF intermediary SP

Intermediary financial service technology refers to the common management and organizational behavior of intermediary service institutions serving financial service technology, and a platform for providing services to financial demand and supply entities. All technical center services work in the fields of credit reporting agencies and correctional services analysis agencies. The branch system uses financial service technology, and branches located in intermediary financial service technology must strictly abide by business norms and standards. And create an administrative office to oversee the educational activities of the cooperatives and regulate the services of intermediary

companies. Many financial services firms provide financial services for a large number of financial services and needs, learning from and competing with each other.

(5) Integrated SP for TF

The integrated SP of TF has two main functions. One is to establish a supervision service center to supervise the SP of TF and ensure its normal, stable and continuous operation. The second is the auxiliary function, which provides convenience for TDO TF through the establishment of education and training centers, scientific and technological literature and information sharing service centers, and college students' entrepreneurship centers. Its frame diagram is shown in Fig. 3.

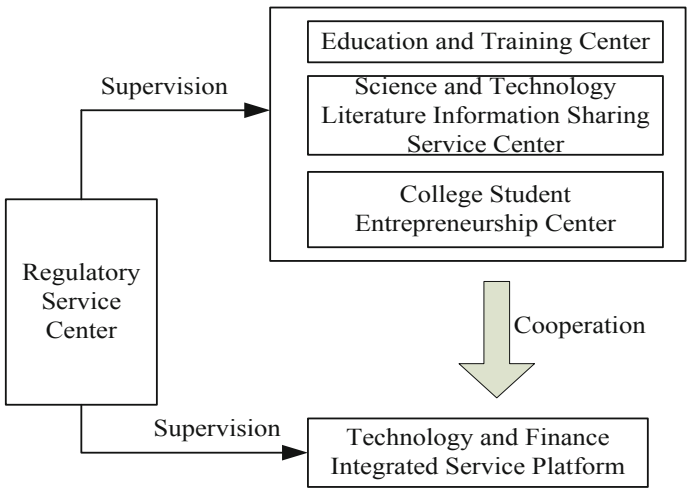


Fig. 3. Integrated SP

4.2 Application of the TF SP in County A

(1) Information resource mining

By building an information resource SP, using technology, and obtaining financial information of krypton gold through big data, cloud computing, etc. At the same time, financial information can be collected to provide re-innovation guidance for technological financial products developed by technology companies, and various types of information can be classified and stored, and then transmitted to platform participants. The government-led TF SP fully integrates the scientific and technological knowledge, business knowledge, financial industry knowledge and other resources owned by public institutions and government social security management to create a public service system where everyone can enjoy resources, and create a page of information resources to facilitate participation to view tech finance information. In this process, the TF SP of County A should improve the ability to obtain and coordinate information, and enhance the ability to support new business activities in the technology industry.

### (2) Investment and financing services

Technology SMEs can apply for government special loans through the SP according to their own situation and capital needs, or obtain emergency financing from financial institutions and financial markets through the platform. In addition, technology-based SMEs can meet their own development needs by recommending other financial institutions, and look for joint ventures with long-term development value. In addition, the financial services industry in County A can issue standards and benchmarks to measure the performance of SMEs, and provide them with management consulting services and advanced TF services, when technology companies in County A can reach different levels of development.

### (3) Credit evaluation

The credit evaluation SP uses data mining algorithms to establish a credit rating model, and through the algorithm, analyzes and evaluates the participants and financial products in the rating system, and provides it to demanders through collaborative office, credit reporting, and system management. Through a unified credit reporting SP, a unified corporate credit evaluation standard is formed, with rating comparability.

### (4) Intermediary agency services

A technical financial service system is required for various personnel such as consulting business and legal business. These types of businesses require the involvement of intermediaries, which are the main intermediaries in the field. Throughout the life cycle of tech companies, but especially in the early stages, they need consulting agencies for guidance on operations and management.

### (5) Authentication service

The certification SP certifies the participating certified companies and displays the relevant information of the certified companies. Through the authentication SP, participants can complete the connection of online authentication services. The platform can increase the external credit of technology companies, standardize the financial process of technology companies, and reduce the financial problems of technology companies, especially startups.

With the advancement in networking and multimedia technologies enables the distribution and sharing of multimedia content widely. In the meantime, piracy becomes increasingly rampant as the customers can easily duplicate and redistribute the received multimedia content to a large audience [3].

## 5 Conclusions

This paper establishes a SP for TDO TF in County A. The SP organically unifies and integrates all participating institutions, which not only satisfies the smooth needs of each subject for communication and interaction channels, but also satisfies the technological needs of each subject for their development. The flow of financial information needs. This improves the resource integration efficiency and information sharing degree of TF to a certain extent, simplifies the process of TF information transmission, promotes flat communication within the TF industry and between enterprises, reduces the cost of maintaining information, and greatly reduces the cost of maintaining information. The market competitiveness of the TF industry has been greatly improved.

## References

1. Zachariadis, M., Hileman, G., Scott, S.V.: Governance and control in distributed ledgers: understanding the challenges facing blockchain technology in financial services. *Inf. Organ.* **29**(2), 105–117 (2019)
2. Tine, B.H.: Fintech: AI powers financial services to improve people's lives. *Comput. Rev.* **60**(4), 182 (2019)
3. Rupa Rani, T.: Open IOT service platform technology with semantic web. *Int. J. Adv. Trends Comput. Sci. Eng.* **9**(1.1 S I), 17–21 (2020). <https://doi.org/10.30534/ijatcse/2020/0491.12020>
4. Lee, J., Ryu, M.H., Lee, D.: A study on the reciprocal relationship between user perception and retailer perception on platform-based mobile payment service. *J. Retail. Consum. Serv.* **48**(MAY), 7–15 (2019)
5. Oh, B.W.: Parallel algorithm for spatial data mining using CUDA. *J. Adv. Inform. Technol. Converg.* **9**(2), 89–97 (2019)
6. Rochd, Y., Hafidi, I.: An Efficient distributed frequent itemset mining algorithm based on spark for big data. *Int. J. Intell. Eng. Syst.* **12**(4), 367–377 (2019)
7. Nandini, G., Rao, N.: Utility frequent patterns mining on large scale data based on apriori mapreduce algorithm. *Int. J. Res.* **3**(8), 19381–19387 (2019)
8. Kouvelis, Panos, Dong, Ling, Turcic, Danko: Advances in supply chain finance and FinTech innovations overview. *Found. Trends@ Technol. Inform. Oper. Manage.* **14**(1–2), 1–4 (2020). <https://doi.org/10.1561/02000000096>
9. Borgogno, O., Colangelo, G.: The data sharing paradox: BigTechs in finance. *Eur. Compet. J.* **16**(2–3), 1–20 (2020)
10. Hendrikse, R., Meeteren, M.V., Bassens, D.: Strategic coupling between finance, technology and the state: cultivating a Fintech ecosystem for incumbent finance. *Environ. Plan. A Econ. Space* **52**(8), 1516–1538 (2020)
11. Sutrisno, S., Fahlefi, D.R., Sarbullah, S.: Penerimaan financial technology berbasis technology acceptance model (Tam). *Jurnal Stie Semarang* **12**(1), 33–46 (2020)
12. Al, A.: Effects of UTAUT 2 model on the use of BCA mobile banking in Indonesia. *Turkish J. Comput. Math. Educ. (TURCOMAT)* **12**(3), 5378–5387 (2021)
13. Anifa, F., Anisa, A., Fadhila, N., et al.: Tingkat Kemudahan dan Manfaat pada Penggunaan Layanan Go-Pay bagi Minat Pengguna di Indonesia. *Organum Jurnal Saintifik Manajemen dan Akuntansi* **3**(1), 37–49 (2020)



# Overlapping Region Extraction Method of 3D Laser Point Cloud Registration on Account of Artificial Intelligence Algorithm

Yaguang Wang, Wuzhan Yu, Yongdong Liang, and Bing Liu (✉)

Construction Branch of State Grid Liaoning Electric Power Supply Co., Ltd., Shenyang,  
Liaoning, China  
13709821139@163.com

**Abstract.** The registration of multi-view laser point cloud is the benchmark step of 3d model establishment, and the overlap of point cloud is very important for registration. In this paper, artificial intelligence algorithm is used to optimize the extraction of overlapping rate of the region. The specific process is: first, the internal geometric structure of the target is completely drawn and blocks are divided from multiple perspectives; second, the ESF multidimensional features of multiple regions are established. Gets the Euclidean distance between individuals. It is the basis of 3d target reconstruction, and the extraction of overlapping regions between point clouds is of great significance to improve the efficiency of registration. This paper studies the extraction method of 3d laser point cloud registration overlap region on account of artificial intelligence algorithm, and points out the extraction method and technical scheme of 3d laser point cloud registration overlap region. The data test shows that the research on the extraction method of the overlapping region for 3D laser point cloud registration on account of artificial intelligence algorithm has excellent performance in the extraction of the overlapping region.

**Keywords:** ARTIFICIAL Intelligence Algorithm · 3D Laser Point Cloud Registration · Overlapping Region Extraction · Method Research

## 1 Introduction

3D laser imaging has the characteristics of long distance, accuracy and no illumination, and can be used in a wide range of target identification, global analysis, local positioning and other fields, covering huge military and economic value [1]. Due to limited visibility and other factors, a single image cannot obtain the whole point cloud from the target. It is usually collected from multiple regions of the target, and the node cloud is combined into a system to form a set of target point aggregation. Cross-cloud alignment is to identify inter-cloud matching points on account of attribute descriptor elements, and to solve inter-cloud variable matrix by analyzing the relationship between a large number of matching points and sample points [2]. Research on the extraction method of 3d laser point cloud registration overlap region on account of artificial intelligence algorithm

effectively promotes the extraction effect of 3d laser point cloud registration overlap region.

As for the research of artificial intelligence algorithm, many scholars at home and abroad have carried on the research to it. In foreign studies, Czako and Z proposed a method to solve the problem of artificial intelligence algorithm selection and adjustment, without manual intervention, in a fully automated way. The method is a hybrid method combining particle swarm optimization. Experiments were carried out on some known data sets to prove the time efficiency and high accuracy of the method [3]. Mourad, M proposed a blind adaptive spectrum sensing algorithm on account of centralized collaborative sensing platform. Then, the adaptive neural fuzzy interference system (ANFIS) technology is applied to the decision-making process to achieve optimal and accurate decision. The simulation process and output results show that the performance of this method is superior to the single sensing technology and other collaborative sensing technology with conventi [4]. Spielberg, Y proposed and formulated a critical-based change Step Number algorithm (CVS) – a flexible step number algorithm that utilizes the criticality function provided by human beings or learns directly from the environment. Testing was done in three different areas, including the AtariPong environment, the Road-Tree environment, and the Shooter environment. It is proved that CVS can surpass popular learning algorithms, such as deep Q-learning and Monte Carlo [5].

The matching points between point clouds are limited to the overlap range of each point cloud sample, and the proportion of matching points is higher than that of matching points in the whole cloud. The transformation of the form is linear and low-dimensional, and there is a mismatch at the most precise point of comparison, full clouds are used to extract and match features, and their calculation errors are larger. Using overlapping regions to extract and match feature points can improve the efficiency and accuracy of distribution domain. Research on the extraction method of 3d laser point cloud registration overlap region on account of artificial intelligence algorithm is beneficial to improve the extraction repetition rate of point cloud registration overlap region.

## **2 Design and Exploration of Extraction Method of Overlapping Region for 3D Laser Point Cloud Registration on Account of Artificial Intelligence Algorithm**

### **2.1 Artificial Intelligence Algorithm**

The acronym for artificial Intelligence is AI [6, 7]. This is a method of operating simulation of the object of study, knowledge expansion of machine intelligence, which involves the theoretical system of intelligence is very cutting-edge science.

The system is a simulation of the whole information process of human thinking and consciousness. There are several mathematical modeling methods on account of artificial intelligence algorithm:

- (1) Qualitative and quantitative mixed system modeling method. A unified qualitative and quantitative algorithm modeling pattern is designed to describe modeling methods and knowledge in sub-domains and to describe the top level, this method

is applicable to the scope of data model to cooperate with the time operation mode, its purpose is to carry out quantitative and qualitative transformation of data.

- (2) Variable structure system modeling method [8, 9]. It includes modeling modes of various parameters, including variable components, dynamic ports and dynamic points. These methods can change the original model of the system to adapt to the multi-range changes of the system and the whole structure.
- (3) Modeling method on account of metamodel. It is a multi-functional modeling model on account of descriptive model, interactive decision model, multi-layer cognitive system model and adaptive system model.
- (4) Modeling method on account of big data intelligence. The complex and difficult points of the system mechanism in the simple mechanism of artificial intelligence can be simulated and analyzed by the genetic mechanism neural network mathematical modeling with the application of professional model system analysis to create a new type of artificial intelligence. A large amount of empirical data and its application requirements can be simulated and analyzed by intelligent modeling method.
- (5) Deep learning simulation modeling method [10, 11]. In the era of artificial intelligence, data is collected by means of collection and explosively increases, deep learning simulates human brain functions, and neural network technology is reused in artificial intelligence, which will provide a strong theoretical basis for developing system models.

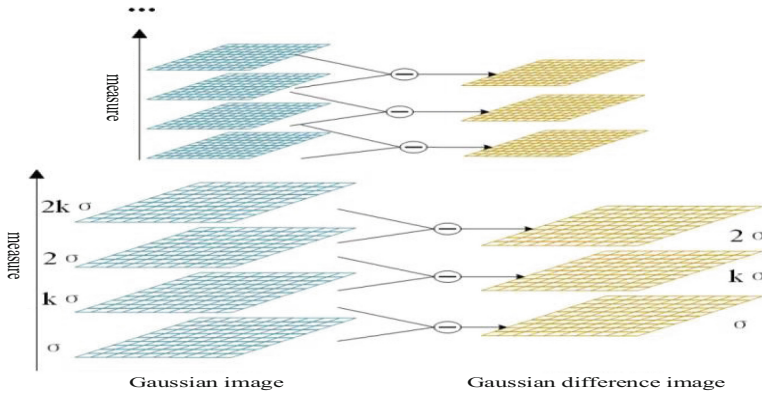
### **3 Research on the Extraction Method of Overlapping Region for 3D Laser Point Cloud Registration on Account of Artificial Intelligence Algorithm**

(1) Feature point extraction and image matching, as shown in Fig. 1, in order to achieve the digital image and laser point cloud registration, premise condition is to obtain a certain number of matching points, the stereo matching points matching to the laser point cloud, so as to realize two different data registration stereo matching points between stereo pairs is the foundation of the same name as point extraction, this belongs to the problem of image matching in digital photogrammetry.

#### **(2) Feature extraction**

Feature extraction is the basis of image analysis and image matching [12, 13]. To achieve the registration of digital image and laser point cloud, it is necessary to obtain a certain number of stereo matching points. The premise of obtaining stereo matching points is point feature extraction. Point feature extraction is mainly carried out by applying various operators as shown in Fig. 1. In the lower part of the first layer in the figure, there are four layers of samples. The second layer in the upper part of the figure has the same flow as the first layer.





**Fig. 1.** Feature point extraction and image matching

(3) Image matching

Image matching is one of the research emphases in the field of computer vision and digital photogrammetry, and its essence is to identify the homonyms between two or more images. The premise of recognition of homonymous points is to determine matching measure, so how to define matching measure is the primary task of image matching.

(4) Feature-based point cloud registration

Point cloud data alignment is an important task of 3D laser scanning data, which is essentially coordinate system alignment [14, 15]. The commonly used cloud data matching algorithm is mainly composed of four numbers and a matching algorithm, six or seven parameters and a recursive algorithm.

**4 Research on the Effect of Extraction Method for Overlapping Region of 3D Laser Point Cloud Registration on Account of Artificial Intelligence Algorithm**

3D point clouds are collected from K angles. 3D point clouds have overlapping areas between them. The point cloud viewed from K angles is initially coordinated in the same system [16, 17]. After preliminary approval, in order to improve the overall, 3D, model accuracy, it is necessary to achieve the optimal level of global compatibility using multidimensional clouds. The objective of global optimization is to determine the rotation transition parameters so that the global fitting e error is minimal. The error formula is:

$$e = \sum_{\substack{m \neq n \\ m=1,2,\dots,K \\ N=1,2,\dots,K \\ i=1,2,\dots,N}} f \tag{1}$$

In the above formulas,  $m$  and  $n$  refer to the point cloud data from the MTH and NTH perspective,  $I$  refers to the point with the same name of the  $i$ th pair, and  $F$  is the error calculation model.

It is assumed that the number of three-dimensional points in the fixed-point cloud is set as  $m$  and  $N$  respectively, and one of all three-dimensional points is searched for the nearest point in the other point cloud. If the distance between them is not greater than a certain threshold value, the two points are defined as approximate points with the same name,  $k$ -D tree is used to detect and traverse all similarity points, assuming the number is  $N$ , and the overlap equation  $W$  of two-point cloud is calculated:

$$W = \frac{N}{\min(m, n)} \quad (2)$$

It is assumed that the set of different scanned sample points is  $P$  and  $Q$ , and the corresponding coordinates are  $O$ -XYZ and  $o$ -XYZ, where  $P$  and  $Q$  are points of the same  $P(x, y, z) \in P, Q(x, y, z) \in Q$  in the coordinates, then the problem becomes shape deformation, and the algorithm formula of multiple data pairs ( $P, Q$ ) of point cloud is as follows:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mu R(\alpha, \beta, \gamma) \begin{bmatrix} x \\ y \\ z \end{bmatrix} + T \quad (3)$$

The above formula  $\mu$  is the parameter of the algorithm formula,  $R$  and  $T$  represent the rotation matrix and translation matrix  $\alpha, \beta$  and  $\gamma$  are the rotation parameter in the algorithm formula.

Image feature assisted point cloud feature extraction.

(1) Hough, extracting straight lines by transformation,

Hough transform is a transformation relation from image space to parameter space. Because of its good robustness to noise interference and object discontinuity, it has been widely used.

(2) Canny, operator extraction edge,

Edge extraction is an important aspect in digital image processing. Edge can be thought of as the image local change is the most vigorous part, was a key part of the human eye image, generally USES the mathematical equation of derivative method for testing, in the image edge has two features: direction along the edge of gray level change is very small, and the direction perpendicular to the edge of gray change drastically, the principle of edge detection are dependent on the two features.

(3) Point cloud feature extraction

Digital image feature extraction is the pixel value of image features, and these pixels have corresponding point sets in the laser point cloud, which constitute the corresponding features in the laser point cloud. Due to the influence of parallax, registration accuracy, image extraction effect and other factors, these point sets may not be able to accurately express the laser point cloud features. By fitting, noise points can be removed to improve the accuracy of point cloud feature extraction. These points are a series of discrete three-dimensional space coordinate points, and the curve fitting of these points can obtain the spatial graph of the building.

### 5 Investigation and Research Analysis of the Extraction Method of Overlapping Region for 3D Laser Point Cloud Registration on Account of Artificial Intelligence Algorithm

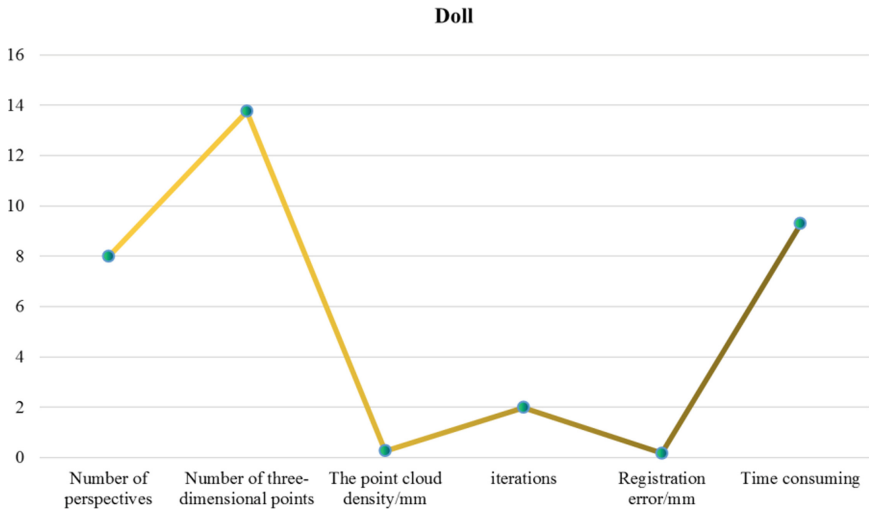
3D laser scanning performed tests and analyses on a set of doll model data. All the toys using laser scanners collected 8 sets of 3D scan data and tested using artificial intelligence algorithms.

Two sets of tests were performed on the survey data and the results were compared. In the first group of tests, 3 or more pairs of similar overlapping sample points were selected to calculate the initial parameters of shape transformation, and then all the scanned data obtained by ICP algorithm for accurate distribution of two-point cloud were formatted into the same system. In the second group of experiments, the point cloud optimization algorithm in this paper was used to analyze three-dimensional point variable parameters from the general equilibrium perspective, and then all the measured data were aligned to a system. Test data are shown in Table 1.

**Table 1.** Extracting test data from overlapping regions of 3D laser point cloud registration

Model	Number of perspectives	Number of three-dimensional points (*100,000)	The point cloud density/mm	iterations	Registration error/mm	Time consuming (*10)
Doll	8	13.75863	0.28	2	0.18	9.3

In the table, the test name extracted from the overlapping area of 3d laser point cloud registration is Doll. Number, of, Perspectives, Number of THREE-DIMENSIONAL



**Fig. 2.** Test images being extracted from overlapping areas

points, Point cloud density (The, Point, Cloud, Density/mm), Iterations, Registration intermediate error (ERROR/MM) and Time Consuming are shown in the table above respectively, and their graphs are shown in Fig. 2.

It can be seen from the figure that the artificial intelligence algorithm is used to extract the overlapping area of 3D laser point cloud registration with only two iterations and the point cloud density is 0.28 mm, indicating that the repetition rate is very good.

Data testing, artificial intelligence algorithm on account of 3d laser point cloud registration overlap region extraction method research, has a good performance in the area of overlap region extraction.

## 6 Conclusions

In this paper, a region segmentation algorithm on account of geometric attributes is used to segment perspective point clouds into multiple modules, and then multi-dimensional attributes are determined for each region, and similar regions are defined as overlapping regions of point clouds. The test results show that the text algorithm can improve the overlap rate between point clouds, and can maintain consistent performance in the simulation test under the condition of strong point cloud heterogeneity. Research on the extraction method of 3d laser point cloud registration overlap region on account of artificial intelligence algorithm promotes the extraction effect of 3d laser point cloud registration overlap region.

## References

1. Jagadev, P., Giri, L.I.: Human respiration monitoring using infrared thermography and artificial intelligence. *Biomed. Phys. Eng. Express* **6**(3), 035007 (2020)
2. Baş, E., Ülker, E.: Comparison between SSA and SSO algorithm inspired in the behavior of the social spider for constrained optimization. *Artif. Intell. Rev.* **54**(7), 5583–5631 (2021). <https://doi.org/10.1007/s10462-021-10035-x>
3. Czako, Z., Sebestyen, G., Hangan, A.: AutomaticAI – a hybrid approach for automatic artificial intelligence algorithm selection and hyperparameter tuning. *Expert Syst. Appl.* **182**(16), 115225 (2021)
4. Mourad Mabrook, M., Khalil, H.A., Hussein, A.I.: Artificial intelligence based cooperative spectrum sensing algorithm for cognitive radio networks. *Procedia Comput. Sci.* **163**, 19–29 (2019). <https://doi.org/10.1016/j.procs.2019.12.081>
5. Spielberg, Y., Azaria, A.: Criticality-based varying step-number algorithm for reinforcement learning. *Int. J. Artif. Intell. Tools* **30**(4), 2150019 (2021)
6. Joe, J.R., Ulla, J.W., Commerford, B.P., et al.: Man versus machine: complex estimates and auditor reliance on artificial intelligence. *J. Account. Res.* **60**(1), 171–201 (2022)
7. Turkoglu, B., Uymaz, S.A., Kaya, E.: Clustering analysis through artificial algae algorithm. *Int. J. Mach. Learn. Cybern.* **13**(4), 1179–1196 (2022)
8. Kachaoui, J., Belangour, A.: Enhanced data lake clustering design on account of k-means algorithm. *Int. J. Adv. Comput. Sci. Appl.* **11**(4), 547–554 (2020)
9. Koulu, Riikka: Proceduralizing control and discretion: human oversight in artificial intelligence policy. *Maastricht J. Euro. Comparat. Law* **27**(6), 720–735 (2020). <https://doi.org/10.1177/1023263X20978649>

10. Et, A.: Modern swarm intelligence based algorithms for solving optimal power flow problem in a regulated power system framework. *Turkish J. Comput. Math. Educ. (TURCOMAT)* **12**(2), 1786–1793 (2021)
11. Mohanty, S.K., Udgata, S.K.: Minimizing the maximum receiver interference in wireless sensor networks using probabilistic interference model. *Eng. Appl. Artific. Intell.* **91**(May), 103563.1–103563.14 (2020)
12. Gajawada, S., Mustafa, H.: Artificial god optimization - a creation. *Comput. inform. Sci.* **13**(1), 41–50 (2020)
13. Khan, R., Debnath, R.: Human distraction detection from video stream using artificial emotional intelligence. *Int. J. Image Graph. Sign. Process.* **12**(2), 19–29 (2020)
14. Fotsing, C., Mbouombouo, N.N., Bobda, C.: Large common planesets-4-points congruent sets for point cloud registration. *Int. J. Geo-Inform.* **9**(Virtual 3D City Models), 647–647 (2020)
15. Eslami, M., Saadatesherst, M.: Imagery network fine registration by reference point cloud data on account of the tie points and planes. *Sensors* **21**(1), 317 (2021)
16. Al-Rawabdeh, A., He, F., Habib, A.: Automated feature-based down-sampling approaches for fine registration of irregular point clouds. *Remote Sens.* **12**(7), 1224 (2020)
17. Islamutdinov, V.F.: Impact of digitalization on the evolution of internal corporate institutions in the northern resource-extraction region. *SHS Web Conf.* **89**(3), 03002 (2020)



# Smart Urban Carbon Emission Management Platform Based on Energy Big Data

Xue Li<sup>1</sup>, Zheyong Piao<sup>1</sup>, Yushan Zheng<sup>1</sup>, Jieping Han<sup>2</sup>, and Rijie Cong<sup>2</sup>(✉)

<sup>1</sup> Baicheng Power Supply Company of State Grid Jilin Electric Power Co., LTD, Baicheng, Jilin, China

<sup>2</sup> Northeast Electric Power University, Jilin, Jilin, China  
569574813@qq.com

**Abstract.** Low-carbon development has become the focus of all aspects in the whole society. China has put forward higher requirements for the carbon emissions of cities and energy-intensive enterprises. This paper discusses the background and current situation of the construction of urban carbon emission management platform. At the same time, the necessity of building a smart urban carbon emission management platform based on energy big data is analyzed. In addition, according to the construction principles of the system platform, the urban smart carbon emission management platform system is constructed, and the main functions of the platform are mainly designed. The application of the platform can help the government control the carbon emissions, new energy consumption and energy efficiency improvement space of regional and key enterprises in real time, and facilitate the realization of the national major strategy of “carbon peak and carbon neutrality”.

**Keywords:** Energy · Big data · Carbon Emissions · Wisdom

## 1 Introduction

At the General debate of the Seventy-fifth session of the United Nations General Assembly, the General Secretary Xi announced that China will adopt more powerful policies and measures to peak its carbon dioxide emissions before 2030 and strive to achieve carbon neutrality by 2060. At present, in order to realize the vision of carbon neutrality, we should also make breakthroughs in the clean and low-carbon utilization of fossil energy and other systematic technologies for vigorously promoting the development of clean energy and facilitating the optimization of energy structure. Innovation is also needed in improving energy efficiency, controlling total energy consumption, and using big data, visualization, Internet and advanced digital technologies to effectively control the Carbon Footprint of key enterprises with high energy consumption in real time.

In early 2021, State Grid issued the action plan of “Carbon Peak and Carbon Neutrality”, proposing to build an energy digital economy platform, accelerate the construction and operation of energy big data center, and promote the implementation of strategic

goals and strategic measures in order to implement the deployment of new digital infrastructure construction of SGC. At the same time, the construction of energy data centers should be vigorously promoted to meet the integrated development trend of energy revolution and digital revolution, and meet the development needs of new energy business forms. Through the construction and operation of city-level energy big data center, it comprehensively integrates energy data such as water, electricity, coal, gas and oil, and excavates the hidden economic and social value behind energy big data. These measures can offer a great significance to the rational allocation and utilization of energy resources, the establishment of a widely interconnected, integrated and open energy Internet ecology, the optimization and transformation of energy structure and the on-demand flow of energy resources, and the establishment of an energy conservation system for industrial development and social life. Meanwhile, they also provide strong technical support for the construction of energy saving cities.

Led by the government, the energy big data center focuses on the construction of new smart cities and the development of smart industries in the digital economy to coordinate all kinds of energy sources in the city, so as to realize the joint construction, sharing and win-win of all participants. The completion and operation of the energy big data center is of great strategic significance to guide the development direction of energy services, and also has significant effects in promoting the transformation and upgrading of territorial industries, promoting employment, increasing tax revenue, improving environmental benefits, and driving social and economic development. The urban smart carbon emission management platform based on energy big data aims to help the government control the carbon emissions, new energy consumption and energy efficiency improvement space of regional and key enterprises in real time, complete the comprehensive evaluation and analysis of carbon emissions of relevant enterprises, and promote the tracking, monitoring and regulation of renewable energy consumption by the government. At the same time, it can predict the future development trend of carbon emissions and new energy consumption of different regions and enterprises, so as to provide auxiliary decision-making for the government and enterprises. The design of the platform has important practical significance for real-time control of key enterprises' carbon emissions, new energy consumption, energy efficiency improvement space and the realization of carbon neutral vision in the future.

## 2 Research Status

Wong, Johnny KW described a prototype architecture and a building project simulation tool for carbon emission prediction using virtual prototyping technology to demonstrate the application of the visualization tool for emission prediction by studying a real public housing construction project in Hong Kong. It is useful for exploring strategies to reduce carbon emissions in the construction sector [1]. Yan Zhang, Jianjiang Zhang, Guoyong Yu started with the development of energy management and control system through intelligent cloud platform in the industrial electricity environment, identified the advantages of intelligent cloud platform, and introduced the actual use of industrial electricity [2]. Lu Jun combined with the current situation analysis of carbon asset management mode and information physical system, applied IoT, AI, big data, 5G and other digital

technologies, through strategic positioning, organization and personnel, process design, rules and regulations, digital means, operation supervision and other paths. She proposed the strategy of building a carbon asset digital management system of group enterprises with “three modules, four levels and eight functions”. It has a more significant advantage of information sharing, and plays a promoting role in the development of carbon asset management business and the activity of carbon trading market in China [3]. Guo Feng, Yang Shangguang, Ren Yi empirically investigated the impact of digital economy development on urban carbon emissions by constructing a city-level digital economy development index and using panel data of 223 cities in China from 2011 to 2019, and identified the action mechanism of green technology innovation. They found that digital economy development could significantly reduce urban carbon emissions. This study provided a new perspective and empirical evidence for understanding the relationship between digital economy development and carbon emissions [4]. Zhou Ji, Xu proud, Liu Xi and Li Jieling proposed to build a “digital intelligent carbon control” platform in Jiangxi Province, China, incorporate the “digital intelligent carbon reduction and control mechanism” into the “No. 1 development project” to improve the digital economy and accelerate the construction of the “1 + 1 + n digital intelligent carbon control” platform system by virtue of digital intelligent. It means to accurately monitor carbon emissions, efficiently analyze carbon data and scientifically support carbon decision-making [5]. Zhang Jinmeng established the linkage model of municipal electricity consumption and carbon emissions, and formed the power-carbon map of 14 cities in Hubei Power Grid of China, which could visually analyze the change trend of carbon emissions in 35 major carbon emission source industries, and opened the data chain of “power-carbon-energy” [6].

In summary, scholars used AI, big data, 5G, virtual simulation and other technologies to obtain the relevant information of enterprises and integrate and analyze the data, so as to achieve the purpose of energy saving and carbon reduction. In terms of the planning and application of the carbon emission control platform, the role and influence of the Internet platform on energy conservation and emission reduction were mainly studied, and relevant suggestions were put forward. But related research lack of carbon emissions to the industrial enterprise visualization analysis and research, the data is static data, lack of real-time dynamic data collection platform. And this research will be based on a great energy data center enterprise dynamic energy consumption data collection, analysis and carries on the depth of mining, help enterprises to gain mining carbon reduction, accurate reduction. At the same time, it can help the government and upstream and downstream enterprises to timely understand the basic information, energy consumption information and carbon emission information of enterprises, so as to realize multi-party docking and make more accurate decisions.

### **3 Platform Constructions**

#### **3.1 Necessity of Platform Construction**

Collection in the face of increasing energy consumption, energy source shortage exist, less access, offline artificial acquisition is difficult to ensure data quality problems. The city level great energy data center arises at the historic moment [7]. How to realize



data acquisition, energy data on energy consumption, power data and industry economy, regional economic analysis, in-depth exploration of the economy behind energy is an urgent problem to be solved [8]. Carbon peaking and carbon neutrality is an extensive and profound systemic economic and social transformation. Actively addressing climate change is an inevitable requirement for China to achieve sustainable development. At present, there are some problems in carbon verification and carbon monitoring, such as poor data quality and inaccurate accounting results. Therefore, the construction of urban smart carbon emission management platform based on energy big data is of great importance.

The overall goal of building a smart urban carbon emission management platform based on energy big data is to make full use of new technologies such as big data, cloud computing, Internet of things, mobile Internet, artificial intelligence and block-chain, build an energy data fusion and sharing mechanism, break through key core technologies, and lead energy data management and technical standards. Gradually realize the convergence and fusion, sharing and exchange, mining and analysis of cold, heat, water, electricity, gas, coal, oil and other energy data. The main function of the platform is to conduct statistical analysis of carbon emissions and new energy consumption of regions, industries and enterprises [9]. The main functions include: total regional carbon emissions, industrial carbon emissions, enterprise carbon emissions, carbon emission intensity, carbon source structure, carbon emission warning, carbon neutralization rate, decision support and other modules. The system features comprehensive evaluation and analysis of relevant enterprises' carbon emissions by comparing national and provincial and municipal local standards, and gives hierarchical visual warning hints by automatic calibration. According to the new energy consumption data and carbon emission reduction data of key enterprises in the data center, the carbon neutrality rate of regions and enterprises is calculated, and the main factors affecting carbon emissions are analyzed. At the same time, the carbon emission data of governments and enterprises are predicted, and decision support is provided for the implementation of carbon emission reduction methods [10].

### 3.2 Principles of Platform Construction

#### (1) Multi-dimension

Platform in the scope of energy data collection including power source - net - load - store "each link information flow and coal, oil, gas, water, heat and other kinds of energy production, transmission, storage and consumption data, and related environmental data, data resource directory, realize data collection and the collection process, continued to expand data gathering of breadth, depth and density.

#### (2) Advanced nature

The design of platform technology should have the ability of processing and storing massive data, and the ability of distributed storage and computing processing, so as to meet the storage and processing requirements of massive real-time data. It can provide urban carbon emission panorama, carbon emission real-time data, carbon emission historical data, carbon flow balance analysis, operation index statistical analysis and other sub-functions.

#### (3) Openness

The platform should have an open publishing and running environment. The application service framework of the system adopts J2EE architecture, which has the advantages of scalability, flexibility, easy maintenance and dynamic deployment. It has the ability to provide low-cost data collection equipment, high-quality energy and carbon emission data processing and analysis, and can flexibly connect with various systems and platforms.

#### (4) Sharing

The platform can provide data resource sharing, enterprise applications on the cloud, carbon emissions data set monitoring service, the value of data mining, Shared smart sensor data and trend of energy consumption, improve the operation efficiency, and pass through water, electricity, gas, heat and other related industry barriers between data, output business consulting report, optimize the business environment, ensuring and improving people's livelihood.

#### (5) Multiple applicability

Through the analysis of user portrait, environmental characteristics and data information content, it provides real-time and customized carbon emission information, industry hot spot sharing, carbon price index push, focus problem analysis and policy interpretation, carbon emission analysis, carbon emission reduction decision for government, industry, financial sector and other departments and members.

### 3.3 Overall Framework of the Platform

The architecture follows the unified construction technical route, structures contain resource foundation support layer, data layer, business application layer, display layer data center platform, through the data link. It establishes standard specification data interface, data mining and integrate the energy value, comprehensive support urban wisdom carbon management platform construction, enhance the level of urban comprehensive energy management and carbon emissions. This framework is the overall framework to achieve "one center, two systems, multiple services". It is equivalent to the significance of great energy data center as the core, with standard system and security system two big system for support, relying on the hardware and software infrastructure platform, data resources platform, business application platform, for the government, industrial users, energy suppliers and social livelihood of the people to provide data analysis, data mining, such as policy more energy and carbon management services. The overall architecture of the platform is shown in Fig. 1.

### 3.4 Main Functions of the Platform (Fig. 2)

#### (1) Total energy consumption

Based on the energy data of the energy big data center, the total energy consumption of regions, industries and enterprises is calculated with reference to various energy discount and standard coal coefficients, which are based on the time dimension (year, month, day), statistical analysis and visual chart display.

#### (2) Energy intensity

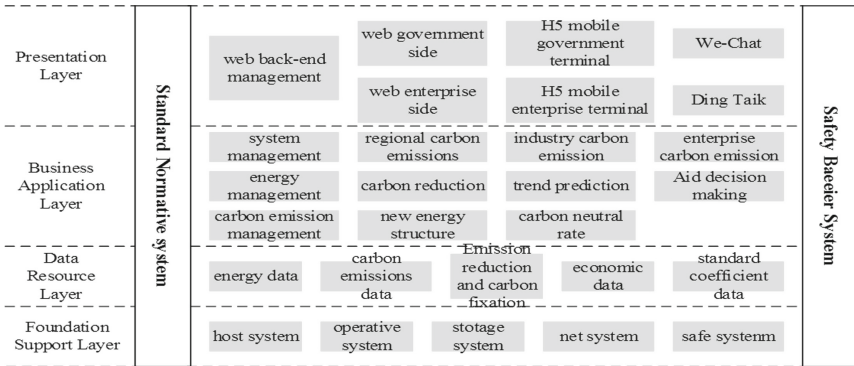


Fig. 1. Overall architecture of the platform

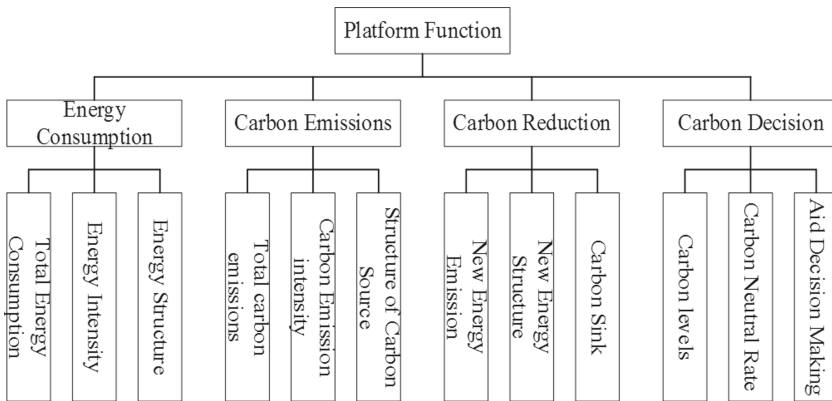


Fig. 2. Main functions of the platform

The energy consumption per unit GDP of cities, industries and enterprises is calculated based on the total energy consumption and GDP data. Monitor the change rate of energy consumption of regions, industries and enterprises.

(3) Energy structure

According to the standard coal quantity of various energy sources, the proportion of energy consumption of cities, industries and enterprises is counted in the form of scale map.

(4) Total carbon emission

Based on the energy data of the energy big data center, the total carbon emissions of regions, industries and enterprises, they are calculated with reference to the carbon emission coefficients of various energy sources, which are based on the time dimension (year, month, day), statistical analysis and visual chart display are performed. The use progress and proportion of carbon quota in regions, industries and enterprises were monitored.

(5) Carbon emission intensity

Calculate the carbon emission intensity per unit of output value according to the total carbon emission value and GDP value of the region, industry and enterprise. The carbon emission intensity per unit of production capacity is calculated based on the total carbon emission value and production capacity data of enterprises.

(6) Carbon source structure

According to the indirect (electricity) and direct carbon emission (heat, coal, oil and gas) sources of regions, industries and enterprises, the proportion of carbon source types was analyzed and visualized according to the time dimension (year, month and day).

(7) New energy emission reduction

The new energy consumption is calculated according to the new energy consumption of regions, industries and enterprises, and the completion of the new energy consumption is counted and displayed according to the time dimension (year, month and day), so as to promote the tracking monitoring and supervision of the renewable energy consumption by the government.

(8) New energy structure

According to the new energy consumption of regions, industries and enterprises, the proportion of three kinds of new energy consumption (photovoltaic, wind power, hydro-power) is analyzed and displayed according to the time dimension (year, month, day), so as to provide data support for the strategic decision of new energy emission reduction.

(9) Carbon sinks

According to the forest cover and grassland cover of the region, the forestry carbon sink of the region was counted in time dimension to provide data support for the strategic decision of forestry emission reduction.

(10) Carbon emission level

The carbon emissions of regions, industries and enterprises were compared with the evaluation levels. Four levels of visual prompts (green, yellow, orange and red) were given by “automatic calibration” and referring to the grade color definition standard. Comprehensive evaluation and analysis of carbon emissions were conducted according to the time dimension (year, month, day).

(11) Carbon neutrality rate

The carbon emission reduction and carbon sink generated by new energy consumption will be counted. The emission reduction of regions, industries and enterprises can be obtained, and the ratio value is formed with the emission as the denominator and the total emission reduction as the numerator to represent the carbon neutrality progress. According to the primary standard, the ranking of regions, industries and enterprises can be formed to effectively check the carbon neutrality progress.

(12) Decision support

Various data analysis tools such as growth rate estimation method, gray system prediction model, moving average method and trend extrapolation prediction method are used to predict the future carbon emission trend, new energy emission reduction situation and carbon emission level of regions, industries and enterprises. Through the horizontal and vertical comparison of the data, the multi-expert swarm wisdom decision-making algorithm is designed, and the convolution learning technology is mainly applied to study the effect of energy conservation and emission reduction incentive policies.

## 4 Benefit Analyses

### (1) Economic benefits

By focusing on enterprise carbon emissions and energy consumption of new energy given trend prediction, effective power enterprises can adjust energy consumption behavior, update equipment, promote enterprise actively given clean energy, enterprise diversification of energy supply, and clean and low carbon, a scientific basis for future energy planning for the enterprise. At the same time, it also provide enterprises with energy conservation and emissions reduction measures and feasible Suggestions of energy saving technology to help enterprises reduce energy consumption and improve energy efficiency. The government can promote the integration of multi-sector businesses and platforms, encourage enterprises to register and use smart carbon emission management platforms as soon as possible, enrich databases, break through data barriers between industries, and truly achieve digital empowerment, quality and efficiency improvement.

### (2) Social benefits

After the completion of the platform, the analysis and prediction of energy use data and carbon emission data can help the government to grasp the economic development status of different regions and different industries, evaluate the scientificity and sustainability of the development mode, so as to provide reference for the government in economic development decisions. Formulate subsidy and regulation policies for the industry, guide efficient industrial restructuring, industrial integration and industrial upgrading, and achieve resource integration and optimization. The utility of energy big data is fully exploited, and the total carbon emission of each index of enterprise carbon source consumption is calculated according to the carbon source data and the national carbon emission coefficient standard. According to the given indicators of the government, the carbon emissions of relevant enterprises are comprehensively evaluated and analyzed, and hierarchical visual early warning is given in the way of automatic calibration. And the national renewable energy consumption responsibility weight is decomposed and allocated, intuitive and quick statistics, display the completion of new energy consumption, will effectively promote the government's renewable energy consumption tracking monitoring and supervision.

### (3) Environmental benefits

For energy data collection and efficient analysis is helpful to promote the development of energy saving energy companies into better, based on the different energy consumption and carbon emissions data is analyzed. It is concluded that the energy consumption and carbon emissions and put forward countermeasures, further form a complete monitoring system, and focus on the information collection and analysis of new energy. It can greatly alleviate the serious problem of energy waste, and promote the whole society to save energy, reduce carbon emissions and energy waste. By analyzing the carbon emission monitoring data of enterprises and comparing the energy consumption and carbon emission of the same type of enterprises, the key enterprises can be accurately monitored and the effective implementation of environmental protection policies can be ensured. Comprehensive energy use efficiency, carbon emission level, user behavior, green distribution and other information can be used to construct a multi-temporal and spatial scale carbon emission management scheme from the two aspects of prevention and treatment to ensure the sustainable development of the city. Comprehensive user

profile, energy output, energy price, carbon emission price and other information can be used to formulate reasonable energy trading policy, carbon emission trading policy and energy subsidy policy, so as to help achieve carbon neutrality at an early date.

## 5 Conclusions

Based on the energy big data, this paper analyzes the necessity of smart carbon emission management platform in cities and designs the system functions. The construction of the platform can effectively promote the regional carbon emission management and improve the efficiency of regional energy utilization. At the same time build digital management platform can assign the wisdom of the carbon energy, using technology such as the big data dig energy value of big data, accurate, intuitive, and timely reflect the carbon emissions of the enterprise and the new energy use. It is advantageous to the energy of their own carbon emissions, clear structure of carbon source, for emission reduction measures to provide data support and decision support.

**Acknowledgement.** This work was supported by the Science and Technology Project of Baicheng Power Supply Company, State Grid Jilin Electric Power Co., LTD. Project Name: Research and Service Project of Energy Carbon Sand Table for Baicheng Industrial Enterprises based on virtual reality technology.

## References

1. Wong, J.K.W., et al.: Toward low-carbon construction processes: the visualisation of predicted emission via virtual prototyping technology. *Autom. Construct.* **33**, 72–78 (2013). <https://doi.org/10.1016/j.autcon.2012.09.014>
2. Zhang, Y., Zhang, J., Yu, G., et al.: Development of an energy management and control system on a smart cloud platform in an industrial power environment. *Int. J. High. Educ. Teach. Theory* **3**(2) (2022)
3. Jun, L.: Construction strategy of carbon assets digital management system of group enterprises. *Xinjiang Oil Gas* **18**(02), 10–15 (2022)
4. Feng, G., Shang-guang, Y., Yi, R.: Digital economy, Green technology innovation and carbon emission: empirical evidence from urban level in China. *J. Shaanxi Normal Univ. (Phil. Social Sci. Edn.)* **51**(03), 45–60 (2022)
5. Zhou, J., Hao, Z., Liu, X., Li, J.: Reflections and suggestions on the construction of “digital intelligence and carbon control” platform system in Jiangxi. *China Natl. Cond. Natl. Strength* **06**, 40–44 (2022)
6. Gallego-Álvarez, S., Segura, L., Martínez-Ferrero, J., et al.: Carbon emission reduction: the impact on the financial and operational performance of international companies. *J. Clean. Product.* **103**, 149–159 (2015)
7. Neuvonen, A., Kaskinen, T., Leppänen, J., et al.: Low-carbon futures and sustainable lifestyles: a backcasting scenario approach. *Futures* **58**, 66–76 (2014)
8. Wang, Y., Yang, H., Sun, R., et al.: Effectiveness of China’s provincial industrial carbon emission reduction and optimization of carbon emission reduction paths in “lagging regions”: Efficiency-cost analysis. *J. Environ. Manag.* **275** (2020)

9. Ji, J., Zhang, Z., Yang, L., et al.: Carbon emission reduction decisions in the retail-/dual-channel supply chain with consumers' preference. *J. Clean. Prod.* **141**, 852–867 (2017)
10. Samarasinghe, D.A.S., Baghaei, N., Stemmet, L.: Persuasive virtual reality: promoting earth buildings in New Zealand. In: Gram-Hansen, S.B., Jonassen, T.S., Midden, C. (eds.) *PER-SUASIVE 2020*. LNCS, vol. 12064, pp. 208–220. Springer, Cham (2020). [https://doi.org/10.1007/978-3-030-45712-9\\_16](https://doi.org/10.1007/978-3-030-45712-9_16)



# A Distribution Network Reconfiguration Method Considering Optimal Distribution of DG Output

Chao Feng<sup>(✉)</sup>, Yuanhong Hou, Shijun Li, Hui Xue, Yulai Yang, and Shunhai Xue

State Grid Qinghai Electric Power Company Huanghua Power Supply Company,  
Jianzha 811200, Qinghai, China  
qh Fengchao@163.com

**Abstract.** Distribution Network Reconfiguration (DNR) is an important means to reduce line loss and improve power supply reliability. It is an important research topic in the field of distribution network to optimize the network topology of distribution network and improve the operation state of the original distribution network by operating the line switch of distribution network. This paper presents a distribution network reconfiguration method. The goal of this method is to minimize the total cost of power system operators (PSOs) changing the distribution network configuration, while taking into account the optimal distribution of distributed generation (DG) output among load centers. In addition, it aims to minimize the amount of DG output wasted due to sub optimal allocation. To achieve these goals, we propose an algorithm that considers the economic and technical constraints of PSO distribution network design.

**Keywords:** Distribution network reconfiguration · Distributed generation · Minimum spanning tree · Heuristic rules

## 1 Introduction

Fossil energy is the cornerstone of the development of human society. However, the unrestrained exploitation of fossil energy for a long time has caused the tension in the supply of fossil energy. Therefore, how to reduce the rate of resource depletion while ensuring high-quality and stable supply of energy in the process of social development has become a major problem faced by countries around the world [1].

Solar energy, wind energy and other natural resources are clean, pollution-free and can be supplemented in a short time. Therefore, developing distributed generation technology and reasonably and efficiently using renewable energy for power supply has become an effective way to solve this problem. With the rapid development of wind power generation, photovoltaic power generation and other power generation technologies, the capacity of distributed power generation is increasing year by year, which has become a hot industry in various countries. Figure 1 shows the change trend of the installed capacity and growth rate of power generation in China at the end of 2019 [2]. It can be seen that the sum of installed capacity of wind power and solar power has exceeded the installed capacity of hydropower, only second to the installed capacity of



traditional thermal power generation, and their annual growth scale is far higher than that of other power generation methods, This shows that in recent years, the distributed power generation represented by wind power generation devices and photovoltaic power generation devices is vigorously developing in China, and the development momentum is still very rapid [3].

The concept of distributed power generation is put forward relative to the characteristics of traditional centralized power supply, which can be understood as the generation equipment or energy storage device that can independently supply electric energy with relatively dispersed geographical locations, independent operation and relatively small capacity distributed in a certain space and connected to the distribution network [4]. Compared with the traditional energy, the distributed power supply is installed closer to the user side, so that it can more meet the needs of users for power generation or power supply, effectively making up for the shortage of centralized power generation.

## 2 Related Work

### 2.1 Research on Distribution Network Reconfiguration Under Normal Operation

When the initial state of the distribution network is normal operation, the main purpose of network reconfiguration is to improve the quality of power supply, reduce network loss and optimize the operation state of the network. With the expansion of distribution network scale and the improvement of requirements for distribution network operation, researchers around the world have carried out a lot of effective research on distribution network reconfiguration under normal operation [5].

In the distribution network optimization model established by relevant scholars under normal operation, common objective functions include reducing network loss, optimizing power supply cost, improving voltage quality, etc. The objective is to reduce the system operation cost, and the operation costs considered include power purchase cost, switching operation cost and distributed power generation operation cost; A distribution network reconfiguration model is proposed to improve the DG absorptive capacity in a single time section, and the network structure at that time is optimized according to the future information; A new distribution network reconfiguration strategy is proposed, whose optimization goal is to reduce the balance of safe distance; Considering the three-phase imbalance problem caused by user load imbalance, a new optimization strategy is proposed to solve the overall imbalance problem of distribution network; Literature takes the minimum network loss of distribution network as the network optimization objective, and comprehensively considers constraints such as power flow constraints, voltage constraints, and main network capacity constraints [6],

A new distribution network reconfiguration model is proposed; In order to improve the utilization rate of renewable energy power generation connected to the power grid, enhance the absorptive capacity of the distribution network system to natural resources, and optimize the DG output in a unit period, a dynamic reconfiguration strategy based on the output similarity of each unit of the distribution network is proposed to maximize the DG utilization [7].

## 2.2 Research on Network Structure and Reactive Power Optimization Considering DG Uncertainty

The rapid development of distributed generation technology, although to a certain extent it has reduced the environmental pollution caused by fossil energy and alleviated the energy crisis, but it has also brought new challenges to the reliable operation of the traditional power grid. Network reconfiguration can better meet this challenge. At the same time, as a part of power system optimization, reactive power optimization of distribution network is of great significance to reduce line loss and improve grid voltage level. This paper focuses on the distribution network reconfiguration and reactive power optimization with DG connected under the consideration of uncertainties such as load and DG in the system.

With the development and popularization of distributed power technology, network reconfiguration considering DG access has gradually become a hot research topic in the academic community. Experts at home and abroad have carried out a lot of research on this issue. A heuristic algorithm is used to solve the reconfiguration problem of distributed generation equipment after grid connection, and the problem of network loss allocation is studied; For distributed power sources such as wind power generation and photovoltaic power generation, a variety of distributed power source models are established. In order to minimize the actual power loss, the decimal coded quantum particle swarm optimization algorithm is used to solve the mathematical model; Taking wind power generation equipment as PQ node, a hybrid algorithm combining differential evolution algorithm, ant colony algorithm and variable neighborhood search algorithm is proposed [8]. The proposed algorithm is used to simultaneously optimize the topology of distribution network and DG output. Although the above studies have analyzed the distribution network reconfiguration from different perspectives, they simply treat the distributed generation as a constant power model, and do not consider the uncertainty of the output of the power generation devices caused by the environmental impact of the distributed generation devices, so there is much room for improvement.

## 3 Reconstruction Based on Minimum Spanning Tree and Heuristic Rules

The distribution network reconfiguration model with the minimum active power loss is:

$$\min f = \sum_{i=1}^n k_i r_i \frac{P_i^2 + Q_i^2}{V_i^2} \quad (1)$$

Where: nb is the number of branches in the system;  $R_i$  is the resistance on branch  $i$ ;  $P$  and  $Q$  are the active power and reactive power flowing through branch  $i$  respectively;  $V_i$  Injects voltage of power node for branch  $i$ ;  $K_i$  indicates the switching state of branch  $i$ . When the branch is closed to 1, the branch is disconnected to 0.

The constraints are as follows:

- 1) Power flow constraint;
- 2) Operational constraints;
- 3) Capacity constraints;
- 4) Network topology constraints

After the reconfiguration, the distribution network topology should meet the radial topology structure, and there should be no isolated nodes.

In order to reduce the number of candidate switches, according to the above heuristic rule 2), only the switches at the low voltage side of the disconnect switch need to be evaluated during the reconstruction process. Definition of optimal switch: when the candidate switch meets the requirement that the network loss after opening the switch is less than the system network loss after opening the adjacent switch at the lower voltage side, the switch is the optimal switch. This greatly reduces the number of candidate switches [9].

Definition of optimal switch: when the candidate switch meets the requirement that the network loss after opening the switch is less than the system network loss after opening the adjacent switch at the lower voltage side, the switch is the optimal switch.

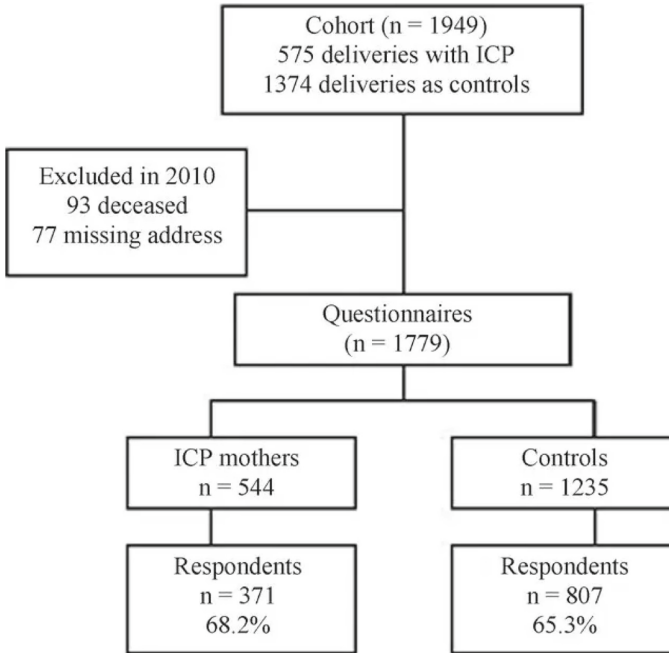
Based on the above improvement of candidate switch evaluation strategy, the following two new candidate switch evaluation strategies are obtained:

- 1) Only when the voltage drop (per unit value)  $\Delta V > \Delta V_i$  on both sides of the switch branch is opened can the network loss be effectively reduced;
- 2) If the candidate switch can transfer the load from the feeder at the lower voltage side to the feeder at the higher voltage side, that is, the system network loss after opening the switch is less than the system network loss after opening the adjacent switch at the lower voltage side, the switch is the optimal switch.

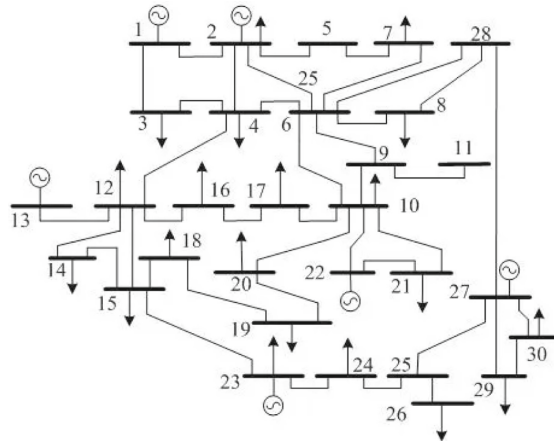
The reconstruction algorithm includes three processes: 1) initial optimization; 2) Secondary optimization; 3) Correction. The process is shown in Fig. 1. The initial optimization generates a better initial solution based on the minimum spanning tree algorithm, so as to reduce the calculation amount of the secondary optimization. The secondary optimization gets close to the optimal solution based on the improved heuristic rules, and the optimal reconstruction scheme can be obtained by modifying.

## 4 Distribution Network Reconfiguration Method for Optimal Distribution of DG Output

In this paper, DG reconfiguration algorithm is divided into two parts: firstly, DG output is optimally allocated based on the power demand of the grid and according to the generation cost and rated capacity of DG; Then, the reconstruction algorithm based on the minimum spanning tree and heuristic rules is used for reconstruction. In the algorithm, the output of DG is treated as a negative load. The improved 33 bus and 69 bus systems with DG are selected as shown in Fig. 2.



**Fig. 1.** Flow chart of distribution network reconfiguration



**Fig. 2.** 33 bus bar test system with DG

The results of the two algorithms are the same, which shows the effectiveness of the algorithm in this paper. In terms of calculation time, the calculation time of algorithm I is short, while algorithm II requires more switches to be evaluated and takes much longer, which shows that the initial optimization using the minimum spanning tree algorithm has an obvious effect on improving the calculation efficiency. In addition, in the initial

optimization process based on the minimum spanning tree algorithm, the weights are obtained based on the closed-loop power flow results, so the reconstruction results do not depend on the initial structure of the network [10].

The network structure and electrical distance are not considered in the DG optimal allocation calculation in this paper. Generally, the electrical connection of the distribution network is relatively short. From the analysis of the 33 node calculation example in this paper and the 69 node calculation example in the following, it can be seen that whether the network structure and electrical distance are considered or not has less significant impact on the network loss than expected, especially when the proportion of DG output to the total load increases, the difference between the two is very small. In fact, the impact of DG on network loss is mainly related to the selection of its initial location. The closer it is to the load center, the more obvious the loss reduction effect will be. Once the DG access location is determined, the impact of its output on network loss is not much related to the electrical distance. Therefore, based on the replication dynamic algorithm in this paper, the optimal allocation of DG output is solved, and the subsequent reconstruction algorithm is used to obtain the network operation architecture with the minimum network loss. The two layers not only ensure the high efficiency of calculation, but also ensure that the global optimal solution of reconstruction (or close to the global optimal solution) is obtained under the premise of optimal allocation of DG output.

## 5 Conclusion

In this paper, a distribution network reconfiguration algorithm considering the optimal distribution of DG output is proposed. First, the replicator dynamic algorithm RD is used to solve the DG output assignment problem with the minimum generation cost of DG, and then the minimum spanning tree algorithm and heuristic method are combined to reconstruct the distribution network to minimize the network loss. The initial optimization of the minimum spanning tree does not depend on the initial topology of the network. At the same time, based on the initial optimal solution generated by the minimum spanning tree and the new heuristic evaluation rules, the number of candidate switches to be evaluated is greatly reduced, and the computational efficiency of the algorithm is greatly improved. The example shows that the optimal reconfiguration solution can be obtained by reconfiguration of distribution network with DG based on the algorithm proposed in this paper. Due to the small number of evaluation switches, the algorithm has high computational efficiency.

## References

1. As, A., Ae, B., Ag, A., et al.: Reconfiguration of electrical distribution network-based DG and capacitors allocations using artificial ecosystem optimizer. Practical case study - ScienceDirect (2021)
2. Liao, J., Wang, R., Xia, S., et al.: Joint Planning of Distribution Network and Electric Vehicle Charging Station Considering Extreme Weather Events. IOP Publishing Ltd (2022)
3. Sun, C., Su, S.: Real-time coordination optimization method of active distribution network considering different stakeholders. IOP Conf. Ser. Earth Environ. Sci. **827**(1), 012024 (6pp) (2021)

4. Shi, Q., Li, F., Dong, J., et al.: Co-optimization of repairs and dynamic network reconfiguration for improved distribution system resilience. *Appl. Energy* **318** (2022)
5. Kim, Y.J.: Supplementary Feedforward Voltage Control in a Reconfigurable Distribution Network (2021)
6. Arivazhagan, T.: Optimal PMU placements using sea lion optimization for adaptable distribution system. *J. Measure. Eng.* **10**(1), 16 (2022)
7. Wenchuan, W.U., Zhang, B., Zihao, L.I., et al.: Method for Planning Distribution Network With Reliability Constraints Based on Feeder Corridor, US20210334429A1[P] (2021)
8. Imra, S.H., Kumar, R.M., Mahmood, S.A.: Optimal integration of battery energy-storage system with high penetration of renewable energy in radial distribution network. *Clean Energy* **3**, 3 (2022)
9. Mohseni, M., Joorabian, M., Ara, A.L.: Distribution system reconfiguration in presence of Internet of Things (2021)
10. Bogdanski, B., Johnsen, B.D., Zahid, F., et al.: System and method for efficient network reconfiguration in fat-trees, US11095498B2[P] (2021)



# A Study of Intersubjectivity in Translation Based on Communicative Behavior Theory and Computer Technology

Lu Liu, Nan Peng<sup>(✉)</sup>, and Jing Guo

Department of Basic Teaching and Research, East University of Heilongjiang,  
Harbin 150060, China  
11357860@126.com

**Abstract.** It is an effective way to overcome these two extremes of translator subjectivity from translation subjectivity to inter subjectivity. Habermas' communicative behavior theory provides a theoretical basis and methodological guidance for the study of inter subjectivity in translation. The author believes that the purpose of this book is to provide a new approach to the study of intersubjectivity in translation. The main idea of this book is to analyze the relationship between the translator and the target language, which can be described as part of the interactive process. This means that both parties are involved in an interactive process. In order to achieve effective communication, the two entities must understand each other's intentions and purposes. The inter subjectivity of translation can well explain the exertion of the translator's subjectivity and the effectiveness of the speech acts of the original and the author in the translation activities, and also ensure the effectiveness of the speech acts of the translator and the translated works.

**Keywords:** Communication behavior · Translation · Intersubjectivity

## 1 Introduction

In recent years, more and more attention has been paid to the subject and Intersubjectivity of translation, and the status of the translator has been gradually promoted, becoming the center of translation activities. Is the translator the only subject in the process of translation? Previous studies on Subjectivity in translation mainly focused on the translator's subjectivity, that is, the characteristics of the subject in the interaction between subject and object. However, subjectivity should also be manifested in the certain relationship between the labor subjects. The characteristics of the interaction between the subjects and the subjects are "inter subjectivity" or "inter subjectivity". Intersubjectivity is the communication between subjects, whose purpose is to break the closeness of possessive subjects [1]. "Habermas regards inter subjectivity as the key to solving three problems: without inter subjectivity, one cannot know whether a person is following a rule; without inter subjectivity, one cannot form" rule consciousness "nor develop"

principle consciousness “from” rule consciousness “and differentiate into” value consciousness; “Without intersubjectivity, we can not defend the legitimacy of the rules [2]. Translation is a communicative process and tool between two language societies. Its purpose is to promote the political, economic and cultural progress of the language society. Its task is to transfer the real world contained in the original works from one language to another without damage. As language is the medium of communication in life, translation should also be the medium of communication, and translation can actually be regarded as a kind of communication activity [3]. This is fully manifested in the inter subjectivity of translation. The study of intersubjectivity is closely related to the value of translation, and can also avoid the abuse of translation and mistranslation. Translation contains the possibility of coordinating the behavior between subjects. Through the mutual process between subjects, the process of language transformation reveals the intersubjectivity [4].

After the linguistic turn and cultural turn of philosophical studies, translation studies have also begun to study from the perspective of language and culture. With the opening of the cultural perspective, translation researchers no longer only focus on the subjectivity of the translator, but expand their research horizons to the author, the translator and the reader, even the publishers and sponsors, as well as the complex group relations among these subjects. Habermas’ theory of communication has a profound impact on the research of Humanities and Social Sciences, and also provides a new philosophical methodological basis for the study of inter subjectivity in translation. From the perspective of Habermas’ communicative behavior theory, this paper will study the principles and norms that should be followed by the various subjects in translation activities when they have group relations [5].

## 2 Related Work

### 2.1 Subject of Translation

As for the study of the subject theory of translation, the scholars of translation mainly focus on three important aspects, namely, the author centered, the text centered and the translator centered. In the introduction of the book translation introduction, it is pointed out that “it is generally believed that the subject of creative treason in literary translation is only the translator. In fact, in addition to the translator, the reader and the receiving environment are also the subject of creative treason. On the basis of comparison and generalization, Xu Jun proposed that” the translator should be regarded as the subject of translation in a narrow sense, and the author, translator and reader should be regarded as the subject of translation in a broad sense He believes that the translation process under the conceptual integration model involves at least five subjects, namely, the translator, the reader, the source language author, the source language and the target language involved in the text, and there is an inter subjective relationship between them [6]. Therefore, under the current research situation, there is no consistent answer to the question of determining the translation subject.

Translation circles generally agree that there are three subjects of Translation: the original author, the translator and the target reader. There is mutual communication and interaction between them. Wang Juan believes that the communication among the three



needs to be completed by some subjects acting as intermediaries, and puts forward two concepts: implied author and implied reader. She cited the concept of implied author put forward by Booth, who said that implied author was often higher than real author in intelligence and moral standards. The implied author, as the author's "second self", is not an ideal and impersonal "ordinary person". He is an implied double of "himself", and there may be many according to the needs of specific works [7]. Through the implied author, the communication between the author and the characters in the works is essentially the interweaving of his personal consciousness. Another concept is the implied reader, as the ideal reader of the author, which restricts his creation. The author's attitude towards the implied readers determines the author's writing strategy and its success.

In the practice of translation, it is a very important task for translators to coordinate the relationship between the subjects of translation. In different cultural backgrounds, conflicts and contradictions between the original author and the target readers are inevitable. The translator should strive to make the two sides communicate with each other in this environment. He should not only respect the original text, but also give full play to his subjectivity and creativity, and complete the task of cross lingual and cross-cultural communication [8].

## **2.2 The Inevitability of the Transformation from Translation Subjectivity to Inter Subjectivity**

According to the above analysis, we find that any centrism is one-sided, narrow and easy to go to extremes. The history of the evolution of the four kinds of centrism in translation studies is the history of one center replacing the other. The author centered theory attempts to reconstruct the author's intention, analyzes the author's motivation and other factors from a deep perspective, and completely ignores the subjectivity of readers and translators.

On the other hand, the text centered theory emphasizes the internal structural system of the text language unilaterally, attempts to establish a universal model of literary research by means of the linguistic model, and neglects the social and cultural nature of the language and the aesthetic psychology of the subject. The reader centered theory focuses on the process of readers' acceptance, less on the form and significance of the text of the works, the creative activities of the writers and their important position, giving people an impression that they ignore the ontological study of literary works, and in essence, it is still the continuation of the monism [9]. As for the translator centered theory, it attempts to completely eliminate the separation of subject and object in the "dualistic opposition", holding that there is no one-to-one absolute relationship between the signifier and the signified, which overemphasizes the subjectivity of the translator and leads to excessive interpretation of the text.

From the above analysis, it can be seen that the shift from subjectivity to intersubjectivity is the development trend of literary studies and translation studies. In translation, on the one hand, each subject element (author, reader and translator) should maintain their relative independence and avoid being dissolved in it. On the other hand, it requires that each subject element should not be self-centered and pursue egocentrism. The relationship among the elements of translation is no longer self-centered, separated and independent from each other, but a relationship of "co existence" (Heidegger language).

The subjective factors are a relationship of equal communication. They recognize each other, respect each other's subjective factors and coexist, and finally form a common subjectivity, that is, intersubjectivity. In other words, the intersubjectivity we understand refers to the state and characteristics of mutual recognition and harmonious coexistence among subjective factors. Only in this way can the subjective factors truly coexist peacefully and effectively play their respective roles.

Therefore, translation activities not only involve the original text and the translated works themselves, but also the smooth process of translation is the result of the interaction and coordination among the original author, the target author and the target reader. The study of intersubjectivity not only covers the uniqueness of each subject in translation activities, but also covers the common characteristics and interactions between subjects [10]. When applied to translation studies, the study of intersubjectivity should pay attention to the interaction between the author, the translator and the target reader as well as the coexistence of the subjectivity of the author, the translator and the target reader in translation activities. From the perspective of Habermas' communicative behavior theory, this paper will study the subjects involved in the translation process and the relationship between them.

### **3 A Study on the Intersubjectivity of Translation Based on Communicative Behavior Theory and Computer Technology**

#### **3.1 Two Extremes of Translation Subjectivity**

Both the linguistic translation research paradigm and the structural linguistic translation research paradigm focus on the author and the text, but cover up the translator's subjectivity. The philosophical basis of the paradigm of philological translation research is the philosophical theory of empiricism. The concept of meaning is referential theory and truth value theory. The influence of this theory on translation studies is to regard the author as the center of meaning and the relationship between the author and the translator as the relationship between the master and the slave. The translator must faithfully convey the author's purpose: the philosophical basis of linguistic translation research paradigm is rationalism or structuralism theory, and its semantic concept is relational theory and determinism. This theory regards the text as the center of meaning and emphasizes that the translation is faithful to the original, The translation is equivalent to the original, and the style of the original is reproduced. The study of translation in structuralist linguistics regards language as a closed and self-contained system, and regards meaning as something set by syntactic semantic laws. In this way, the translator's subjective factors involved in translation activities are excluded. Both of these paradigms require the translator to reproduce the content and style of the original text faithfully, and the translator's subjectivity is completely submerged. There are many metaphors about the identity and status of the translator in traditional translation theories. Tan Zaixi made a comprehensive summary of these metaphors. Yu Guangzhong compared the translator to a wizard between God and man. All these metaphors, without exception, show the cover of the translator's subjectivity.

"Deconstruction holds that the meaning of a text is open, intertextual and non original". The redefinition and interpretation of meaning by deconstruction directly overturns

the practice of structuralism to solidify and static meaning, and overturns the authority and absoluteness of meaning given by structuralism. In translation, people doubt the authority of the original text and the author. The skepticism tendency of deconstruction has also brought some negative effects, such as “denying all rationality, overemphasizing the subject consciousness, or simply banishing the subject”, “even when criticizing the certainty and clarity of the meaning of structural linguistics, it completely dispels the regularity of the language, and makes the interpretive activity become an endless interpretation and unlimited extension”. Here, due to the uncertainty of the meaning of the original text, if the translator does not grasp it properly, he will interpret the original text arbitrarily, abuse the translator’s rights, and lead to the excessive publicity of the translator’s subjectivity.

In a word, the post structuralist translation theory with deconstruction as the core excessively publicizes the translator’s subjectivity, which makes the translator’s subjectivity play freely in translation and completely subverts the concepts of original work, author and faithfulness in the structuralist translation theory. Moreover, both the concealment of the translator’s subjectivity and the excessive publicity of the translator’s subjectivity show the disadvantages of the translator’s Subjectivity: the closed single subject theory does not get rid of the “subject object” framework, because they focus on one part of the three subjects of the author, the target reader and the translator in the whole chain of translation activities, while ignoring the other subjects. Therefore, more and more studies are trying to break this closed and single subjectivity, so as to make the study of subjectivity of translation move towards the study of inter subjectivity of translation, as shown in Fig. 1.

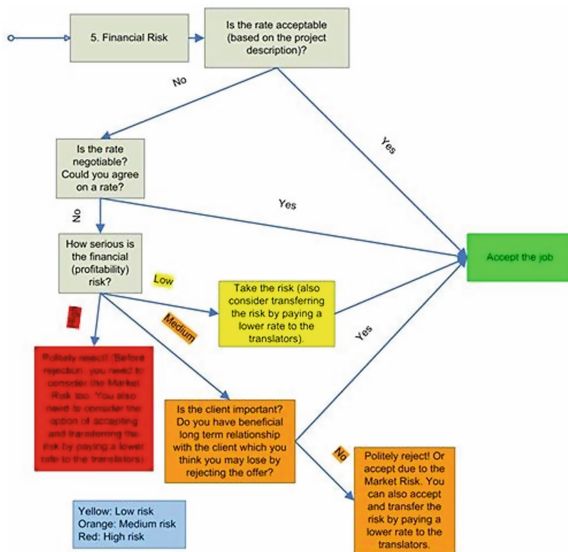


Fig. 1. A study on the intersubjectivity of computer technology in Translation

### 3.2 Communicative Behavior Theory and Intersubjectivity of Translation

How to overcome the above two extremes of translation subjectivity and make it reach an ideal intermediate state: that is, the translator's subjectivity can be reflected without being publicized? Habermas' communicative behavior theory affirms the recognizability of the speaker's subject in the process of communication and emphasizes the intersubjectivity in the process of communication, which is conducive to the equal communication between the translator and the author; Moreover, the theory of communicative behavior and communicative rationality have positive constructive significance and important implications for translatology. Therefore, this theory provides a philosophical and methodological basis for solving the above problems.

Language is the foundation of Habermas' communicative behavior theory. He first set about studying language and established his unique linguistic philosophy theory, universal pragmatics. Habermas believes that "the task of universal pragmatics is to determine and reconstruct the basic premise of communication"; The starting point of reconstructing universal pragmatics is communicative experience [understanding], "understanding refers to the meaning of discourse", "On the basis of the connections between the symbolic subjects established with other individuals, interpreters who take understanding meaning as their own duty are basically engaged in experience as reference persons in the process of communication". "Since the understanding experience is faced with a world built up by symbols, the inter subjective relationship between the understanding subject and the understood object has become the inter subjective relationship between the subjects who establish the world. Habermas takes" speech "as the basis of universal pragmatics, and is also the object of universal pragmatics.

## 4 Conclusion

As a new topic in the field of translation studies, the research results of inter subjectivity of translation can better reveal the essence of translation, explain the accuracy of translation strategies, highlight the value of subjects in translation activities, and put all subjects on the platform of equal interaction. Only in this way can the study of inter subjectivity of translation be further developed. Habermas' communicative behavior theory affirms the recognizability of the speaker's subject in the process of communication and emphasizes the intersubjectivity in the process of communication, which is conducive to the equal communication and interaction between the translator and the author. The inter subjectivity of translation refers to the interaction between the author, translator and target readers in the whole chain of translation activities in a specific context, and is the common existence of the subjectivity of the author, translator and target readers in translation activities. Therefore, it is undoubtedly the new trend of translation theory research from subjectivity to intersubjectivity.

**Acknowledgements.** 1. Research project of East University of Heilongjiang in 2020: Intersubjectivity Research in Translation from the perspective of practical Philosophy (Item No.: HDFKY200208).

2. Research project of East University of Heilongjiang in 2020: Novel Translation Studies from the Perspective of New Ecological Theory (Item No.: HDFKY200229).

3. Research project of East University of Heilongjiang in 2020: Application of ethnic Diasporic Theory in English Literary Criticism (Item No.: HDFKY200211).

4. Key Project of the 14th Five-Year Plan of Education Science in Heilongjiang Province in 2022: Research on Interactive Teaching Mode and Application of Online Courses from the Perspective of Cognitive Load (Item No.: GJB1422479).

5. One of the First Foreign Language Education Reform and Innovation Projects in Heilongjiang Province, Project: A Study on the Reform of Blended Foreign Language Teaching Mode from the Perspective of Curriculum Ideology and Politics (Item No.: HWX2022030-B).

## References

1. Latysheva, S.: Intersubjective approach to pre-translation analysis of architectural criticism: reaching after the balance. *J. Teach. Engl. Specif. Acad. Purp.* (2021)
2. Bi, Y., Zhang, D.: A study on female image alienation translation in Chinese classical literature based on computer corpus. *J. Phys. Conf. Ser.* **1915**(2), 022039 (2021)
3. Liu, M.: A study of the translation of 《three-body problem》 from the ambiguity in gender translation (2021)
4. Nanayakkara, A.: Adaptation in audiovisual translation; a study based on the Sinhalese version of a children's cartoon series "Ferdie, the ant". In: South Eastern University International Arts Research Symposium (2021)
5. Zhang, M., Huang, Z.: Crowdsourcing used in higher education: an empirical study on a sustainable translation teaching mode based on crowdsourced translation. *Sustainability* **14** (2022)
6. Isbister, K., Cottrell, P., Cecchet, A., et al.: Design not lost in translation: a case study of an intimate-space socially assistive robot for emotion regulation (2021)
7. Weng, M.Y.: Dynamic contextual adaptation in subtitle translation: a case study of red cliff (2) (2022)
8. Tang, Y.: A study on the English translation and introduction of YU Hua's novels from the perspective of medio-translatology (5). Science Publishing Group (2021)
9. Palmquist, A.: Lost in translation: a study of (mis)conceptions, (mis)communication and concerns when implementing gamification in corporate (re)training. In: 54th Hawaii International Conference on System Sciences, HICSS 2021 (2021)
10. Qin, R., Yu, C.: A Study of translation methods adopted in the yangs' version of "master gao" from the perspective of manipulation theory. *Int. J. Engl. Linguist.* (4) (2021)



# Analysis and Research on Children's Cognitive Psychological System Based on Digital Learning Terminal

Zixin Yang<sup>(✉)</sup>

North Sichuan Medical College, Nanchong 637000, Sichuan, China  
yangzx@sdhs.edu.com

**Abstract.** Children's understanding of things mainly depends on intuitive feelings, and their attention is extremely unstable. Allowing children to improve their attention when learning on the digital platform and achieve the purpose of teaching in fun is an important problem that needs to be solved for children's digital learning terminal. The analysis and research of children's cognitive psychological system based on digital learning terminal is the research on the operation mode of human mind. This is a process involving observation, analysis and synthesis. The main purpose of this process is to understand the working mechanism of the brain and make it work better. This helps improve the quality of life of people with mental disorders or any mental related problems. The first stage of this process involves observing what happens when a person does something with his brain. In most cases, these observations are made by using an electroencephalograph that records electrical activity.

**Keywords:** Children · Digital learning · Cognitive psychology

## 1 Introduction

With the rapid development of science and technology and the continuous improvement of people's consumption level, the products combining digital and mobile technology have been widely accepted by the public. Not only adults use the digital mobile platform to engage in various work, study, entertainment and life activities, but also more and



**Fig. 1.** Children use digital early education equipment

more children gradually begin to widely contact the digital mobile platform, which has become an increasingly obvious social phenomenon, and the scientization of education has also become a trend, As shown in Fig. 1:

The most basic learning of human beings is to absorb all external information through various sensory organs in the body, then transmit it through the nervous system, and then produce integrated information in the brain, so there is conscious cognition. After children have rich sensory perception, they gradually develop the ability of abstract thinking and independent thinking [1]. Preschool children begin to simulate and learn in the virtual world through these interactive media textbooks. Multimedia interactive teaching products that combine teaching with entertainment must be easy to learn and make users feel the satisfaction of learning. Early childhood education products generally have the nature of intelligence, enlightenment and education. Early childhood education products that integrate the advantages of digitalization have potential huge market demand and purchasing power in today's era. Compared with other digital products, at present, the vast majority of early childhood education products are based on multimedia CDs played by computers [2]. There are some problems, such as less development, insufficient attention, weak product interaction interface and children's interaction, limited by time, space and space, and lack of relatively mature and guiding principles and methods. The products based on digital mobile platform are not only small and portable, but also relatively flexible in time and place of use. Using the mobile digital technology platform to develop and design a good interactive interface that can promote children's early education, attract children's interest in learning, improve the learning effect, achieve the purpose of interactive learning effect, and correctly guide the education of preschool children has considerable exploratory and development value [3].

## 2 Related Work

### 2.1 Development of Children's Cognitive Psychology

Cognition is a process of continuous development. Under the influence of congenital factors and acquired environment, people's cognition of surrounding things will continue to change according to environment, experience, new knowledge and other factors. Because children are in the rapid growth period of individuals, their cognitive ability and characteristics develop and change more rapidly than adults. According to Piaget's division of children's cognitive development stages, their cognitive development is divided into four stages: perceptual movement, pre operation, specific operation and formal operation [4].

Children's cognitive development process is a process from sensibility to rationality. With the continuous growth of age, children's mind is constantly maturing, and their thinking and cognitive mode is also changing from simple to complex. According to relevant research, generally speaking, children over 5 years old have a deeper understanding of the visual guidance system, and the directional recognition system will have practical operational significance for children over 5 years old; For younger children, MI said, they often need to be accompanied by adults to identify directions. Combined with children's cognitive ability at different stages, the main age groups of children in children's space and the feasibility of the design of guidance system, this paper mainly analyzes

the cognitive characteristics of children aged 6–12. From the age of 6 to 12, children at this stage enter the specific operation stage, that is, they have made great progress in specific image thinking and logical reasoning, and their cognitive level of guidance has developed greatly [5]. Starting from the information processing model, this chapter analyzes the cognitive psychological characteristics of children's multi-channel senses and children's cognition, such as memory, thinking, imagination, curiosity, imitation and game psychology, through the form of dynamic development [6].

## 2.2 Digital Learning

Digital learning is the integration of digital technology and teaching courses, which refers to the establishment of digital learning platforms such as the Internet in the field of education to assist learning, also known as e-learning. The definition of the concept of digital learning by the U.S. Department of education can be summarized as follows: "Digital learning refers to learning and teaching activities mainly through the Internet. It makes full use of the learning environment provided by modern information technology, which has a new communication mechanism and rich resources, to realize a new learning method; this learning method will change the role of teachers and the relationship between teachers and students in traditional teaching, thus fundamentally changing the teaching structure and the essence of education". Digital learning originally refers to network-based learning, but in practical applications, digital learning based on multimedia materials is usually included in the scope of digital learning [7].

Digital learning refers to the process in which learners use digital learning resources to learn in a digital way in a digital learning environment, as shown in Fig. 2. It contains three basic elements: digital learning environment, digital learning resources and digital learning methods. Among them, digital learning environment mainly refers to the requirement to be equipped with multimedia computer, Internet and other hardware equipment; Digital learning resources refer to multimedia materials that can be operated in multimedia computer or network environment after digital processing, including digital video, digital audio, multimedia software, website, online learning system, computer simulation, data files, databases, etc. [8], and can find and process information through user autonomy, cooperation, creation and other ways; Different from traditional learning methods, digital learning methods use digital platforms and digital resources to carry out consultation, discussion and cooperative learning between professors and learners, and learn through the collection and utilization of resources, exploration, discovery, creation and display of knowledge.



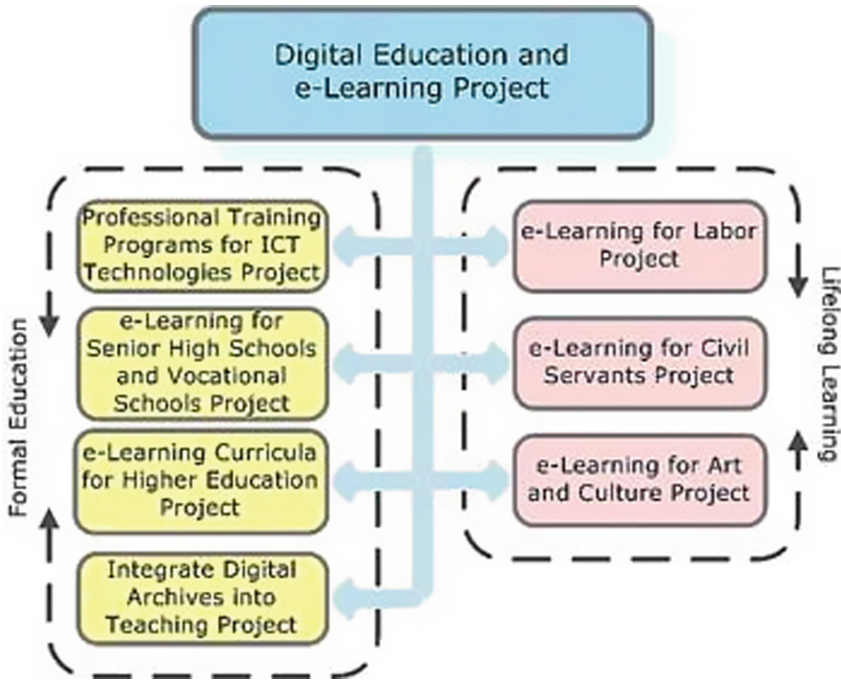


Fig. 2. Digital learning mode

### 3 Children's Cognitive Psychological Characteristics

#### (1) Memory characteristics

Children's memory is based on children's senses, which is the mapping and accumulation of the world felt by sensory organs. Compared with preschool children, children aged 6–12 years old have a qualitative transition in memory development, showing the following characteristics: (1) from the dominance of mechanical memory to the dominance of understanding memory; (2) From unconscious memory to conscious memory; (3) From the dominance of specific image memory to the gradual growth and development of abstract memory of words. Children of this age begin to have logical reasoning, systematization and goal clarification in their memory. Therefore, in the corresponding guide design, its role in the transition of children's memory development should be strengthened, and the guide should be more in line with children's memory development characteristics by using elements such as graphics, symbols and colors with logical relevance. As shown in Fig. 3, the familiar things and scenes of children are transformed into concrete graphics after cartoon processing [9]. This graphic with relevance to reality enables children to recall the meaning they express at the first time they see it, so as to understand the logo information they want to convey.



Fig. 3. Design of child identification system

(2) Thinking characteristics

Children's thinking ability is reflected in the ability to summarize, compare, classify and solve problems. Under the guidance of school teaching, school-age children begin to gradually have the ability of abstract thinking. Psychologist zhuzhixian proposed that the basic feature of pupils' thinking development is the gradual transition from concrete image thinking to abstract thinking; But this kind of abstract thinking is directly related to perceptual experience to a large extent, and still has a large component of concrete visualization. In the design of children's space guidance system, the transmission of information should be intuitive and direct, and different information should have obvious differences, so that children can quickly find and understand the corresponding guidance information.

(3) Curious psychological characteristics

Children's curiosity about the unknown makes them more likely to be attracted by things they don't understand than adults. This curiosity inspires children to explore the world around them. When children are in a public space, excellent psychological system design can convey a spatial information that arouses children's desire to explore, and also reduce children's tension about unfamiliar space.

#### 4 Analysis of Children's Cognitive Psychological System Based on Digital Learning Terminal

Children's concept of the world is closely linked with "activities". In the eyes of young children, everything that can move is alive and alive. On the one hand, children use

“moving and immovable” to distinguish the “living and inactive” countries of objects; On the other hand, moving things can always easily attract children’s attention and interest. This is because the cognitive development in early childhood is dominated by unintentional attention, and the development of unintentional memory is better than intentional memory. Children around 3 years old especially like to watch advertisements. They can watch interesting advertisements for 10 min without moving [10]. It can be seen from this that children’s attention is always clearly selective, and they can even accept the information that interests them. They are characterized by strong curiosity, rich imagination and specific thinking image.

According to the above analysis and summary, a children’s application software that combines teaching with fun is designed. The overall style is simple and generous, the interface layout (as shown in Fig. 4) is reasonable and effective, the colors are bright and the saturation is high, and the interactive experience is rich and colorful.

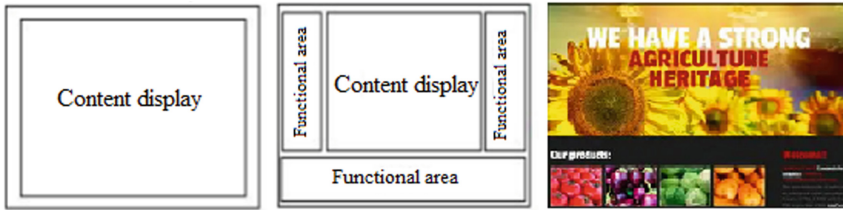


Fig. 4. Schematic diagram of interface layout structure

As shown in Fig. 3, the interface layout is optimized by taking advantage of the operation advantages of the software itself, so that the window area can be maximized and more content can be displayed. At the same time, children’s attention will not be distracted by the beating of multiple windows. The position of the main buttons is fixed, so that it is easier for children to operate, and they don’t have to spend energy on memorizing and looking for button position information alive.

## 5 Conclusion

By studying and understanding the cognitive development characteristics of preschool children and the location of children’s attention, this paper summarizes children’s cognitive characteristics from the aspects of shape, color and attention, so that children can be liked visually in the interface design, so that children can no longer enter the world because of the restrictions of text symbols, and turn learning into a participatory process, making the learning of teaching in fun more natural and easier.

**Acknowledgements.** Interpretation of College Students’ Psychological Painting and Research on Psychological Intervention (No. CSXL-202BO3).

## References

1. Dvoryatkina, S.N., Zhuk, L.V., Smirnov, E.I., et al.: Didactic model of development of research activities of schoolchildren in a hybrid intellectual learning environment (2021)
2. Xiaoxian, S., Meicheng, W.: Spatial color analysis of kindergarten interior design based on children's psychological activities in digital environment, p. 05011 (2021)
3. Su, J.: Scene matching method for children's psychological distress based on deep learning algorithm. *Complexity* (2021)
4. Xiang, Y., Cipriani, A., Teng, T., et al.: Comparative efficacy and acceptability of psychotherapies for post-traumatic stress disorder in children and adolescents: a systematic review and network meta-analysis. *Evid. Based Ment. Health* (4) (2021)
5. Luo, P., Zhang, X.: Simulation of psychological course satisfaction based on android mobile system and neural network. *Microprocess. Microsyst.* **81**, 103751 (2021)
6. Qi, Z., Hui, H., Bo, C.F, et al.: Analysis of cognitive impairment in schizophrenia based on machine learning: interaction between psychological stress and immune system. *Neurosci. Lett.* (2021)
7. Nickel, B., Copp, T., Brennan, M., et al.: The impact of breast density information or notification on women's cognitive, psychological, and behavioral outcomes: a systematic review. *JNCI J. Natl. Cancer Inst.* (2021)
8. Li, X., Karuppiah, M., Shanmugam, B.: Psychological perceptual analysis based on dance therapy using artificial intelligence techniques. *Int. J. Artif. Intell. Tools* (2021)
9. Degenhart, L., Zonatto, V., Lavarda, C.: Effects of psychological capital and managerial attitudes on the relationship between budgetary participation and performance (2022)
10. Sang, H.: Analysis and research of psychological education based on data mining technology. *Secur. Commun. Netw.* **2021**(7), 1–8 (2021)



# Analysis of E-Commerce Live Broadcast Marketing Strategy Based on Big Data Algorithm

LinLin Sun<sup>(✉)</sup>

School of Economics and Management, Hainan College of Vocation and Technique,  
Haikou 570216, Hainan, China  
15120666209@139.com

**Abstract.** E-commerce live broadcast marketing strategy analysis based on big data algorithm is a new method to analyze consumer behavior and trends, which can be used to analyze consumer preferences and make accurate predictions. The analysis method is based on big data algorithms, such as machine learning and artificial intelligence. It can be used for many purposes, such as analyzing website traffic or social media activities. The analysis process is very simple: first, collect all the necessary information about your enterprise, including its name, URL, email address and phone number. Then, the system analyzes this information and shows you its impact on business performance. Through analysis, we can predict future consumer behavior with high accuracy. This technology has been applied in many industries, such as finance and insurance, tourism, retail and so on. This technology will bring more opportunities for enterprises in these fields.

**Keywords:** big data · Marketing strategy · E-commerce live broadcast

## 1 Introduction

Webcast, a new interactive communication method, has produced a wonderful ‘chemical reaction’ after being combined with e-commerce, which has brought surprising sales results. In April last year, Maybelline New York invited spokesperson Yang Ying and 50 netizens to broadcast live online at the same time. More than 10000 lipsticks were sold in two hours of live broadcast, which translated into an actual sales of 1.42 million yuan. This sales performance shocked the e-commerce industry. Everyone witnessed this “sales miracle” and saw the great potential of live broadcast marketing [1]. Then many e-commerce enterprises introduced this new way to carry out e-commerce marketing. Today, with the increasingly fierce market competition, e-commerce companies hope to use the east wind of live broadcasting to drive sales. Old e-commerce companies want to use the live broadcasting platform to continue to tap the consumption potential and further expand the market. New businesses also hope to gain a place in the market by combining live broadcasting.

E-commerce live broadcasting is like the next outlet for the development of e-commerce. E-commerce enterprises have come in droves and joined the live broadcasting

army. Although this marketing method has only been introduced for a short time, the economic benefits it can bring have been unanimously recognized by e-commerce. Marketing and promotion with the help of live broadcasting platform will be the focus of e-commerce development in the next period. As e-commerce live broadcasting is a new thing, it is still in the stage of exploration and development [2]. With the development of this marketing method, problems have gradually emerged. For the sustainable and healthy development of e-commerce live broadcasting, the author combs and analyzes the current e-commerce live broadcasting marketing strategy, points out the existing problems, and puts forward optimization paths, so as to provide some reference for future development.

The new way of information exchange created by webcast and the user resources gathered by the platform have great commercial value for e-commerce, and the webcast platform has also become the focus of e-commerce marketing. The use of live webcast platform for effective marketing communication, import traffic for e-commerce, promote the realization of traffic, realize the unity of communication benefits and economic benefits, and become a new business growth breakthrough for e-commerce [3]. At present, major e-commerce companies such as tmall, vipshop and Jumei premium products have started the live broadcast mode. Therefore, it is particularly important to study the marketing strategy of e-commerce on the online live broadcast platform.

## 2 Related Work

### 2.1 Research on E-Commerce Marketing Strategy

At present, the research on e-commerce marketing in China is mostly focused on the practical application of marketing communication, and there is less theoretical research. A more prominent aspect is the marketing research of mobile e-commerce. With the development of mobile Internet technology, mobile e-commerce occupies an increasingly important position in the e-commerce market. At present, empirical research on mobile e-commerce is mainly focused on. For example, in the mobile e-commerce marketing strategy, Yang JianZheng believes that to strengthen the marketing of mobile e-commerce, we should consider four aspects: creating demand, highlighting characteristics, strengthening publicity, and developing projects. The article "Research on the application of mobile e-commerce interactive marketing mode" by Liao Weihong summarizes four main models of mobile e-commerce interactive marketing, including SMS website, mobile QR code, mobile business district, mobile search, etc. [4].

With the development of e-commerce, the market share of cross-border e-commerce continues to expand and has become an important part of the e-commerce market. The marketing research of cross-border e-commerce mainly focuses on brand marketing and marketing strategy marketing. Chen Huan pointed out in the article "an analysis of the independent brand marketing strategy of cross-border e-commerce enterprises" that at present, China's cross-border e-commerce is developing rapidly. Behind the rapid growth, problems such as price competition and product homogenization are also highlighted. This article analyzes the competitive advantages brought by independent brand marketing to cross-border E-commerce, and puts forward brand marketing strategies for the sustainable development of cross-border e-commerce [5]. Chen Zhizhong pointed

out in the “analysis of viral marketing strategy of cross-border e-commerce platforms” that viral marketing has efficient information dissemination, explosive communication process and effective marketing effect. As a powerful marketing method that conforms to the trend of the Internet era, viral marketing will definitely contribute to the development of cross-border e-commerce platforms.

## 2.2 Research on Webcast Platform

For the research on the problems of webcast platform, this kind of research discusses the communication problems, ethical problems and legal copyright problems in webcast from the perspective of media criticism and legal system. For example, Li Junxian’s “the formation and elimination of barrage language violence mechanism in webcast”, the article believes that the instability of the communication role and interaction relationship in the webcast platform, the concealment of the media ecology of the webcast platform and the lag of the management of the webcast platform, etc., lead to the problem of barrage language violence. Purifying the barrage language environment of the webcast platform requires the joint efforts of operators, anchors and audiences, Nor can it leave the close cooperation of all social parties in today’s Internet environment. In “cold thinking under the upsurge of network live broadcasting platform in the era of Pan entertainment”, Huang Yi believes that the entertainment characteristics of mass media make the over consumed live broadcasting show pleasing the public with vulgar forms and contents, which is reducing people’s cultural taste. At the same time, consumerism and pan entertainment are disintegrating the traditional social and cultural value system. In the cold thinking of the hot webcast, Zhao Yuqiao discussed the ethical vacancy of the anchor, netizens and live broadcasting platform involved in webcast from the perspective of ethics, and made a specific analysis of all parties involved in standardizing webcast from the perspective of media ethics [6]. Zhou Gaojian’s “analysis of the legal protection of the copyright of online game live broadcast” discusses the ownership of rights in the process of online game live broadcast from the perspective of jurisprudence, and believes that we should clarify the right attributes of online game live broadcast programs, clarify the ownership of rights, and strengthen the legal protection of online game live broadcast platforms.

Case analysis of webcast. This kind of research mainly analyzes the characteristics of webcast based on specific cases, and reveals the communication value and marketing value of webcast platform. In “talking about the marketing advantages of webcast - Taking Durex’s interactive marketing as an example”, Wu Desheng pointed out that the linear broadcast of webcast platform can obtain continuous attention, and interactive communication improves the degree of information involvement.

At present, the research on the webcast platform is mostly focused on the platform value and the disputes brought by the webcast. The economic value contained in the webcast platform has attracted people’s attention and recognition. How to develop the webcast platform is still a process that needs continuous exploration for people. There is less research on the current e-commerce webcast marketing, which is more concerned. This paper enriches the research on this topic by discussing the e-commerce webcast marketing methods.

### 3 New Situation of E-Commerce Marketing - E-Commerce Live Broadcast

E-commerce marketing activities came into being with the emergence of e-commerce. The popularity of mobile Internet and the continuous improvement of people's consumption level have made online shopping an indispensable part of people's daily life. Old e-commerce companies should consolidate their market position, while new e-commerce enterprises continue to join in, which has intensified the competition in the e-commerce industry. Major e-commerce companies have launched various marketing activities, event marketing A series of marketing activities such as festival marketing are emerging one after another, especially after tmall created the "double 11" sales myth, China's e-commerce enterprises are more than happy with festival marketing. There are more than 20 e-commerce shopping festivals. Facing the "frenzy" of various e-commerce marketing activities, the consumer market is becoming increasingly tired. When the promotion is weak, e-commerce began to explore a new marketing tool - e-commerce live broadcast [7].

#### (1) Development of live broadcast platform

The emergence and development of the live broadcast platform are inseparable from online games. In 2005, the online game Warcraft fire, huanju times company was optimistic about business opportunities, and then launched YY voice for players to make voice calls in the team. YY voice can be called the earliest prototype of the domestic live broadcast industry, and then joined various functions such as singing and chatting. YY live broadcast was launched. At the time of YY voice fire, live broadcast platforms such as 9158 and six rooms were launched one after another, which was well known to the public. By 2011, the popularity of e-sports game hero League directly promoted the development of webcast. In the next few years, live broadcasting platforms have sprung up. With the development of mobile Internet, Yingke, Zanthoxylum bungeanum, Yizhu, douyu and other mobile live broadcasting platforms have been launched one after another. At one time, the network live broadcasting platform has been in flames. 2016 was called the first year of mobile live broadcasting [8]. According to incomplete statistics, the number of live broadcasting platforms has exceeded 200, and the number of platform users has exceeded 300 million.

#### (2) Driving force of e-commerce webcast marketing

Under the background of increasingly fierce competition and gradual disappearance of e-commerce dividends, live broadcast marketing has opened a new marketing door for e-commerce. Interactivity, entertainment and sociality are the remarkable characteristics of the live broadcast platform. The interaction between users and live broadcast generally includes on-site interaction and off-site interaction. The relationship between the two is based on mutual trust. There is a strong relationship link between the anchor and users, which plays a great role in promoting the realization of traffic. At the same time, the entertainment and sociality of the live broadcast platform meet the personalized needs of consumers. Marketing activities on the webcast platform are of great significance to the development of e-commerce.



With the continuous development of mobile Internet technology and the wide rise of social networks, netizens have obtained channels for information dissemination and interaction, which provides technical support for the rise of netizens; In terms of public demand, the consumption concept of the public is constantly changing, and young groups pursue the value identity in the virtual environment and the cultural context. Online celebrity economy is a unique economic phenomenon in the Internet age [9]. The realization of online celebrity economy is to turn fans into purchasing power. The relationship between online celebrities and fans is based on trust. Because of their fashion sensitivity, online celebrities are easy to demonstrate and lead others and form others' imitation objects. The combination of online celebrity influence and business produces online celebrity economy, as shown in Fig. 1.

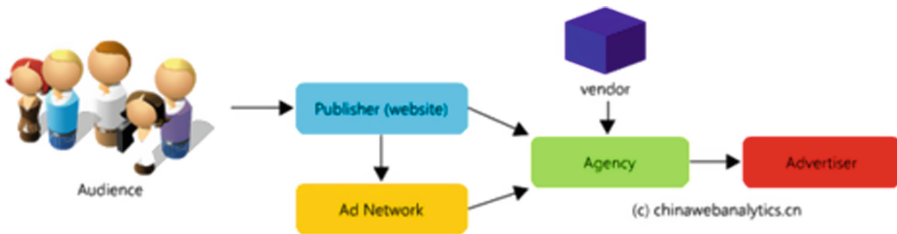


Fig. 1. Online industry chain

#### 4 E-Commerce Live Broadcast Marketing Strategy Based on Big Data Algorithm

- (1) Online marketing to create entertainment shopping  
The entertainment of shopping mode is to infiltrate the entertainment elements into the shopping process of consumers, form a new shopping mode, and let consumers obtain a happy consumption experience. On the one hand, the rise of shopping entertainment is due to the increasingly fierce competition in the supplier market. At present, the competition in the e-commerce industry is fierce. In addition to established e-commerce enterprises, some rising stars are also constantly pouring into this field. In order to develop in such a fierce competitive environment, it is bound to seek new ways to innovate supply. Cross border e-commerce live broadcasting with its entertainment elements is just a good breakthrough. With the development of economy and the continuous upgrading of people's consumption demand, traditional functional consumption can no longer meet people's consumption demand, and people are more inclined to spiritual consumption and personalized consumption. The continuous upgrading of consumer consumption demand forces the e-commerce industry to continue to innovate. For e-commerce, the addition of entertainment elements optimizes users' shopping experience and not only meets users' consumption demand, It also enables users to obtain spiritual pleasure [10].

## (2) Scenario marketing stimulates consumers' shopping needs

The word "Scene" was originally often used in the concept of film and television, "refers to a certain character's action in a certain time and space or the specific life picture formed by the character's relationship. In short, it refers to a group of continuous shots taken in a single place". The scene in scenario marketing is composed of three core elements: people, things and fields, while technical elements, space-time elements, subject elements and social elements play a role around these three core elements. Scenario marketing is a new marketing concept and marketing method, which is developed on the basis of traditional advertising marketing. In fact, scenario marketing is a marketing activity aimed at the needs of consumers in the real environment. This demand is more psychological and exists in the subconscious of consumers. Scenario marketing is a means to stimulate consumers' shopping needs, as shown in Fig. 2.

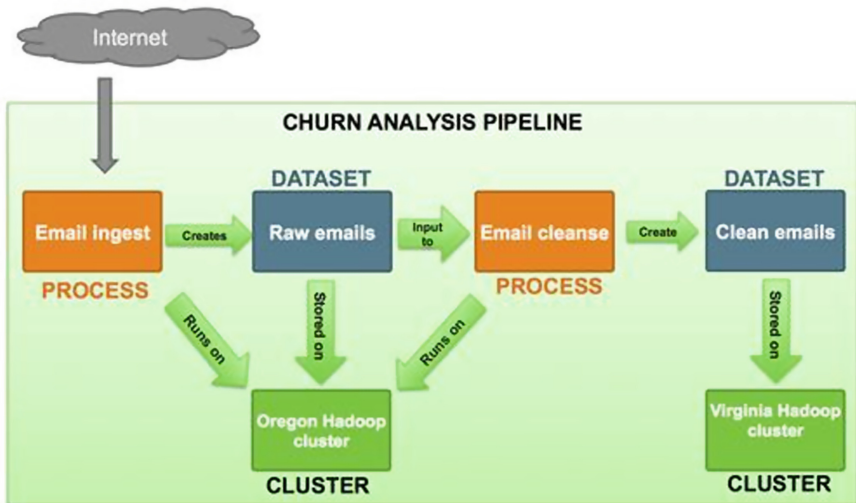


Fig. 2. Scenario marketing inspires consumers to shop

## (3) New user purchase experience to improve traffic conversion rate

The conversion rate refers to the ratio of the number of completed conversion behaviors to the total number of clicks of the promotion information in a statistical cycle. The traffic conversion rate is the ratio of the number of web page clicks to purchasing power. "For e-commerce, traffic is customer traffic, which is an important indicator to measure the comprehensive ability of e-commerce. Traffic is the lifeline of e-commerce. In order to attract customers to the store, e-commerce carries out publicity and promotion through online media, social media and other channels, opens up traffic portals, and hopes to win traffic dividends. With the development of e-commerce industry and the continuous progress of mobile Internet technology, online shopping has become the norm for consumers, and consumers have become more and more Add 'picky', personalized consumption is more obvious,

and they are gradually tired of the almost normal e-commerce marketing methods. For e-commerce, traffic is not equal to the purchase volume, especially in today's extremely competitive market, consumers are facing more and more choices, and using effective publicity strategies to attract traffic is far from enough. How to provide personalized services and give consumers a new shopping experience plays a great role in improving the traffic conversion rate and promoting the realization of traffic.

## 5 Conclusion

With the development of mobile Internet, the e-commerce market continues to expand, and e-commerce has penetrated into all aspects of people's life. People can buy their favorite goods without leaving home. It breaks the boundaries of time and space, and has a revolutionary impact on people's production and life. The e-commerce industry is developing rapidly and its scale continues to expand. The huge business opportunities contained in the e-commerce market make the e-commerce competition increasingly fierce. E-commerce live broadcast has innovated the user experience. E-commerce live broadcast marketing is a major innovation in e-commerce marketing activities. Carrying out e-commerce live broadcast marketing has opened a new marketing door for e-commerce. Although it is currently in the initial stage of development, this model has brought huge dividends to e-commerce, and its economic value has been widely recognized by the e-commerce industry. E-commerce live broadcasting is bound to become a key development direction of e-commerce in the future.

**Acknowledgement.** Hainan Province Education Science 13th Five Year Plan Project in 2019: Research on multi-level personnel training system of vocational skills based on Logistics Management 1 + X certificate pilot system, project number: QJY20191065, project leader: Sun Linlin.

The General Project of Higher Education Teaching Reform in Hainan Province in 2020: Practical Research on the Integration of Courses and Certificates in the Pilot Work of "1 + X" Certificate System in Vocational Education — Taking the Pilot Work of "1 + X" Vocational Skill Level in Logistics Management as an Example project leader: Dong Jiajia.

## References

1. Fei, M., Tan, H., Peng, X., et al.: Promoting or attenuating? An eye-tracking study on the role of social cues in e-commerce livestreaming (2021)
2. Xue, J., Li, J., Han, Y.: Evaluation and emotional analysis of mobile phone sales of JD e-commerce platform based on LDA model. *J. Phys. Conf. Ser.* **1861**(1), 012076 (2021)
3. Hu, X., Liu, J.: Research on e-commerce visual marketing analysis based on internet big data. *J. Phys. Conf. Ser.* **1865**(4), 042094 (2021)
4. Wieting, M., Sapi, G.: Algorithms in the marketplace: an empirical analysis of automated pricing in e-commerce. Working Papers (2021)
5. Dreossi, T., Ballardini, G., Gupta, P., et al.: Analysis of E-commerce ranking signals via signal temporal logic (2021)

6. Zhou, M., Huang, J., Wu, K., et al.: Characterizing Chinese consumers' intention to use live e-commerce shopping. *Technol. Soc.* **67** (2021)
7. Cielo, A.C., Raiol, T., Silva, E., et al.: Implementation of the e-SUS primary care strategy: an analysis based on official data (2022)
8. Kubrusly, J., Neves, A.L., Marques, T.L.: A Statistical analysis of textual e-commerce reviews using tree-based methods **12**(3), 16 (2022)
9. Jiang, H.: Strategy decision making in e-commerce: a case study of Amazon (2021)
10. Zhang, Z., Yang, Z.: Application analysis of e-commerce web design based on cinema 4D technology. *J. Phys. Conf. Ser.* **1848**(1), 012155, 4 p. (2021)



# Analysis of Flow Data of E-Commerce Project Incubation Park Based on Computer Algorithm

Xingrong Zhang<sup>(✉)</sup>

Entrepreneurship College of Chengdu Polytechnic, Chengdu 610041, Sichuan, China  
rr20220106@126.com

**Abstract.** The development of computer technology and the increasing popularity of the Internet have led to the rapid development of e-commerce, the continuous emergence of e-commerce enterprises and the doubling of the number of users, resulting in the rapid rise of multi-source traffic data of e-commerce websites. The traditional e-commerce traffic data processing method stores limited data and supports few data source formats. It is urgent to introduce a modern e-commerce traffic big data processing system to meet the needs of enterprise traffic data processing. The flow data analysis of e-commerce Project Incubation Park Based on computer algorithm is a method to analyze and classify the transactions of online shopping centers. The main purpose of this project is to study the transaction mode, frequency and amount in a specific time period. This analysis will help to determine the efficiency level of a specific area or department and can be used as a benchmark for improvement. Compared with traditional methods such as manual calculation or mathematical formula, computer algorithm has many advantages. They are easy to use, fast and accurate solutions.

**Keywords:** Computer algorithm · E-commerce projects · Incubation Park · Flow data

## 1 Introduction

In recent years, China's e-commerce has developed rapidly. With the increasing popularity and development of emerging technologies such as the Internet of things, cloud computing and mobile terminals, e-commerce, as a new business model, is rapidly integrating with the real economy. It has become an important way to allocate resources under the conditions of networking, informatization, internationalization and marketization, and has a profound impact on people's daily production and life, Become an important force leading economic and social development and progress [1]. By the end of 2015, there were about 600 million Internet users in China, of which nearly 300 million had done shopping on e-commerce websites. The number of large and small e-commerce enterprises has exceeded 10000. The huge market potential has attracted more and more traditional retail enterprises to develop to e-commerce. However, this also makes the competition in the e-commerce industry increasingly fierce. E-commerce enterprises should fully understand their own development status, respond in time when

the market changes, and adjust quickly when their own development is limited [2]. The traffic data of e-commerce websites is an important basis for the strategic deployment of each e-commerce enterprise.

With the rapid development of small and micro enterprises and small and micro enterprise incubators, some problems that can not be ignored have also been exposed. The growth of quantity has not brought about the synchronous improvement of quality, and small and micro enterprises are still facing the grim reality of low survival rate. According to the survey report of the Democratic Progressive Central Committee, the average survival cycle of small and micro enterprises in China is only half that of European and American countries. Compared with developed countries, China's small and micro enterprise incubation parks still have a big gap in incubation quality, incubation efficiency and sustainable development [3]. The phenomenon of emphasizing form over substance is common, mostly staying in the provision of infrastructure such as venues, lacking professional team management, effective service system and support mechanism; A large number of small and micro business incubators rely too much on government investment, lack their own "hematopoietic function" and have a deep-rooted awareness of "finding the government when it is difficult", resulting in the disadvantages of single investor, inconsistent rights and responsibilities, unclear income model, inadequate financing service function and so on [4]. The above problems have affected the entrepreneurial success rate of small and micro enterprises to a certain extent, and have become an important factor restricting the efficient operation and stable construction and development of China's small and micro enterprise incubators. There is still a certain gap between the incubation quality and efficiency and the original intention of the establishment of the incubators.

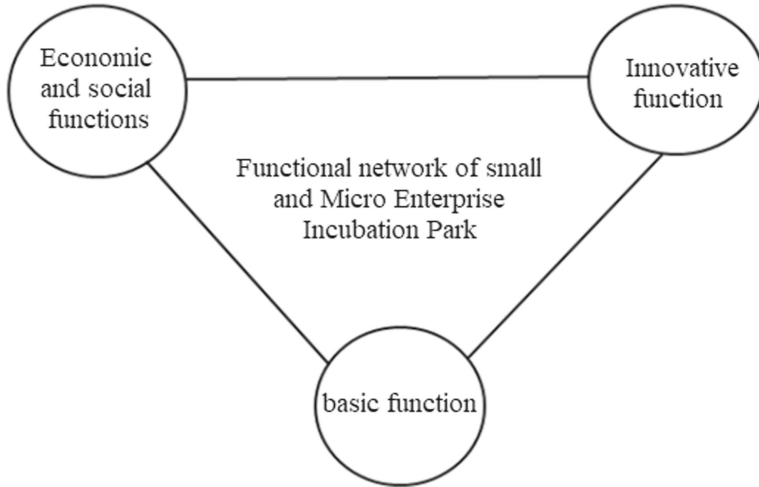
## 2 Related Work

### 2.1 Definition of Incubator

In European and American countries, the incubator is defined as an organizational form that can provide various services to start-ups in the early stage to improve the ability of enterprises to adapt to the market, which can greatly improve the success rate and survival rate of enterprises. The American incubator Association defines an incubator as a place to support the growth of enterprises and provide services in terms of capital, management, technology and market. In 2001, American scholars divided small and micro business incubators into three stages according to their functions. In the - stage, small and micro business incubators mainly provide hardware facilities, including plants, office space, water and electricity, etc.; The small and micro Business Incubation Park in the second stage is based on the first stage and adds consulting and guidance services; In the third stage, the small and Micro Enterprise Incubation Park has added the related functions of financing and venture capital on the basis of the first two stages. In 2003, Peter engberts, a famous Dutch scholar, divided the responsibilities of the government and the incubator, believing that the government only acts as a guide to provide financial and policy support for the incubator and start-ups, while as an independent form of economic organization, the incubator should be operated and managed in the form of enterprises [5].

The Ministry of science and technology of the people's Republic of China positioned the Incubation Park as a special service institution for cultivating small and medium-sized enterprises. Through its own resources, the Incubation Park provides business sites, infrastructure and basic services required for entrepreneurship for enterprises participating in the incubation, so as to reduce the entrepreneurial threshold, reduce capital consumption, reduce the risk of failure, improve the entrepreneurial success rate of enterprises, and cultivate entrepreneurs with high comprehensive quality at the same time. As early as 2000, Fu Chunmei, a scholar, defined the Incubation Park as a new form of organization, between the enterprise and the market. The main function is to provide convenient hardware facilities and consulting guidance for start-ups, and provide various services to promote entrepreneurial success according to the specific needs of enterprises, so as to reduce the entrepreneurial cost investment and improve the success rate of enterprises.

In 2007, sun Mengyao, a scholar, positioned the small and micro Business Incubation Park as a sharing platform founded by the government. It has both system guarantee and technical service functions, and is used to help the growth of small and micro enterprises, so as to reduce the investment of capital, site, infrastructure and consulting services in the initial stage of entrepreneurship, and reduce the entrepreneurial pressure of small and micro enterprises [6]. In 2010, scholar Zhao Guanbing defined the small and Micro Enterprise Incubation Park as a new economic organization form focusing on public welfare, which is mainly invested and built by the government, coordinates the participation of all social forces, provides preferential places, public infrastructure and professional training services, helps the development and growth of start-ups, reduces the entrepreneurial threshold of small and micro enterprises and shortens the entrepreneurial cycle. Liu Guihua pointed out that in 2012, the main body of microenterprise incubation not only needs the help of its own professional incubation facilities, but also needs the help of the outside world to realize the sustainable development of microenterprises. The main function of the business incubator is to provide the entrepreneurial entities with the hardware facilities and high-quality shared services required in the entrepreneurial stage, so as to reduce the entrepreneurial threshold and improve the success rate of entrepreneurship [7]. The business incubator mainly has three functions, as shown in Fig. 1:



**Fig. 1.** Small and micro enterprise incubation and functional network diagram

## 2.2 Operation Management Mechanism

Operation mechanism refers to the way and principle of mutual connection, interaction and interaction between various elements within the organization and between the organization and the external environment in the process of organization operation. It organically connects various parts in a certain operation mode to make them operate in a coordinated and orderly manner, so as to achieve the goal of the organization. It is a specific operation mode to coordinate the relationship between various parts and elements. Specifically, it is the design of incubation action process and system. Through the process design of the interaction action between various elements, we can improve the work efficiency in the incubation process, and standardize the behavior of each subject through scientific and effective system, so as to make the incubation process more orderly [8].

Management mechanism refers to the internal structure of the management system and the interaction mechanism between various elements. The implementation subject of the management mechanism is the management system and its elements. Under the guidance of the operation principle, coordinate the efficient linkage between various parts and timely adjust the factors that are not conducive to the normal operation of the system, so as to ensure the realization of the function of the whole management system. The management mechanism must adapt to the internal structure and function of the organization, interact and restrict each other. On the one hand, the form of organizational structure determines the type of management mechanism. The adjustment of organizational structure will inevitably lead to the change of the type and effect of management mechanism. On the other hand, the management mechanism has a reaction to the organizational structure. Once the management mechanism is formed, it will actively regulate and restrict the organizational structure and organizational behavior according to certain laws.



### **3 Computer Algorithm**

#### **3.1 Algorithm Concept**

The concept of algorithm has existed for a long time, but so far there is still a lack of a generally recognized definition. Intuitively speaking, the algorithm is a continuous sequence of rules executed in a certain order in order to complete a specific task. In its 2016 report, OECD defined an algorithm as a precise list of simple operations, which are mechanically and systematically applied to a set of tokens or objects (such as the configuration of chess pieces, numbers, cake ingredients, etc.). The initial state of the token is input, and the final state is output. “Algorithms can be embodied in languages, charts, codes and other forms, and various programs used in daily life also reflect the application of algorithms. With the development of computer science, algorithms have been developed to automatically process repetitive tasks including complex operations and data processing [9]. The latest development of artificial intelligence and machine learning has brought the algorithm to a new level, enabling computers to solve complex problems and make predictions and decisions more effectively than humans.

#### **3.2 Application of Algorithm**

The application of algorithms is becoming more and more extensive. Both government departments and enterprises are actively using algorithms to predict, analyze and optimize business processes, so as to support business decisions. One of the most commonly used functions of the algorithm is to price services or products. For example, many online trading platforms have been using automatic sales - price determination algorithms for several years. These platforms allow settled merchants to distinguish the market through dynamic pricing. This pricing method is also widely used in tourism, hotel reservation, retail, sports and entertainment. Since businesses can also use algorithms to optimize prices based on available inventory and expected demand, price algorithms also dominate the online sales of many goods. Computer algorithms are also reported in the insurance industry. For example, the so-called “market factor” algorithm used by Allstate is to optimize the price by measuring the possibility of price comparison before users buy insurance. The use of this algorithm has been criticized because it promotes a selective pricing method that deviates from risk, making the premium range possible from 90% to 800% of the standard price.

### **4 Analysis of Flow Data of E-Commerce Project Incubation Park Based on Computer Algorithm**

#### **4.1 Reshape the Strict and Objective Entry Screening Mechanism**

Strategic choice theory emphasizes the important role of choice in enterprise operation and management. The funds, venues and human resources of the small and Micro Enterprise Incubation Park are relatively limited, and the contradiction between the scarcity of resources and the incubation needs of small and micro enterprises always exists. The problems of blind introduction and “no rejection” in the admission screening of the

park directly affect the effective realization of the strategic positioning of the park and exacerbate the tension between demand and resources. To solve this problem, we must adhere to the theoretical guidance of strategic choice, implement differentiation strategy and specialization strategy, select entrepreneurs and entrepreneurial projects with growth potential as incubation objects, highlight the advantages of county resources, and achieve “no one has me, no one has me”, so that limited resources can produce higher returns.

The small and Micro Enterprise Incubation Park should make full use of the production and service elements of the park, improve the transparent, perfect and fair entry screening criteria, establish a rigorous and objective evaluation team, timely conduct comprehensive evaluation, review and decision-making on the entrepreneurial subjects applying for entry, one enterprise, one policy and one letter, sign the entry agreement, clarify the responsibilities, rights and interests of both parties, and standardize, guide and restrict the behavior of both parties. The selection of incubation objects directly affects the realization of the incubation function of the park. Therefore, careful selection of incubation objects is the top priority to ensure the quality and efficiency of park operation.

#### **4.2 Improve the Comprehensive and Efficient Cultivation and Incubation Mechanism**

Operation and management theory reveals that operation and management activities are the core functions of economic organizations and the guarantee mechanism to realize the transformation from input to output. As an emerging economic organization, small and Micro Enterprise Incubation Park must strengthen the planning, organization, implementation, supervision and inspection of the operation process of the park in order to ensure that the invested space, funds, policies, services, information and other resources can quickly and effectively promote the growth of small and micro enterprises [10]. Understand the development objectives, growth needs, weaknesses and weaknesses of incubating enterprises, coordinate and solve the problems in the development of funds, training, talents and information, effectively control the incubation quality, cost investment and time, run flexible management through the whole incubation process, and effectively improve the entrepreneurial quality and competitiveness of enterprises, so as to feed the park and society and achieve the organic unity of quality, efficiency and benefit.

- (1) Apply for support from special funds and subsidies. In order to break the limitation that the current assistance funds of the small and Micro Enterprise Incubation Park are invested solely by the county government, it is suggested that the park focus on combing the national policies and special funds related to encouraging entrepreneurship and innovation and promoting supply side structural reform, make full use of the existing policies, and help the small and micro enterprises settled in the Incubation Park to apply for special funds and various financial subsidies set up by the state, autonomous regions and autonomous prefectures to promote the development of small and micro enterprises. Support the development of incubated small and micro enterprises in the park by means of free financial assistance, loan discount

and post entrepreneurship subsidy, so as to turn the national preferential policies and financial support into the development power of small and micro enterprises. We will improve supporting financial funds, raise social funds, set up a small and micro enterprise incubation industry development fund, and invest in small and micro enterprises in advantageous industries in the early and medium-term.

- (2) Broaden financing channels. In view of the fact that the economic foundation of small and micro enterprises is generally weak and the effectiveness of the government's single "blood transfusion" is limited, we should actively open up financing channels to solve the capital dilemma and activate the enterprise's own "hematopoietic" function with multi-channel "blood transfusion". We will build microfinance companies and technology and financial service centers, strengthen the financial connection between banks and enterprises and venture capital, build a financing platform for small and micro enterprise incubation parks, attract banks, microfinance companies and venture capital companies to participate, broaden financing channels for small and micro enterprises, and provide financial support for incubating enterprises in the early stage of entrepreneurship and lack of funds by means of loans and equity investment. The park should actively coordinate various financial entities, expand the scope and amount of credit loans for small and micro enterprises, expand the collateral scope of mortgage loans, and carry out intellectual property mortgage services; Actively coordinate the cooperation between insurance companies and banks and other financial institutions, introduce property insurance to serve small and micro enterprises, and carry out loan guarantee insurance business for small and micro enterprises, that is, for small and micro enterprises that have financing needs but can not provide guarantors or corresponding amount of mortgages and pledges, apply for loan liability insurance from insurance companies. In case of overdue repayment or failure of repayment of loans by small and micro enterprises, The insurance company shall bear the liability for compensation or advance payment according to the limit agreed in the insurance contract to solve the problems of difficult credit and guarantee.

## 5 Conclusion

Combining theory with practice, this paper puts forward a scheme to optimize the operation and management mechanism of small and Micro Enterprise Incubation Park. In view of the problems and shortcomings existing in the operation and management mechanism of the small and Micro Enterprise Incubation Park, combined with the specific local reality, and focusing on the four mechanisms, this paper highlights the responsibilities of the government of the investor, the park of the specific management service party and the settled small and micro enterprises, and puts forward improvement measures and optimization schemes, in order to give full play to the "baton" function of the government in mastering macro-control Promote the incubation capacity and incubation quality and efficiency of small and micro business incubators, stimulate the original driving force of small and micro enterprises in the park, improve production capacity and profitability, improve the success rate of entrepreneurship, and realize the co construction, win-win and sharing of enterprises, parks, society and government.

## References

1. Bonetto, S., Mosca, P., Vagnon, F., et al.: New application of open source data and rock engineering system for debris flow susceptibility analysis. *J. Mt. Sci.* **18**(12), 3200–3217 (2021)
2. Zippo, R., Stea, G.: Computationally efficient worst-case analysis of flow-controlled networks with network calculus (2022)
3. Tirelli, I.: Data-driven analysis of flow visualization (2020)
4. Kolaczyk, E.D., Csárdi, G.: Analysis of network flow data (2020)
5. Hu, X., Liu, J.: Research on e-commerce visual marketing analysis based on internet big data. *J. Phys. Conf. Ser.* **1865**(4), 042094, 6 p. 2021
6. Pinasang, I.G., Tulung, J., Saerang, R.T.: The analysis of e-commerce retargeting strategy toward student purchase intention (2020)
7. Smojver, V., Torga, M., Zovak, G.: Exploring knowledge flow within a technology domain by conducting a dynamic analysis of a patent co-citation network. *J. Knowl. Manag.* (2020). ahead-of-print (ahead-of-print)
8. Li, B., Garicano-Mena, J., Zheng, Y., et al.: Dynamic mode decomposition analysis of spatially agglomerated flow databases. *Energies* **13** (2020)
9. Yu, R., Wu, C., Yan, B., et al.: Analysis of the Impact of big data on e-commerce in cloud computing environment. *Complexity* **2021**(2), 1–12 (2021)
10. Hu, J., Ouyang, T., Wei, W.X., et al.: How do manufacturing enterprises construct e-commerce platforms for sustainable development? A case study of resource orchestration. *Sustainability* **12**(16), 6640 (2020)



# Analysis of Teaching Management Data Based on Decision Tree Algorithm

Dan Yu (✉)

Liaoning Urban Construction Technical College, Shenyang 110122, Liaoning, China  
baby-lucky@hotmail.com

**Abstract.** Teaching management data analysis based on decision tree algorithm is a method of analyzing data to identify the most important factors. Any type of statistical method can be used for analysis. The results of this analysis will help make decisions, such as which teachers should be retained, which teachers should be transferred, and how many resources each teacher needs. The decision was based on data collected through observation, questionnaires and other surveys. Logistic regression is a process used to analyze the relationship between two variables. Decision tree is used to represent data and find out which variables are directly related to each other. It can be used to find out which variable affects another variable and how it does so. This analysis is helpful to understand the relationship between variables and to use statistical methods to predict future results.

**Keywords:** Decision tree algorithm · Teaching management · Data analysis

## 1 Introduction

In fact, the monitoring and analysis of teaching quality in the Teaching Department of colleges and universities and the monitoring and analysis of teaching quality in distance continuing education are inseparable from the university teaching research data analysis system. Analysis of the learning situation of daily teaching in Colleges and universities [1]. It is usually carried out after the end of the teaching stage, but it often ignores the correlation analysis of students' usual class attendance, homework, learning attitude, partial subjects and other factors closely related to students' performance.

The university teaching and research data analysis system is a data service for all members of the university teaching organization (including managers, teachers, students, etc.). Based on big data, it analyzes and solves the “why” and “how” problems, and gets through all links of the teaching cycle, such as daily teaching, teaching quality evaluation, semester examination, performance analysis, etc. through analysis, presentation and data mining, it can drill down the data and analyze the image data The abnormal early warning shall be carried out in-depth, layer by layer and in-depth insight, so as to achieve the objectives of teaching and personnel training, and assist the education and teaching reform.

Through the big data system, more valuable comprehensive information of teachers and students can be collected, which is convenient for the school to master the information of all staff, and it is also an important support for students to carry out self-management. On the other hand, the big data analysis system can save the data so that it is not easy to lose it, which provides convenience for reuse for a long time in the future. Due to the characteristics of cyberspace, big data analysis can save everyone's information independently and find it quickly. Compared with the traditional file management mode, it is convenient and fast, and it is an important way to promote the school's data-based information management.

## 2 Related Work

### 2.1 Decision Tree Algorithm

Classification is a very important task in data mining, which is most widely used in business. The purpose of classification is to learn a classification function or classification model (classifier). The model can map the data in the database to a given category. Both classification and regression can be used for prediction. The purpose of prediction is to automatically derive the extended description of the given data from the historical data records, so as to predict the future data [2].

To construct a classifier, we need a training sample data set as input. The training sample data set, also known as the training set, is composed of data records. Each record contains several attributes to form an eigenvector.

$$e_j = -k \sum_{i=1}^n f_{ij} \ln f_{ij} \quad (1)$$

$$W_j = d_i / \sum_{j=1}^m d_j \quad (2)$$

Each record of the training set also has a specific class label corresponding to it. This class label is the input of the system, usually some previous experience data. A specific sample can be in the form of a sample vector:  $(V, v \dots v_a: C)$ . Here,  $V$  represents the field character, and  $C$  represents the value of the class label attribute. The purpose of classification is to analyze the input data and find an accurate description or model for each class through the characteristics reflected in the data in the training set. This description is often represented by predicates. Figure 1 below shows the data characteristic model.

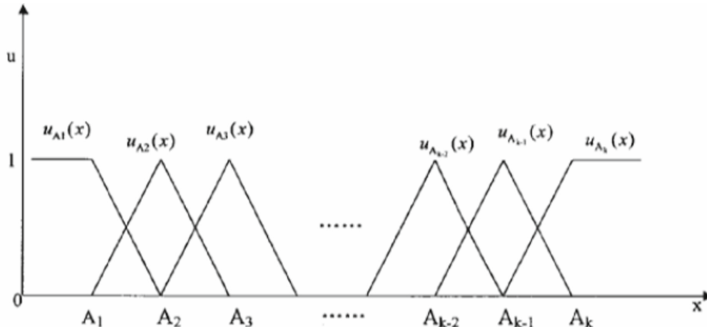


Fig. 1. Data characteristic model

The generated class description is used to classify the future test data. Although the class label of these future test data is unknown, the class of these new data can still be predicted. We can also have a better understanding of each class in the data.

The classification model of decision tree is widely used because of its unique advantages. First, users do not need to know a lot of background knowledge [3]. As long as the training cases can be expressed in the form of attribute conclusion, they can learn with the algorithm; Secondly, the decision tree model has high efficiency, which is more suitable for the large amount of data in the training set; Thirdly, the classification model is a tree structure, which is simple and intuitive. It can convert the path to each leaf node into rules in the form of if-then, which is easy to understand; Finally, the decision tree method has high classification accuracy. The goal of classification is to accurately describe the class or build a model by analyzing the data in the training set, and then use it to classify other data in the database or upgrade it to classification rules.

### 2.2 Analysis of Teaching Management Data

Big data analysis can conduct intelligent multi-dimensional analysis on students' classroom performance and examination conditions on the basis of collecting students' overall learning data, and provide score evaluation for students through the given evaluation system to help students provide accurate data basis for later behavior adjustment and planning. The application of big data analysis to the teaching management system can help students provide more comprehensive information and data analysis, so as to help students form targeted plans and arrangements for learning according to their interests and hobbies, so as to provide necessary help for students' learning [4]. In addition, big data analysis can also analyze the basic situation of the school from the data content obtained at ordinary times, such as teachers, venues and course query, It can give students more information they want to obtain, and cooperate with the operation of the teaching system with the instantaneous and fast data system. It is also a necessary guarantee for more efficient management of the school.

In education, especially in school education, data has become the most significant indicator of teaching improvement. Generally, these data mainly refer to teaching resources, teachers and students, etc. if these data are recorded and compared in a paper

way, it is inevitable that there will be mistakes and omissions, resulting in decisions that sometimes affect the development of the school.

The Ministry of education has issued a notice that colleges and universities are gradually implementing grid management. Many colleges and universities have begun to implement grid management. ABI is also completing this project, and the information of each department needs to be effectively combined [5]. However, now colleges and universities basically have their own independent systems for each department. The data standards are inconsistent, and the systems cannot be related. The management of teachers and students in Colleges and universities is a huge and abstract work. The arrival of the information age has brought new methods and means to this work. Using big data to manage teachers and students can not only fundamentally solve the massive consumption of human and financial resources, but also make the analysis more accurate and provide a more accurate basis for decision-making.

### **3 Analysis of Teaching Management Data Based on Decision Tree Algorithm**

First of all, we should understand the core functions of the “university teaching and research by data analysis system”.

1. Group dynamic portrait  
From the perspective of dynamic development and change, data are used to dynamically describe the group characteristics of “students” and “teachers”, so as to provide the latest analysis and decision-making basis for “students’ teaching in groups” and “teachers’ management in groups”.
2. Dynamic situation tracking and early warning  
Pay attention to the learning situation, track each student’s real-time attendance rate, class participation, homework completion rate & excellent rate, predict the academic trend, give early warning of homework plagiarism, and give early warning of failing classes, so as to help the school adjust the teaching process.
3. Teaching quality evaluation and tracking  
Regular teacher-student evaluation, peer evaluation, quality quantification and trend comparison, positive incentives for teachers and early warning of abnormal teaching quality [6].
4. Performance analysis and early warning  
Multi dimensional insight into teaching problems, examination paper proposition analysis, continuous improvement, and verification of the effectiveness of decision-making measures.

1) Proposition quality analysis. It supports the analysis of single volume and single question Differentiation & change, difficulty & change, score rate and average level, wrong question difficulty and knowledge. 2) Student performance analysis. Students’ individual scores in various subjects in class, grade position, change trend, as well as knowledge point strength analysis, partial subject analysis, academic prediction and actual result comparison. 3) Teacher performance analysis. The overall situation of



students in the class (score rate, analysis of answers to difficult questions, difficulty of wrong questions - Analysis of knowledge points), comparison with other classes, etc. 4) Performance analysis of college and school level managers [7]. Multi dimensional data insight (the public courses/core courses are analyzed in terms of subjects, average score distribution, disciplines, colleges, teachers and classes to gain insight into the essence).

Then, we will look at the value of “university teaching and research by data analysis system”.

1. The whole process of data. Get through the teaching data cycles and links such as daily teaching, teaching quality evaluation, semester examination, score analysis, etc.
2. Full role. Provide data application scenarios for all role members (students, subject teachers, head teachers/counselors, subject principals, school level and school level managers, etc.) of the university teaching organization [8].
3. Data visualization and insight. Through analysis, presentation and data mining, data can be drilled down, and abnormal early warning can be stripped of its cocoon, deepened layer by layer, and deep data insight can be carried out.
4. Anytime, anywhere. Users can master teaching trends and abnormal warnings anytime and anywhere, support wechat business notification, approval to do, and real-time viewing of analysis data.
5. API integration. Provide API interface to facilitate data docking with educational administration system and third-party system.

## 4 Simulation Analysis

Data are generally obtained through scientific experiments, verification, statistics and other means, and are used for scientific research, decision-making and other purposes. Then, through comprehensive, accurate and systematic collection, classification and storage of these data, and strict statistical testing of these data, we can get very convincing conclusions [9]. When making big data, it is necessary to strictly carry out scheme design and statistical test, otherwise the data obtained is meaningless and of little value.

In education, especially in school education, data has become the most significant indicator of teaching improvement. Generally, these data mainly refer to teaching resources, teachers and students, etc. if these data are recorded and compared in a paper way, it is inevitable that there will be mistakes and omissions, resulting in decisions that sometimes affect the development of the school. It is probably the simplest building tutorial, and can be completed with a little Python foundation. With the help of open source tools, the mature system that has been developed for several months can be completed within one hour.

An excellent data analysis platform must first meet the functions of data query, statistics, multi-dimensional analysis, data report, etc. Unfortunately [10], in their first year of work, many analysts are buried in SQL statements and complete their work in the form of sql+excel, but they do not use efficient tools. Figure 2 below shows the data analysis setup code.

```
def giniimpurity_2(rows):  
    total = len(rows)  
    counts = uniquecounts(rows)  
    imp = 0  
    for k1 in counts.keys():  
        p1 = float(counts[k1])/total  
        imp+= p1*(1-p1)  
    return imp
```

Fig. 2. Data analysis setup code

## 5 Conclusion

Teaching management data analysis based on decision tree algorithm is a technology used to analyze the performance of school teachers. School administrators and teachers use it to understand their efficiency in their work. This analysis helps them make decisions about future employment or reemployment. The technology uses a set of rules that apply to data collected from teachers' performance during their tenure in the school. These rules are helpful to identify teachers' behavior patterns and propose improvement methods.

## References

1. Zhang, Z, Zhao, Z., Yeom, D.S.: Decision tree algorithm-based model and computer simulation for evaluating the effectiveness of physical education in universities (2020)
2. Zhao, Y.: Research on the application of university teaching management evaluation system based on Apriori algorithm. *J. Phys. Conf. Ser.* **1883**(1), 012033, 6 p. (2021)
3. Gao, K.: Evaluation of college English teaching quality based on particle swarm optimization algorithm. In: CONF-CDS 2021: The 2nd International Conference on Computing and Data Science (2021)
4. Li, J.: Application of intelligent fuzzy decision tree algorithm in English teaching model improvement. *Complexity* (2021)
5. Diao, S.: The reform of teaching management mode based on artificial intelligence in the era of big data. *J. Phys. Conf. Ser.* **1533**(4), 042050, 7 p. (2020)
6. Qomariyah, N.N., Heriyanni, E., Fajar, A.N., et al.: Comparative analysis of decision tree algorithm for learning ordinal data expressed as pairwise comparisons. In: 2020 8th International Conference on Information and Communication Technology (ICoICT) (2020)
7. Xin, Y., Wang, C., Dong, Y., et al.: Management and entrepreneurship management mechanism of college students based on support vector machine algorithm. *Comput. Intell.* **IP**(IP), 1–13 (2020)
8. Niu, Q.: Optimization of teaching management system based on association rules algorithm. *Complexity* **2021** (2021)
9. Xu, S., Liang, L., Ji, C.: College public sports culture practice based on decision tree algorithm. *Pers. Ubiquitous Comput.* **24**(7) (2020)
10. Huang, X.: Analysis of public physical education teaching and quality evaluation in colleges and universities based on decision tree algorithm. In: WAIE 2020: 2020 2nd International Workshop on Artificial Intelligence and Education (2020)



# Analysis on Reasonable Planning and Design of Port Waters and Channels

Wenting Liu<sup>(✉)</sup> and Zhiwei Liu

Power China Zhongnan Engineering Corporation Limited, Changsha, Hunan, China  
WentingLiu@msdi.cn

**Abstract.** In recent years, the scale of operations in the shipping industry has continued to expand, and the demand for the expansion of port waters and waterways has continued to increase. The previously used port waters and waterways have to bear more load. This article will specifically plan and analyze the scientific and reasonable design of port waters and waterways. On the premise of shipping stability, the previous design scheme of port water channel is improved, which provides a certain design reference for the scientific and reasonable planning of port water channel in the future.

**Keywords:** Port · Waterway Design · Planning

## 1 Introduction

Aiming at the shipping status of the East No. 1 berth of Douwei port area of Meizhou Bay Port, this paper makes a planning analysis on the scientific and reasonable improvement and reconstruction of the channel, improves the previous port water channel design scheme on the premise of shipping stability, and provides a certain design reference for the scientific and reasonable planning of the port water channel in the future. East No. 1 berth located in zoumali planned port area in the north of Qinglanshan in Meizhou Bay port. With the increasing scale of ships, the navigation conditions can not meet the existing development requirements. This article carries out relevant planning and design for the waters of the East No. 1 port to improve the navigation safety and navigation efficiency of ships in the port waters.

## 2 Overview of Waterway Design

The navigable capacity of the port channel is related to the transport potential of the port water channel, and the land element involved is mainly the service rate of the ship terminal. The main land elements involved are the service rate of the ship terminal, and the water area elements include the channel, berthing area, anchorage, landing point, dock berthing water area, whirling water area, etc. [1]. In the design concept of port and waterway, the following specific standards must be followed: first, to ensure the stability of the passage of ships; second, it is better to select the straight-line layout

method that can directly enter the port for channel network setting; third, appropriate widening shall be carried out on the overall scale of the necessary channel to reduce the subsequent expenditures and maintenance costs as much as possible; fourth, minimize late expenditure and maintenance costs; fifth, comprehensively consider the environment around the port and protect the ecological environment [2].

### **3 Requirements for Reasonable Planning and Design of Waterways**

#### **3.1 The Problem**

According to the analysis of investigations carried out by Chinese ports today, in terms of port construction, the role of ports is relatively single, and the number of ports with comprehensive and multi-purpose is far from enough, it can not meet the needs of today's continuously improving port excellent service and port independent innovation [3].

#### **3.2 The Influence of Geological Conditions and Natural Environment on the Route**

When designing and planning the port and waterway, it is necessary to comprehensively consider the geological conditions of the planning site and the intervention and influence of natural factors on the overall planning of the route, such as hydrology and meteorology, basin water depth, etc., and use the existing location of natural water resources to make overall planning for the channel as much as possible [4, 5].

#### **3.3 Key Points of Channel Axis Placement**

In the actual design of the port channel axis, it is necessary to avoid the emergence of S-shaped bends, and try to ensure that the channel axis is as smooth as possible to ensure the stable entry and exit of the ship when actually passing, and avoid problems such as the ship touching the wall due to turning [6].

In the planning process, When the channel axis is close to lighthouses, piers, observation platforms and other buildings, it is required to comprehensively consider the impact of wind speed, water flow and other factors on the potential flow transformation near the buildings, which will lead to yaw when passing through, in the design concept, the adjustment of the distance between the channel sideline and the building shall be considered, and the distance shall not be less than 3.5 times of the building.

### **4 The Theory of Reasonable Planning and Design of Port Waters and Channels**

The design of the port waterway system is essentially the composite result of several factors that interact and influence each other in port transportation. From the perspective of differences, the design factors of the port production system can be divided into two major factors: land area and water area; from the basic function From the perspective of classification, the port waterway system can be divided into production and operation factors and auxiliary factors [7].

## 5 Specific Design of Port Waters and Channels

### 5.1 Route Selection

No. 1 berth is located in the planned port area of Zoomali in the northern part of Qinglan Mountain in Meizhou Bay Port. With the increasing size of ships, the navigation conditions of the port can no longer meet the existing development requirements. Therefore, it is necessary to carry out correlations to the port waters and waterways. Planning and design [9].

Under the influence of realistic conditions, assuming that the port channel cannot be leveled and the channel has to be turned, it is necessary to strictly control the size of the turning direction to ensure that the turning direction is adjusted within  $30^\circ$ . If the actual conditions are still not satisfied, then the overall width of the channel and the turning radius must be enlarged to reduce the stress effect of the ship during the turn [10] (Fig. 1).

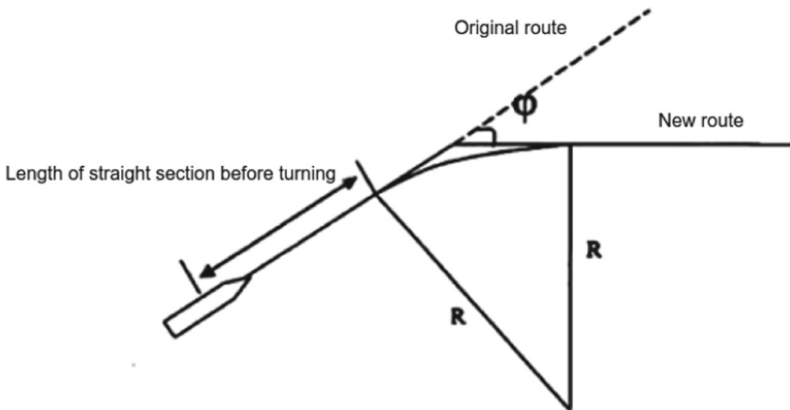


Fig. 1. Channel axis turning

According to the local objective natural conditions, the selected channel axis shall be considered. The position deviation of a large ship under the action of inertia is also large. Therefore, the selection of the route must be reduced and the current speed  $\geq 1$  km/h. The river courses of h cross each other.

### 5.2 Determining the Width of the Channel and the Track Zone

The determination of the total width of the channel mainly includes real ship observation methods, ship handling simulation device experiments, etc. Now, the single channel in the waters of East No. 1 Port is widened to dual channels, so that the density of ships entering and leaving the port every day exceeds that of all-day shipping. At the time, the frequency of ships is 90.0% as the index value, which is calculated based on the total width of the track zone, the total width of the bottom of the ship's channel, and the

total width of the ship. The formula for the reasonable total width of a single channel is calculated as:

$$W = A + 2c \tag{1}$$

The calculation formula for the effective width of the dual channel is:

$$W = 2A + b + 2c \tag{2}$$

In the formula,  $W$  refers to the navigable width of the channel,  $A$  refers to the width of the track zone,  $c$  refers to the surplus width between the ship and the bottom line of the channel, and  $b$  refers to the surplus width between ships. The basic dimensions of the channel design are shown in the figure below (Fig. 2).

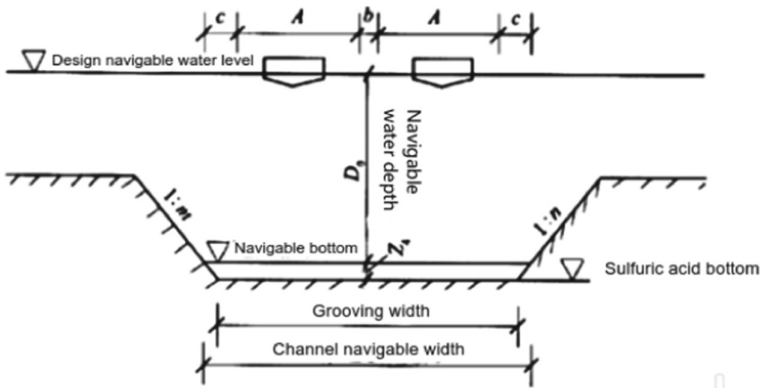


Fig. 2. Basic dimensions of channel design

When the ship is passing through the channel, to ensure that the new route is consistent with the direction of the river, many practical problems have to be solved, such as the wind speed on the river, the flow rate of the river itself, and the speed of the propeller of the ship. Once the ship drifts, the ship will appear. The situation of swinging from side to side. In order to control the course of the ship in the channel and minimize the impact of wind and current on the ship, the width of the track band is generally calculated using the deflection angle of wind and current pressure. The calculation formula of the track width  $A$  is:

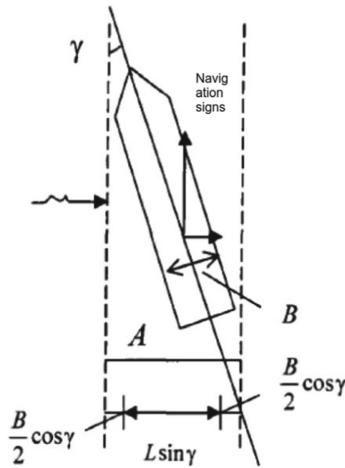
$$A = n(L\sin \gamma + B) \tag{3}$$

where  $A$  is the width of the track zone,  $n$  represents the multiple of the propagation drift, the  $B$  represents the designed boat width, and  $\gamma$  represents the deflection angle of wind pressure. There is still a certain relationship between the designed ship width and wind current pressure deflection angle. The specific relationship is shown in the table below (Table 1).

**Table 1.** The relationship between the designed ship width and wind current pressure deflection angle

Wind force	Cross wind $\leq 7$ 级				
Crossflow $V$ (m/s)	$V \leq 0.10$	$0.10 < V \leq 0.25$	$0.25 < V \leq 0.50$	$0.50 < V \leq 0.75$	$0.75 < V \leq 1.00$
$n$	1.81	1.75	1.69	1.59	1.45
$\gamma(^{\circ})$	3	5	7	10	14

According to related calculations, when planning the width of the track belt of East No. 1 Port, the width of the track belt is set between  $2.0B$  and  $4.5B$  (Fig. 3).



**Fig. 3.** Wind current pressure deflection angle and ship's track direction

### 5.3 Channel Water Depth Determination

The determination of the channel depth is mainly to meet the minimum safe depth specified for the full-load draught navigation of the planned ship under certain conditions. It is affected by factors such as water level, ship navigation, loading, and channel bottom quality. Therefore, the determination of the channel depth Specifically, it is the comprehensive consideration of ship type full draught value, hull sinking value when the ship is sailing, the lowest sufficient depth under the keel when sailing, the sufficient depth of waves, and the sufficient depth of the ship's loading trim. The design and planning of the channel depth is the depth of the shipping water The sum of the depth of wealth.

## 6 Evaluation on Reasonable Planning of Port Waters and Channels

In order to better perform objective assessment of the port waters and waterways, we can start with the safety of the ship navigation system in the port waters. The ship navigation system can be regarded as an organic whole combining people, ships, and water environment, from the perspective of the relationship between the navigation environment and the ship. Research on the aspects; natural conditions directly and indirectly affect the operation of the ship, and the channel can also use the shallow water effect to control the hull position and restrict the ship's movement speed and orientation.

## 7 Conclusion

Port waters and waterways are a key component of the ship's navigation process. Its overall planning and design should be built on the basis of corresponding laws, regulations and norms. It is analyzed in combination with various factors such as hydrometeorology, geographical elements, navigation weather, safety management and so on, scientific and reasonable route selection, route depth control and route water depth measurement. The above elements evaluate the safety of port water routes from two aspects: safety, reliability and danger, and give certain reference to the safe and stable navigation of ships, and then design and plan better port water routes.

## References

1. Guan, S.: Analysis of reasonable planning and design of port waters and channels. *Sci. Chin.* (17) (2017)
2. Luo, H.: Research on reasonable planning and design of port waters and waterways. *Pearl River Water Transp.* **000**(012), 31–32 (2014)
3. Yu, J.: Research on Reasonable Planning and Design of Port Waters and Channels. Dalian Maritime University (2008)
4. Xu, J.: Analysis on the planning, design and safety of port waters and waterways. *Sci. Technol. Outlook* **000**(021), 161 (2015)
5. Gao, Q.: Analysis on the issues of channel planning in port waters. *Sci. Technol. Inf.* **000**(035), 27 (2009)
6. Sun, X., Liu, X.: On the planning and design of port waters and waterways. *China Water Transp. (Second Half Mon.)* **17**(009), 65–66 (2017)
7. Liu, H.: On the planning, design and safety of port waters and navigation channels. *Cit. Ed. Eng. Technol.* **000**(010), 85 (2015)
8. Qina, W.C.: Research on the planning and design of port waters and navigation channels and safety evaluation. *Driv. Gard.* **000**(004), 80 (2019)
9. Yu, J.: Research on Reasonable Planning and Design of Port Waters and Channels. Dalian Maritime University (2009)
10. Ye, Y.: Analysis of countermeasures to improve the passage capacity of inland river ports. *Sci. Technol. Innov. Appl.* **22**(22), 173 (2014)





# Ant Colony Algorithm Theory and Its Application in Control Engineering

Chunxiang Huang, Nenjun Ben<sup>(✉)</sup>, and Guojun Yan

Institute of Intelligent Manufacturing,  
Yancheng Polytechnic College, Yancheng 224005, Jiangsu, China  
ycgyhcx@126.com

**Abstract.** Ant colony algorithm theory is an algorithm that helps to find a path or route for a given task. It is also used in control engineering to find the best way to control the traffic flow on the road. This method can be applied to various problems, such as routing, scheduling and so on. The main advantage of this algorithm is that compared with other algorithms (such as Dijkstra algorithm and a\* search algorithm), it requires very little computing time. This process involves several steps. These steps are explained. This paper will introduce some basic concepts of ant colony algorithm (ACA) theory and its application in control engineering. We will also outline some existing work on ant colony algorithm (ACA) for various control system applications, including real-time scheduling and traffic flow.

**Keywords:** Ant colony Control engineering · dispatch

## 1 Introduction

Ant colony algorithm is essentially a kind of nonlinear control, which belongs to the category of intelligent control. Ant colony algorithm is characterized by both systematic theory and a large number of practical applications. The development of ant colony algorithm initially encountered great resistance in the West; However, in the East, especially in Japan, it has been rapidly and widely promoted and applied. Over the past 30 years, ant colony algorithm has made great progress in theory and technology, and has become a very active and fruitful branch in the field of automatic control. Examples of its typical applications involve many aspects of production and life, such as fuzzy washing machines, fuzzy air conditioners, fuzzy microwave ovens, fuzzy vacuum cleaners, fuzzy cameras and video recorders in household electrical equipment; Examples applied in the field of industrial control include fuzzy water purification treatment, fuzzy fermentation process, fuzzy chemical reaction kettle, fuzzy cement kiln control, etc.; Applications in special systems and other aspects include fuzzy subway stop, parking control, fuzzy car driving, fuzzy elevator, escalator control [1], fuzzy steam engine and robot control.

The wide application of fuzzy theory in the field of control is entirely determined by the characteristics of ant colony algorithm itself. Ant colony algorithm uses human language information to simulate human thinking, so it is easy to understand, simple to

design and convenient to maintain. Based on the control rules containing fuzzy information, the control system composed of ant colony algorithm is more stable and robust than the conventional control system. When improving the characteristics of the system, the ant colony algorithm system does not need to adjust parameters only like the conventional control system, but also can modify the characteristics of the system by changing the control rules, membership functions, reasoning methods and decision-making methods. Therefore, the design, adjustment and maintenance of ant colony algorithm become simple. In the conventional control algorithm, small errors and parameter drift may cause the system out of control, while the ant colony algorithm system based on control rules is very sensitive to the change of a certain rule, and the system has strong anti-interference ability [2].

## 2 Related Work

### 2.1 Ant Colony Algorithm

At the point of calculation, it is necessary to initialize parameters, such as ant number  $m$ , pheromone factor  $\alpha$ , heuristic function factor  $\beta$ , pheromone volatilization factor  $\rho$ , pheromone constant  $Q$ , maximum iteration number  $T$ , etc.

The parameter selection of ant colony algorithm requires experience or trial and error, so in terms of parameter setting, attention should be paid to:

If the number of ants  $m$  is set too large, it will make the pheromone on each path tend to average, weaken the positive feedback effect, and slow down the convergence speed; If the number of ants is set too small, the pheromone concentration of some paths that have never been searched may be reduced to 0, so that the algorithm converges prematurely and the global optimality of the solution will be reduced. Generally, the number of ants is set to 1.5 times the target number [3].

If the pheromone constant  $Q$  is set too much, the search range of ant colony will be reduced, resulting in premature convergence of the algorithm, and the population will fall into local optimization; If the setting is too small, the difference of information content in each path will be small, and it is easy to fall into chaos. The pheromone constant is generally taken as [10100] according to experience [4].

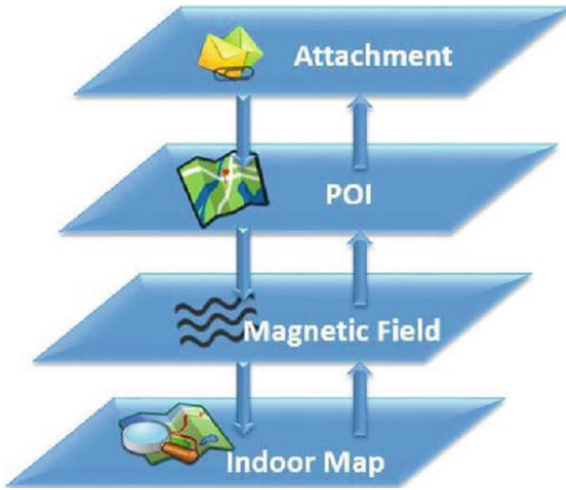
If the maximum number of iterations is set too high, the algorithm will run too long; Setting too small will lead to fewer alternative paths and make the population fall into local optimization. The maximum number of iterations is generally [100500], and the recommended value is 200.

$$P_{ij}^k(t) = \begin{cases} \frac{\tau_{ij}^\alpha(t) \eta_{ij}^\beta}{\sum_{j \in N_j^k} \tau_{ij}^\alpha(t) \eta_{ij}^\beta} & \\ 0 & \end{cases} \quad (1)$$

$$E^N = \frac{1}{2} \sum_{n=1}^N \sum_{k=1}^c (t_k^n - y_k^n)^2 \quad (2)$$

Pheromone factor indicates the relative importance of pheromone stars accumulated on the path in guiding ant colony search. If the parameter setting is too large, the ants

are more likely to choose the path they have walked before, which is easy to weaken the randomness of the algorithm; If the parameter is set too small, the search range of ant colony will be too small, which will make the algorithm converge prematurely and make the population fall into local optimization. Generally, the value is between [1, 4]. The flow of ant colony algorithm is shown in Fig. 1 below.



**Fig. 1.** Ant colony algorithm process

The heuristic function factor indicates the relative importance of heuristic information in guiding ant colony search. If the parameter is set too large, the convergence speed will be accelerated, but it is easy to fall into local optimization; If the parameter is set too small, the randomness of ant colony search will become larger, and it is difficult to find the optimal solution. According to experience, the value range of this parameter is generally between [0.5].

### 2.2 Control Technology

In recent years, control technology has developed rapidly and attracted extensive attention. Because the control has strong robustness and flexibility, it shows obvious advantages in the control of some nonlinear systems and systems that are difficult to establish accurate mathematical models. Control is based on human experience and knowledge, so prior knowledge is particularly important when selecting appropriate quantitative factors, scale factors, membership functions and control rules. However, this kind of prior knowledge is often not comprehensive enough, especially for some complex and nonlinear systems, it is impossible to obtain detailed or accurate prior knowledge, which brings certain difficulties for the effective implementation of control and the improvement of accuracy [5].

The optimization of control rules is a bottleneck problem in the practical application of controller. If this problem is not solved, the application and popularization of control

will be greatly limited. In order to solve this problem, people have been studying the methods and techniques of automatic generation and optimization of design parameters and control rules [6]. For various methods of controller parameter optimization, genetic algorithm has the advantages of wide adaptability and convenient parallel optimization of multi-objective and multi parameter problems. In addition to its good optimization ability, it is a main research direction of many scholars at present. Using genetic algorithm to optimize the control parameters of the controller, such as quantitative factors, saves the process of collecting expert control experience, overcomes the randomness of manual design of the controller, and expands the method of controller design and the application field of control.

### 3 Ant Colony Algorithm Theory and its Application in Control Engineering

In the process of fuzzy controller optimization, the optimization of fuzzy quantization factors, fuzzy membership function, fuzzy rules or the simultaneous optimization of multiple objectives can be used as needed. Different genetic coding methods can be adopted according to different optimization objectives.

For example, only optimizing the shape of fuzzy quantization factor or membership function is actually the optimization of several values, so the genetic coding method is relatively flexible, and binary coding, decimal coding and other methods can be used. When optimizing fuzzy control rules, because the number of general fuzzy rules is relatively large, the rule range is relatively small, and only the number of quantization levels of the output Analects, decimal coding is more convenient [7], Of course, this is not absolute, and specific problems need to be analyzed.

The simulated annealing operation first needs to design the initial temperature. At a higher initial temperature, the poor individuals can be introduced, and as the temperature decreases, it will be difficult to accept the poor individuals, and the probability of selecting the optimized individuals increases. There are many methods to deal with simulated annealing, such as linear cooling or exponential cooling, such as  $tx = t*k$  (where  $0 < k < 1$  is a constant), etc. The coupling system model of control engineering is shown in Fig. 2 below.

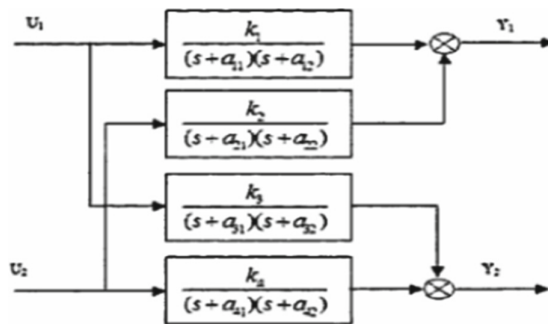


Fig. 2. Coupling system model of control engineering

Fuzzy control is a new control method that uses expert experience to construct language information and simulates the unique fuzzy and uncertain information processing methods in human thinking, behavior and language. Fuzzy control is a rule-based control method [8]. The formulation of control rules adopts the control experience of field operators or the knowledge of experts. In the design, there is no need to establish an accurate mathematical model of the controlled object, and the design is simple. Compared with traditional control, the system has fast response, short adjustment time, strong adaptability and robustness, and has a certain level of intelligence. Therefore, it is suitable for solving the control problems of nonlinear, time-varying and time-delay systems that are difficult to be solved by conventional control. PD type two input and single output fuzzy controller is the most widely used in practical engineering applications [9]. However, PD Fuzzy controller has some shortcomings, such as low steady-state accuracy and easy to produce oscillation, which limits the wide application of fuzzy control.

## 4 PID Control

PID control algorithm is simple and reliable, but it has the shortcomings of low control accuracy and difficult to set parameters, and the control effect is not good for the object with uncertainty. Fuzzy control does not depend on the model of the control object, and has great adaptability and robustness. It is very suitable for control occasions such as model uncertainty, parameter change and so on. Grey predictive control has the advantages of simple modeling, fast calculation speed, convenient real-time control, and can predict the future trend of the object, and can make decision and control in advance. Combining grey predictive control with fuzzy control, grey predictive control is used to predict the error and error change rate of the system, and this predictive change value is used as the input of the fuzzy controller. Through fuzzy reasoning and decision-making, the advanced control decision is obtained, which is conducive to reducing the overshoot and error of the control. On this basis, combined with the traditional PID controller, the PID controller is used as a conventional feedback control to stabilize the system and achieve the preliminary control effect [10]. Then through the combination of grey predictive control and fuzzy control, the robustness and adaptive ability of the system are used to improve the anti disturbance ability and adaptive ability of the system, so as to compensate the uncertainty and external interference of the system and improve the accuracy of system control. In this way, we can achieve a higher control effect by learning from each other.

## 5 Conclusion

Ant colony algorithm is a self-organizing system that has been used in the field of control engineering for many years to solve complex problems. The basic idea behind the algorithm is to divide the problem into smaller sub problems, and then solve each sub problem independently. In this way, solutions that are not obvious or difficult to implement can be found. Ant colony algorithm (ACO) is an example of meta heuristic search technology, which has been widely used in optimization problems with large search space, such as combinatorial optimization, scheduling, routing and other similar

problems. ACO can be seen as an extension of ant's ability to cooperate to achieve better solutions than any independent solution.

**Acknowledgement.** 2021 Jiangsu Province Policy Guidance Plan (Special for Introducing Foreign Talents) Fund Project: Key technology research and development of wall-climbing robots for large ferromagnetic surfaces such as ships (BX2021051).

## References

1. Kiseleva, M., Kuznetsov, N., Leonov, G.: Theory of differential inclusions and its application in mechanics. In: Clempner, J.B., Yu, W. (eds.) *New Perspectives and Applications of Modern Control Theory*, pp. 219–239. Springer, Cham (2018). [https://doi.org/10.1007/978-3-319-62464-8\\_9](https://doi.org/10.1007/978-3-319-62464-8_9)
2. Zhang, H., Yue, Q., Yuan, Q.: Illusion of control theory and its application and prospect in information system. *J. Mod. Inf.* (2018)
3. Zhang, J., Wang, X., Ju, M., et al.: An improved sparsity adaptive matching pursuit algorithm and its application in shock wave testing. *Math. Probl. Eng.* **2021**(8), 1–10 (2021)
4. Iranpour, E., Sharifian, S.: A distributed load balancing and admission control algorithm based on fuzzy type-2 and game theory for large-scale SaaS cloud architectures. *Futur. Gener. Comput. Syst.* **86**, 81–98 (2018)
5. Jiang, F., Wang, L., Bai, L.: An improved whale algorithm and its application in truss optimization **18**(3), 12 (2021)
6. Qiao, B., Li, H., Wu, X.: Intelligent-assist algorithm for remote shared-control driving based on game theory. *Journal of Shanghai Jiaotong University (Science)* **26**(5), 615–625 (2021)
7. Zhu, M., Yan, X., Yuan, Q.: Structural holes theory and its application and prospect in social network studies. *J. Intell.* **37**, 190–197 (2018)
8. Guan, G., Liao, W., Mao, Z., et al.: Comparison and improvement of time delay parameter algorithm for feedforward volume compensation control in canal system. *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering* **34**(24), 72–80 (2018)
9. Andrei, P., Valentin, P.: Extremum-open-source optimization library and its application to optimal control synthesis problem. In: *2018 IV International Conference on Information Technologies in Engineering Education (Inforino)* (2018)
10. Shen, G.F., Jing-Ming, L.I., Zhao, S.P.: Research on GIHS-TSFNN parallel learning algorithm and its application. *Mathematics in Practice and Theory* (2019)



# Application and Optimization of Style Transfer Algorithm in Contemporary Cultural and Creative Products

Jinxia Chen<sup>(✉)</sup> and Gui Wang

Guilin Institute of Information Technology, Guilin 541004, Guangxi, China  
cjx873@163.com

**Abstract.** With the development of economic globalization and information age, cultural and creative industries have gradually become the pillar industries of a country or region. In recent years, China has developed rapidly in the field of cultural and creative industries, but there are also many problems. Therefore, it is of great significance to study the development status and existing problems of China's cultural and creative industries and put forward relevant suggestions to improve the development status of China's cultural and creative industries and enhance China's international competitiveness. The application of style transformation is a new phenomenon developed by artificial intelligence and computer vision. The main purpose of this technology is to create an image that looks like it was taken from a real person or object without using a real image. In recent years, this technology has many applications in various fields such as art, photography and architecture. However, the most important application is clothing design. This can be achieved by creating unique clothes of different styles and colors, or by creating unique hairstyles at home or even on social media platforms such as instagram or Facebook.

**Keywords:** Style migration algorithm · Contemporary culture · Creative products

## 1 Introduction

As a newly rising industry, the rapid development of cultural and creative industries benefits from the comprehensive assistance of economic globalization. Its core is independent creativity. Its main content is that the main culture of a country or a nation develops intellectual property rights through corresponding creative formation, technological development and industrialization and makes profits. Creative industry was first proposed by Britain, and then many countries and regions have put forward relevant concepts that are consistent with their own cultural background [1]. Cultural and creative industries mainly include creative groups in the fields of radio, film and television, animation, audio-visual, media, visual arts, performing arts, technology and design, advertising and decoration, clothing design, software and computer services. The understanding of cultural and creative industries in major countries and regions in the world

can be broadly divided into three types: the “creative type” represented by the United Kingdom; The “copyright type” represented by the United States and the “cultural type” represented by China and South Korea [2].

Since the development of the world economy, the main driving force of social development and economic growth has changed from the early industrial promotion to the Internet based cultural industry and creative economy at the present stage. In order to adapt to the pace of rapid economic development in the current era, various countries and regions in the world have put forward cultural and creative industry development policies adapted to their own national conditions and characteristics to varying degrees. It is hoped that through the rapid development of the industry, social employment and the vitality of the national economy can be promoted again [3]. On the other hand, the cultural and creative industries not only show unparalleled advantages in economic development, but also show strong attraction in disseminating the excellent culture of the country and enhancing the country’s identity and overall competitiveness in the international community. In addition, the cultural and creative industries have highlighted the irreplaceable importance of other industries in terms of added value income generation, environmental protection and the incremental effect of marginal income. For this reason, cultural and creative products today are undoubtedly better than traditional industries and their products in terms of characteristics of the times and competitive advantages [4]. Today, with the integration of the world economy, the exchange and cooperation between various countries and regions and industries are also deepening, and this phenomenon is particularly obvious in the development of cultural and creative industries. The trade of cultural and creative products in various countries and regions has gradually become a new engine of world economic development.

## 2 Related Work

### 2.1 Research Status of Cultural and Creative Products

In terms of the research on the international competitiveness of cultural and creative products trade, Hu Fei and Ge Qiuying studied the current situation of China’s cultural and creative products export trade. At the same time, they quantitatively analyzed the competitive advantage of China’s cultural and creative products export trade in the mainland through multiple indicators and angles, and then put forward corresponding reform measures and improvement measures for the problems and shortcomings found; Fang Zhong and Zhang Huarong use the methods of comparative analysis and empirical analysis to study the trade of cultural and creative products between China and South Korea, and draw the conclusion that the international competitiveness of China’s cultural and creative products trade lags behind South Korea, but the trade development between the two countries will be more positive; Chen Weixiong made an empirical analysis on the trade of cultural and creative products in Chinese Mainland by using a variety of research indicators [5]. He came to the conclusion that the trade of cultural and creative products in Chinese Mainland is developing actively and has been more widely recognized internationally, but there is also an imbalance in trade development layout. He also put forward development strategies for this conclusion.



In terms of the current situation and policies of cultural and creative products trade, Liu Xiaohui studied the laws of the development of China's cultural and creative products from the perspective of international trade, and further pointed out that China should actively learn from developed countries and countries or regions with strong competitiveness in the development of cultural and creative products trade in the world, so as to promote the positive development of domestic trade and become more competitive in the international market; By comparing and analyzing the current situation of the development of cultural and creative industries in China and Britain, Zhao Jinjin believes that although China's export trade volume in this industry is large, the income is not optimistic, and even the abnormal phenomenon of trade deficit appears, and then puts forward suggestions for improvement; Yang Lixin starts with an empirical analysis of the cultural and creative industries in the United States, and compares the current situation of China's industries [6]. Finally, she concludes that the rapid development of a country's economy in the current era can not be separated from the contribution of the cultural and creative industries.

## 2.2 Related Concepts of Cultural and Creative Industries

Britain was the first country in the world to pay attention to cultural and creative industries. In the 1990s, it first introduced the concept of "creativity" into the cultural policy documents issued by the government. Subsequently, the term "creative industry" and its specific concepts were listed in detail for the first time in a report entitled "Atlas of British creative industries" issued by the British government, The report believes that the products and smart industries created by individuals through the sublimation of their own ideas and talents and the accurate application of rapidly developing technologies can be called "creative industries" [7]. According to the above explanation of the connotation of this industry, the British government specifically classified music, performing arts, television broadcasting, art and cultural relics trading, handicrafts, architecture, design, advertising, fashion design, film, software, interactive leisure software, publishing and other industries as creative industries.

In the current era, the understanding of the scope of creative industries and the formulation of relevant definitions and norms are different among countries or regions in the world. There is no connection between them and sometimes there is a big gap. In view of this, the United Nations Conference on Trade and development (UNCTAD) gave the definition and coverage of creative industries accepted and recognized by most countries or regions in 2007 on the basis of various definitions of creative industries formulated by various countries and regions in the past. According to the definition of UNCTAD, creative industries should have the following characteristics: (1) the creation, production and distribution of goods and services with creativity and intellectual capital as the primary input; (2) It constitutes a set of knowledge-based activities, focusing on but not limited to art and culture, which can generate potential income from international trade and intellectual property rights; (3) Tangible products and intangible intellectual or artistic services with creative content, economic value and market objectives; (4) It is an intersection of art, service and industrial sectors; (5) It has become a relatively active emerging sector in international trade. In terms of the coverage and extension of creative industries, the United Nations Conference on Trade and development (UNCTAD) divides

creative industries into four groups: cultural heritage, art, media and functional products, and gives specific classified products under each group.

### 3 Research on Style Migration Algorithm

The content of this section starts from the style migration algorithm, understands the essence of style migration technology through literature review, extracts its value in photographing beautification applications according to the basic architecture of style migration technology, and summarizes the image processing methods, user viscosity, user privacy and security, and finally summarizes a set of industrial problem solving methods and applicable scenarios based on the technical value of style migration algorithm.

#### 3.1 Concept and Essence of Style Transfer Technology

Style migration algorithm is an optimization technology. It can mix a content image and a style reference image (such as a work of a famous painter) together, so that the output image has the general content of the content image, but has the texture style of the style image. This is to optimize the output image by matching the content statistics of the content image and the style statistics of the style reference image, and these statistics can be obtained from each image through the convolutional neural network [8]. As shown in combination Fig. 1, the main performance of style transfer technology is shown: figure (a) is the content image, we selected an ordinary real-life architectural photo, figure (b) is the style reference image, and we selected the art painting “shout” created by Edward monk. Figure (c) is the output image, which is jointly generated by figure (a) and figure (b). The output image retains the basic characteristics of the original pattern of the content image, With the painting style of style reference image, it looks like a new painting work created on the architectural basis of style reference image and content image. Now we can intuitively feel the performance of image style transfer technology [9].



Fig. 1. Application of style transfer algorithm

### **3.2 Applicable Field Record of Style Transfer Technology**

Style transfer technology is actually a branch of computer vision. The research on image analysis is of great significance to image understanding and image representation. In some of our commonly used beautification applications, we can use the style change of some filter backgrounds, as well as the popular face comics and face paintings in recent years. As long as one photo can have a two-dimensional face or the characters in the oil painting, in addition, there is the transfer of face makeup, which transfers the makeup of other people's faces to their own photos. Style transfer technology is often used as an auxiliary means to help the computer improve the performance of visual tasks, such as pedestrian recognition, and help the computer find the same face again more effectively. Style transfer technology helps us understand images better and analyze image features from a special perspective.

## **4 Application and Optimization of Style Transfer Algorithm in Contemporary Cultural and Creative Products**

The style migration algorithm enables the machine to simulate the visual effects of people through intelligent technology, and turns the photos into art paintings. The application of style migration class converts the technology from web page to mobile phone. As the name implies, the photographing beautification application is an application category that helps users process images. Now we want to use the technology of style migration as a function of image processing to generate mobile phone applications. This application helps users to process images in the way of image style conversion, and users can get pictures that make any of their real photos have painting characteristics. At the same time, it provides materials for art paintings to further improve the visual experience of users [10].

First of all, from the analysis of market necessity: there are many endless experience modes since the popularity of style transfer technology. After people understand the technical effect, they need a platform to realize this effect. In short, mobile phone applications are the best carrier. Some users who need or want to try this technology need a fast and convenient form. Secondly, the application market of photographing and beautification is very wide, but in recent years, the application function lacks certain uniqueness, and a special function is needed to bring a new experience to the user. Secondly, from the analysis of basic conditions: the style transfer technology has been relatively mature, and it is the most important thing to enrich this application on the basis of this technology. Now we need to diversify the application in the form of expression, and improve the user's use experience with more smooth interaction.

## **5 Conclusion**

Cultural accumulation is the behavioral basis of a country or a nation. It represents the cognition of the whole group to itself and the response to the environment. Since entering the new century, all industries and fields in the world have shown great concern about the importance of cultural trade. Among them, cultural and creative industries, as one of the

industries with the most economic benefits and development prospects in this field, have attracted the attention of the academic community and become its main research field. Because of the extensive and continuous attention of the academic community and the increasingly important contribution of the industry to economic development, various countries and regions began to vigorously develop cultural and creative industries. As the specific materialized representative of the industry, the international trade of cultural and creative products has become an important manifestation of the international control and leadership of a country or region.

**Acknowledgements.** The Basic Ability Enhancement Program for Young and Middle-aged Teachers of Guangxi in 2022: Research on the design of Red Cultural Products in North Guangxi under the integration of culture and tourism (ProjectNo.: 2022KY1618).

## References

1. Ding, W., Yang, X.: Comparative study on the growth power of cultural and creative products in museums (2022). [https://doi.org/10.1007/978-3-031-06391-6\\_26](https://doi.org/10.1007/978-3-031-06391-6_26)
2. Li, X., Lin, B.: The development and design of artificial intelligence in cultural and creative products. *Math. Prob. Eng.* **2021**, 1–10 (2021)
3. Wei, X., Zhang, Y.: Design strategy and issues of realising tourism cultural and creative products in northeast China (2021)
4. Zhang, J., Kong, D.: Application and optimization of algorithm recommendation in mobile audio APP. *J. Phys. Conf. Ser.* **1848**(1), 012025 (2021)
5. Meng, X.: Optimization of cultural and creative product design based on simulated annealing algorithm. *Complexity* **2021**, 1–10 (2021)
6. Chen, X., Xiao, S.: Multi-objective and parallel particle swarm optimization algorithm for container-based microservice scheduling. *Sensors (Basel, Switzerland)* **21**(18), 6212 (2021)
7. Mn, A., Hj, B., Na, B.: Half a century experience in rate of penetration management: Application of machine learning methods and optimization algorithms - A review (2021)
8. Zhang, J.: Exploring the application of traditional elements in cultural and creative product design. *Art Des. Rev.* **9**(4), 9 (2021)
9. Liu, H.: Application and optimization of project management in the process of construction engineering management. In: *IPEC 2021: 2021 2nd Asia-Pacific Conference on Image Processing, Electronics and Computers* (2021)
10. Chen, J.: Optimization and application of particle swarm algorithm in software engineering. Springer Cham (2022). [https://doi.org/10.1007/978-3-031-05484-6\\_77](https://doi.org/10.1007/978-3-031-05484-6_77)



# Application of Ant Colony Algorithm in Chinese Language and Literature Education

Qianyu Ma<sup>(✉)</sup>

Northwest Minzu University, Lanzhou 730030, Gansu, China  
zcysym@126.com

**Abstract.** In this study, we apply ant colony algorithm to Chinese language and literature education. We believe that this is a good way to improve the quality of Chinese language and literature education. This method can be used by teachers and students in daily life. It also has great application potential and can be applied to other fields, such as computer science, biology, chemistry and so on. This method can be used to teach pupils and adults with different proficiency levels. The main idea behind this method is that it can help learners learn the meaning of words more easily through memory, not just how to write words. In this way, learners will better understand what they are writing and why they are writing it, and will be able to remember it longer than just being taught how to twist.

**Keywords:** Ant colony Language and literature · education

## 1 Introduction

Language is not a transparent existence, not just a tool of communication and a shell of thought. Without language communication, there would be no human society; Without human society, there will be no thinking consciousness. Language is the foundation of all knowledge discourse, which affects people's thinking, emotion and intuition. Our thinking is unconsciously controlled by language, so in a sense, language is the thought itself, the thinking process and the thinking content. Language is intrinsically related to cultural concepts and national spirit. Different languages have formed different cultural forms. Classical Chinese style and vernacular style have created Chinese classical literature and modern literature with different styles respectively. Language is not a stable rational system. It will change with the development of human society [1]. Language shows different values in different historical periods. Language is not unpredictable and difficult to understand. It has not only the systematicness and structure of human civilization, but also the fluidity and dialogue of culture. Like culture, language in the transitional period is often noisy.

In other words, linguistic knowledge is, in my opinion, the most fundamental part of language education. This link does not mean that I can ignore a language if I learn it well and give my experience to you, because education itself is also a philosophical problem and a process of psychological game. Through the understanding of the nature of human language and the structure and formation of various languages in the world,

the problem of what to teach is first solved; Secondly, through the psychological analysis of human beings at different ages, the problem of how to teach is solved [2]; Finally, the understanding of philosophical thought also solves the problem of why to teach. The combination of the three can be called real language education. Based on this, this paper studies the application of ant colony algorithm in Chinese language and literature education.

## 2 Related work

### 2.1 Ant Colony Algorithm

Modern biology has found that ants in nature mainly rely on pheromone, a chemical produced by ants, to transmit information. This is different from humans and other advanced species, which mainly transmit information through vision and hearing. In fact, many ants' visual and auditory systems have completely degenerated. In fact, they are "blind" and "deaf". Just as homing pigeons use magnetic fields to locate, ants also use this pheromone to locate the path between the ant nest and food, so as not to get lost [3]. And if the pheromone content on a certain path is relatively high, then more ants will gather on this path, which is a typical positive feedback.

$$D_r = \frac{\sum_{x_j \in Z_r} \sum_{x_i \in Z_r} D(x_i, x_j)}{n_r^2} \tag{1}$$

This way of locating the shortest path between ant nest and food source through pheromone is the inspiration of ant colony algorithm. The path analysis of ant colony algorithm is shown in Fig. 1 below.

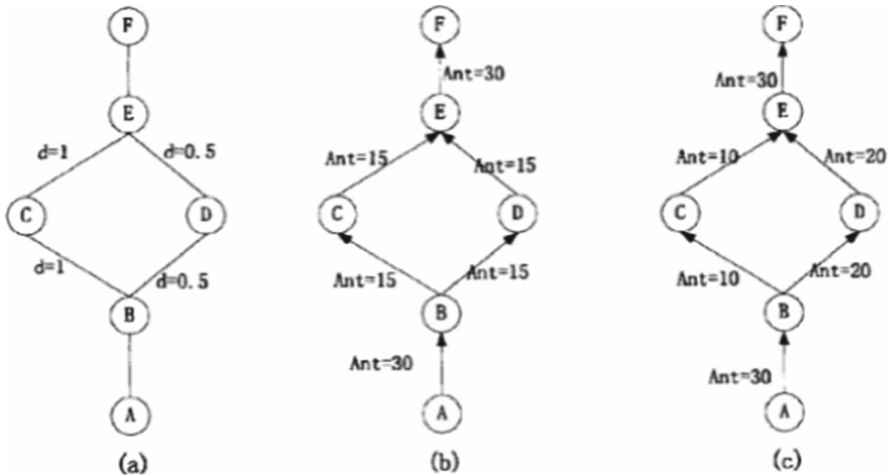


Fig. 1. Path analysis of ant colony algorithm

Suppose point a is the ant nest, point F is the food source, and CD is the obstacle between the two points. Ants at points B and e should choose the way forward, and the

distance between points EF and ab can be ignored. Assume that the path  $bce = 2$ ,  $bde = 1$ . Every certain time, 30 ants start from point A [4]. Assuming that their speed is 1 and they move forward at a uniform speed, they will reach point B after a period of time. As shown in Fig. 1 (a), ants will leave a certain amount of pheromones on the passage they pass, and these pheromones will not be affected by the outside world, but will volatilize evenly at a fixed rate.

At time  $t = -0$ , there is no pheromone interfering with ants on both BCE and BDE paths. Therefore, ants on point B are not affected by pheromones and randomly choose paths. We assume that the ratio of choosing BCE and BDE paths is the same, that is, there are 15 ants on each path. As shown in Fig. 1 (b).

At  $-1$ , all 15 ants on BDE have reached point E and begin to return; The 15 ants on the BCE path have just reached point C:

At time  $t = 2$ , 15 ants on BDE have returned to point B, while ants on BCE have just begun to return:

At this time, if there are 30 new ants from point a to point F, because the BDE path has been passed twice, and the pheromone content is twice that of BCE, then later ants will choose the BDE route more, and with the passage of time, until all ants choose the BDE route [5].

In this experiment, more ants pass through the BCE path than the BDE path, so the pheromones left by the ants on the BCE path are far more than those on the BDE path. According to the positive feedback characteristics of the ant colony, later ants tend to choose a shorter BCE path. After a certain number of iterations, all ants will choose a shorter BCE path at last. Compared with the first experiment, in this experiment, the initial random fluctuation phenomenon is less, mainly because the pheromone released by ants plays a leading role.

## 2.2 Language and Literature Education

Understanding of popular culture. "Teaching Chinese as a foreign language is for foreign students according to its disciplinary nature, and teaching Chinese as a second language." Culture is a unique symbol of a country and national spirit. Since the great discussion of culture began in the last century, the cultural issues in teaching Chinese as a foreign language have always been an important topic in the field of Chinese as a foreign language. The "culture" we want to discuss here refers to the cultural teaching in teaching Chinese as a foreign language, because teaching Chinese as a foreign language is still a kind of language teaching in the final analysis. We should introduce cultural factors into teaching Chinese as a foreign language, eliminate the cultural obstacles in the communication between foreign students and Chinese people in Chinese, and overcome the communication difficulties caused by the differences in values, thinking patterns, cultural backgrounds, and moral and ethical concepts. In the process of the development of the times, popular culture embodies a strong vitality. Popular culture is represented by popular music, film and television literature and art, popular novels and network culture [6]. Popular culture has penetrated into people's lives and imperceptibly promoted language learners to better master the language.

Popular culture has an important influence on Teaching Chinese as a foreign language. The post-80s and post-90s generation both grew up in the global pop culture

environment, and increasingly rely on intuitive, vivid and diverse media means. Pop culture has a great impact on them. Especially for foreign students, Chinese itself is difficult to learn, and does not have the international status of English. There must be continuous stimulation from the outside world to ensure their enthusiasm for learning. Popular culture plays a positive role in promoting the teaching of Chinese as a foreign language. At present, the United States, Germany and Japan are all popular with Chinese fever. For European and American countries, their learning Chinese is mainly to understand Chinese culture and the past, present and future of Chinese culture. For Japan and South Korea, learning Chinese is because their own culture and language come down in one continuous line with Chinese culture. The introduction of popular culture is one of the simple ways to stimulate students' external learning motivation. Generally speaking, vivid, interesting and even fashionable knowledge is more likely to arouse their desire to learn [7]. Chinese popular culture effectively combines the interest of Chinese learning with the practical use function, so that people can experience the interest of classroom activities in Chinese as a foreign language, so as to promote students to use their familiar knowledge in popular culture to participate in classroom activities and produce meaningful language communication effects.

### 3 Application of Ant Colony Algorithm in Chinese Language and Literature Education

From the probability formula of the basic ant colony algorithm, it can be seen that when  $\alpha = 0$ , the ant colony algorithm will be transformed into the traditional greedy algorithm; And when  $\beta = 0$ , the algorithm evolves into a purely positive feedback heuristic search algorithm [8]. Our goal is to let the ants coordinate their work under certain heuristic strategies to complete the overall complex Work. From the various applications of ant colony algorithm, the selection of heuristic factors is also quite exquisite.

First of all, we can think of the evaluation function in algorithm A:

$$SSE = \sum_{l=1}^K \sum_{x \in L_l} Dist(x, Z_l)^2 \quad (2)$$

The read language and literature Chinese character string is transformed into a sequence of words headed by each Chinese character in the string in turn in memory, and the sequence of words headed by this Chinese character is combined. In a word, preprocessing the string and sorting out the processed data requires a general Chinese dictionary.

Our electronic thesaurus supports the first word hash. In order to save time, the same first word in the thesaurus is arranged by the pointer according to the length from long to short [9]; And write down the maximum length of such words for later comparison with the length of sentences.

Similarly, the release of pheromones will also face the following problems. There are three options: (1) Pheromones are released on the Chinese language and literature characters of the words. Because the length of the words cut each time is different, for a specific ant, it is assumed that the total amount of pheromones is  $Q$ . If the pheromone



is evenly released to Chinese language and literature Chinese characters after cyclic segmentation, due to a certain number of words, it will cause the pheromone to have no discrimination, form a positive feedback mechanism, and recognize words; Or, after cutting a word, the pheromone Q will be released to all Chinese language and literature Chinese characters of the word on average. Although it can form strong differences in pheromones on Chinese language and literature Chinese characters, it may form pairs of words in fact, and the pheromones on Chinese language and literature Chinese characters are different, which makes it difficult to identify words. Therefore, this scheme is not feasible. (2). Imagine that there is a path between Chinese characters in Chinese language and literature, and pheromones are released on the connection between words. This scheme is easy to recognize that words cannot form words in sentences, and for word recognition, its essence is similar to (1), so this scheme is not feasible. (3). The pheromone is released on the cut word. Because each ant cuts different words, even on the same word, the pheromones released by ants are different. In this way, the sum of pheromones on each word is different [10]. The larger the pheromone, the greater the probability of being selected in the next iteration, which is the embodiment of the positive feedback mechanism.

## 4 Conclusion

The application of ant colony algorithm in Chinese language and literature education is a method to find the most appropriate words, sentences or paragraphs for a given text. It is based on the idea that all texts are composed of a few “key” words, which determine its meaning. Ant colony algorithm searches these keywords in each sentence or paragraph and finds their frequency distribution. The most common one is found first, so it can be used as an effective tool to improve students’ understanding of the text.

## References

1. Research on literature aesthetic education based on neural network algorithm. *J. Phys. Conf. Ser.* **1881**(4), 042066 (2021)
2. Wang, Z.: The evaluation model of teaching quality of modern and contemporary literature in Chinese higher education based on support vector regression algorithm (SVRA), pp. 452–455 (2018)
3. Jiang, T., Yang, X., Yang, Y., et al.: Wavelet method optimised by ant colony algorithm used for extracting stable and unstable signals in intelligent substations **7**(2), 9 (2022)
4. Wang, X.Q.: Research on the implementation of aesthetic education in Chinese language and literature teaching. *Journal of Heilongjiang College of Education* (2019)
5. Liu, L.L.: Research on the reform of education in Chinese language and literature undergraduate course. *Education Teaching Forum* (2019)
6. Weiqiang, W.U., School, H.A.: On the problems in Chinese language and literature education and countermeasures. *The Guide of Science & Education* (2019)
7. Qian, W.: A consideration on strengthening Chinese classics education in Chinese language and literature subject. *Chinese Culture Research* (2018)
8. Liu, Y.: Analyses the Chinese language and literature education and college students’ humanistic quality education. *The Theory and Practice of Innovation and Entrepreneurship* (2018)

9. Zhang, C.L.: The problems existing in Chinese language and literature education and its countermeasures. Heilongjiang Science (2018)
10. Yang, G.X.: The development of aesthetic education in Chinese language and literature teaching. Heilongjiang Science (2018)



# Application of BP Neural Network Algorithm in Teaching Evaluation and Learning Prediction

Zhangrong Hu<sup>(✉)</sup>

China West Normal University, Nanchong 637002, Sichuan, China  
h zr2568@cwnu.edu.cn

**Abstract.** The application of BP neural network algorithm in teaching evaluation and learning prediction is to predict the effectiveness of teachers. In this case, the obtained data is used as the input of BP neural network algorithm. This output is used to evaluate teachers and predict their effectiveness based on previous research results. How does it work? The method is divided into two parts: training phase and testing phase. The training phase includes three steps: pretreatment, training and post-processing. The main purpose of this paper is to provide practical solutions for teaching evaluation and learning prediction. It has been proved that combining BP neural network algorithm with other algorithms can improve the quality of teaching evaluation and learning prediction. Our method is more effective than other methods, because it can not only predict students' performance, but also predict teachers' performance of students' future achievements according to their past achievements.

**Keywords:** BP neural network · Teaching evaluation · Learning prediction

## 1 Introduction

Types of teaching evaluation the content of this part is mainly single choice questions. The examination form is case type single choice questions or concept type single choice questions. Occasionally, multiple choice questions will be examined; This part of the content should focus on the concept distinction and understanding between different types of teaching evaluation.

Learning forecasting, a branch of economic forecasting, refers to the measurement of the future development prospect of the learning market based on accurate survey statistics and teaching information, starting from the historical status quo and regularity of the learning market and using scientific methods. In the prediction research of teaching system, learning prediction is a very hot topic, because the learning market has the characteristics of high income and high risk. With the development of teaching, people are constantly exploring its internal laws, gradually deepening their understanding of teaching laws, and producing a variety of teaching prediction methods [1]. However, as a complex giant system with many influencing factors and various uncertainties, the learning market often shows strong nonlinear characteristics. In addition, the amount of information processed by teaching modeling and prediction is often very large, and

there are high requirements for algorithms. It is because of these complex factors that the prediction of teaching is often difficult.

Traditional learning technology analysis methods include moving average method, point chart method, K-line chart method, etc. they can predict the general trend of stock index changes over a period of time, but the changes of short-term learning are often the information that investors are more interested in. In addition, the traditional learning technology analysis method also needs to know various parameters in advance, and how these parameters should be modified under what circumstances [2]. The requirements for the predicted object are too specific and strict.

## 2 Related Work

### 2.1 BP Neural Network Algorithm

BP (back propagation) algorithm is one of the most important algorithms in neural network deep learning. Understanding BP algorithm can help us better understand the essence of neural network deep learning model training, which belongs to internal skill cultivation.

BP (back propagation) neural network is a concept proposed by scientists led by Rumelhart and McClelland in 1986. It is a multilayer feedforward neural network trained according to the error back propagation algorithm. It is one of the most widely used neural network models. Minsky and Papert pointed out in their influential book "perceptron" that a simple perceptron can only solve linear problems, and a network that can solve nonlinear problems should have a perception layer, but there is no reasonable theoretical basis for the learning rules of hidden layer neurons.

From the perspective of perceptron learning rules, the adjustment of its weight depends on the difference between the expected output and the actual output:

$$E^N = \frac{1}{2} \sum_{n=1}^N \sum_{k=1}^c (t_k^n - y_k^n)^2 \quad (1)$$

$$Z_{i,j,n}(K, X) = \sum_{C=1}^C K_{cn} \bullet X_{i,j,n} \quad (2)$$

But for the nodes of each hidden layer, there is no known expected output, so the learning rule can not be used to adjust the weight of the hidden layer.

The basic idea of BP algorithm is that the learning process is composed of two processes: signal forward propagation and error back propagation.

In forward propagation, the characteristics of the sample are input from the input layer, and the signal is processed by each hidden layer, and finally transmitted from the output layer. For the error between the actual output and the expected output of the network, the error signal is transmitted back layer by layer from the last layer, so as to obtain the error learning signal of each layer, and then the weight of each layer of neurons is modified according to the error learning signal.

This kind of signal forward propagation and error back propagation, and then the process of adjusting the weight of each layer is repeated [3]. The process of weight

adjustment is the process of network learning and training. This process is carried out until the network output error is reduced below the preset threshold or exceeds the preset maximum training times. Figure 1 below shows the weight adjustment of BP neural network.

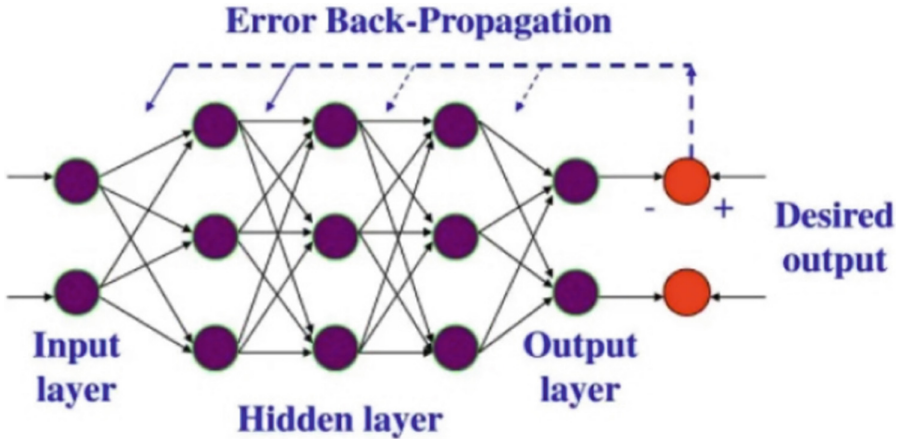


Fig. 1. Weight adjustment of BP neural network

## 2.2 Teaching Evaluation

Classroom performance evaluation is a process evaluation. Process evaluation requires that the evaluation subjects interact, the evaluation methods are dynamic and sustainable, the evaluation contents are multi-dimensional, and the “two skins” of evaluation and teaching are avoided. Schools need to consider how to organically combine evaluation and teaching, rather than simply fill in the evaluation results of a period of time with a few pages of written evaluation form at the end of the semester. Let teachers have motivation, students are interested, teaching is effective, the classroom is changed, and truly record a vivid process.

Now let’s introduce a simple and easy-to-use classroom performance evaluation tool that has been tested by practice [4].

Features:

1. simple

Import the list of students on the platform, set up evaluation indicators and prize exchange mechanism, and you can complete the evaluation and incentive of students as long as you “click” easily at ordinary times.

## 2. valid

Periodic evaluation indicators can be set according to the teaching requirements to make the objectives more clear; The last two minutes of the class can be used for public evaluation in the classroom to make the evaluation more timely and open; You can open the parents' terminal to push evaluation messages every day to make parents more active.

## 3. high tech

According to the usual evaluation data, it can automatically generate class and student personal reports and comments in the form of growth manual.

### **3 Application of BP Neural Network Algorithm in Teaching Evaluation and Learning Prediction**

The neural network model has the characteristics of massive parallelism, storage distribution, structural variability, high nonlinearity, self-study habit and self-organization, and can approach those functions that best describe the laws of sample data. Regardless of the form of these functions, the neural network has a wide range of adaptability, Learning ability and mapping ability. Through learning and mastering the dependency between data, it shows certain advantages in learning prediction. It has great advantages over the traditional methods that rely on derivation of mathematical models and parameter optimization, which are very accurate and bring limitations [5]. The teaching trend is highly nonlinear, and the transaction price and trading volume contain a large number of internal laws and characteristics that determine the teaching changes. Through the study of historical transaction data, the neural network can autonomously find out the laws and characteristics (S') between parameters from the complex data and depict these laws and characteristics. Therefore, by contrast, it has a good effect on the prediction of teaching trend.

According to the function of teaching evaluation, the basic types of teaching evaluation can be divided into diagnostic evaluation, formative evaluation and summative evaluation [6]; According to the evaluation criteria, it can be divided into absolute evaluation, relative evaluation and individual difference evaluation; According to the evaluation subject, it can be divided into internal evaluation and external evaluation.

Diagnostic evaluation is an evaluation conducted at the beginning of a semester or at the beginning of a unit teaching in order to understand students' learning readiness and factors affecting learning. It includes a variety of what is commonly known as a quiz.

The main functions of diagnostic evaluation are: to check the students' readiness for learning; Decide on the appropriate placement of students; Identify the causes of students' learning difficulties.

Formative evaluation is the evaluation of students' learning process and results in order to improve and perfect teaching activities in the teaching process. It includes oral questions and written tests for students in the teaching of a class or a topic.

Teaching itself and the variables that affect teaching show nonlinear characteristics, so it is required to have a strong ability to deal with nonlinear problems. Most

of the existing mature technologies are to solve linear problems and univariate nonlinear problems [7]. There is a lack of effective analysis tools for complex and general multivariable nonlinear problems. Deterministic linear cusps can only produce simple behavior. Deterministic nonlinear cusps can produce chaos. Teaching system is a multivariable nonlinear problem, Although nonlinear mathematics and dissipative structure theory provide some tools for describing nonlinear dynamic systems, there are still many problems in practical application in teaching empirical analysis [8].

## 4 Simulation Analysis

The complexity of teaching behavior and the ability to influence future events make the prediction error quite large, and it rises sharply with the increase of time. Different from other physical systems, in the teaching system, the best matching of sample data can not guarantee the best prediction, that is, the minimum error criterion of modeling data is not the best criterion to improve the prediction accuracy, and the past and present performance of a prediction method can not explain its future prediction results. Figure 2 below shows the learning prediction implementation code.

```
# generate raw data
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline

def f_fun(x):
    return 10 * np.sin(2 * np.pi * x)
num_samples = 32
X = np.linspace(-0.5, 0.5, num_samples).reshape(-1, 1)
y_true = f_fun(X)
y = y_true + np.random.randn(num_samples, 1)

plt.scatter(X, y, marker='+', color='y', label='Training data')
plt.plot(X, y_true, label='Truth')
plt.grid()
plt.legend()
```

**Fig. 2.** Learning prediction implementation code

In contrast, neural network has achieved good results in learning and prediction. It has been proved mathematically that artificial neural network can approach those functions that best describe the laws of sample data, regardless of their forms. The learning ability of neural network and the dependence system between data through learning show certain advantages in learning prediction [9]. It has great advantages over the traditional methods

that rely on the derivation of mathematical models and parameter optimization, which are very accurate and bring limitations, The teaching trend is highly nonlinear, and the teaching contains a large number of internal laws and characteristics that determine the teaching changes. Through the study of historical transaction data, the artificial neural network can independently find the laws and characteristics between parameters from the complex data, and depict these laws and characteristics [10]. Therefore, it has a good effect on the prediction of teaching trend.

## 5 Conclusion

The application of BP neural network algorithm in teaching evaluation and learning prediction is to predict the future performance of a given student. This will help teachers improve their teaching skills. How to apply BP neural network algorithm to teaching evaluation and learning prediction? We need two datasets, one for training and the other for testing. The training set shall contain all data points required by the training model, while the test set shall only contain those data points that cannot be seen during its training. In order to get a good result, we need to adjust our parameters, such as the number of hidden nodes, activation function, etc.

## References

1. Zhang, R., Liu, M., Zhang, Q., et al.: A network security situation prediction algorithm based on BP neural network optimized by SOA (2020). [https://doi.org/10.1007/978-3-030-57881-7\\_37](https://doi.org/10.1007/978-3-030-57881-7_37)
2. Xu, X., Xie, G.: Convergence analysis of multilayer BP neural network with momentum term. *J. Phys. Conf. Ser.* **1650**(3), 032123 (2020)
3. Yu, L., Xie, L., Liu, C., et al.: Optimization of BP neural network model by chaotic krill herd algorithm. *Alex. Eng. J.* **61**(12), 9769–9777 (2022)
4. Chenghao, Q., Lei, N., Tingyi, Z., Yifei, Z., Yuting, T.: Function analysis of deep learning in BP neural network structure design. In: Xu, Z., Alrabaee, S., Loyola-González, O., Zhang, X., Cahyani, N.D.W., Ab Rahman, N.H. (eds.) *CSIA 2022. LNDECT*, vol. 125, pp. 622–627. Springer, Cham (2022). [https://doi.org/10.1007/978-3-030-97874-7\\_81](https://doi.org/10.1007/978-3-030-97874-7_81)
5. Yao, L., Ren, L., Gong, G.: Evaluation of chloride diffusion in concrete using PSO-BP and BP neural network. *IOP Conf. Ser. Earth Environ. Sci.* **687**(1), 012037 (2021)
6. Xu, S.: BP neural network-based detection of soil and water structure in mountainous areas and the mechanism of wearing fatigue in running sports. *Arab. J. Geosci.* **14**(11), 1–15 (2021)
7. Meng, L., Gu, Z., Li, F., et al.: Application of BP neural network on quantitative evaluation of curriculum reform in advanced education-a case study of signal and system. *J. Phys. Conf. Ser.* **1774**(1), 012047 (2021)
8. Deng, H., Zhang, W.X., Liang, Z.F.: Application of BP neural network and convolutional neural network (CNN) in bearing fault diagnosis. *IOP Conf. Ser. Mater. Sci. Eng.* **1043**(4), 042026 (2021)
9. Peng, Y.: LLL aided MIMO detection algorithm based on BP neural network optimization. In: *ICAIS 2021: 2021 2nd International Conference on Artificial Intelligence and Information Systems* (2021)
10. Tang, Y., Su, J., Khan, M.A.: Research on sentiment analysis of network forum based on BP neural network. *Mobile Netw. Appl.* **26**, 174–183 (2021). <https://doi.org/10.1007/s11036-020-01697-y>





# Application of Computer Aided Technology in Kindergarten Interior Design Under the “New Crown” Epidemic

Shaowei Liu<sup>1,2</sup>(✉)

<sup>1</sup> Guangdong University of Science and Technology, Dongguan 523083, Guangdong, China  
liushaowei@gdust.edu.cn

<sup>2</sup> University of Perpetual Help System DALTA, 1740 Las Piñas, Philippines

**Abstract.** During the prevention and control of the new crown pneumonia epidemic, how to exercise scientifically and appropriately for children to promote physical fitness and develop healthy living habits has become an important research topic at present. This article combines the characteristics of homes during the prevention and control period of the new crown pneumonia epidemic, and proposes the idea of classifying children’s interior design from six aspects: total strength, cardiorespiratory endurance, physical coordination, responsiveness, balance stability, and flexibility and stretching, and specific game cases for each category. Provide evidence with exercise guidance, and make scientific recommendations for children’s home exercise load on this basis.

**Keywords:** Epidemic Prevention and Control · Young Children · Indoor

## 1 Introduction

0 ~ 6 years old is a critical period for children’s development, and a large number of neuronal connections change greatly during this period [1, 2]. The brain is the product of the interaction between the environment and genes. Learning, experience, stress, etc. will change the brain [3, 4]. Scientific research on the brain has found that the plasticity of the brain is lifelong, and that a person constantly reshapes his brain due to new experiences throughout his life [5, 6]. The function of the brain can be replaced when it is used and discarded. For example, in the learning of body movements, imitation is the most primitive and effective learning [7, 8]. The “mirror neuron” in the brain is like a mirror. It responds to external stimuli in the brain, and may imitate it in action. Therefore, through in-depth understanding of the learning mechanism of the brain, master the growth and development of the group of children grasp the game, visualization, and life features of the children’s sports design, follow the cognitive ability of the children’s brain, and strive to enrich the variety [9, 10]. Appropriate sports with simple and diverse forms lay the foundation for children’s all-round development.

As we all know, exercise is the best way to strengthen your body and an important means to enhance your immune system. In the special period of the prevention and control

of the new crown pneumonia epidemic, young children need to persevere in exercise and continuously improve their physical fitness. Based on the current situation of the broad masses of people fighting against the new crown pneumonia epidemic at home, taking into account the characteristics of home sports, the selected exercise form is simple and easy to use, without special equipment, and conforms to the growth and development characteristics of young children. Starting from the fun and interactive principles of children's game design, various simple and easy indoor contents are recommended. The game design includes both action descriptions and exercise recommendations to help children at home imitate games, actively participate, and keep exercising.

## **2 The Basic Elements of the Interior Space Design of the Kindergarten**

Kindergarten interior space design is very professional. How to find a breakthrough from the routine, innovate and explore the development direction of the kindergarten interior space design in the future is a question that all interior space designers should think deeply. The author believes that the interior space design of kindergartens should start with the following six basic elements:

- (1) The use of different materials. Not only must it meet the high-strength, high-weather resistance requirements of the kindergarten space, but also be healthy and environmentally friendly, which is an important expression of the atmosphere of the use space.
- (2) Color design of interior space. Proper use of color can often have unexpected effects. In the interior space design of the kindergarten, it is necessary to choose colors with close tones, and use less contrasting colors and dark and depressing colors such as black and ripe brown. Use the change of color to give the space a sense of rhythm and rhythm to achieve the harmony of space and color.
- (3) Based on the psychological and physical peculiarities of the users, the lighting design should also be properly measured. Too strong light will produce a strong stimulus to people, and too dark light will give people a feeling of dim, dull and even frustrated. The brightness of the light in the kindergarten space should maintain a soft and natural experience to help balance the mental state and emotions of patients and kindergarten teachers.
- (4) Furniture and furnishing design. In addition to aesthetics, the furniture in the interior space design of the kindergarten is more important to meet the behavioral needs of doctors and patients. With the improvement of the public's requirements for the environmental quality of kindergartens, the interior space design of many kindergartens pays great attention to furnishings and soft decoration design, and even displays artworks.
- (5) Indoor greening landscape design is also increasingly used in the interior space design of kindergartens to soften the spatial form and enhance the viewing and comfort of the environment.
- (6) Pay attention to the acoustic environment. The acoustic environment of the kindergarten public space will have a huge impact on the emotions and psychology of the

kindergarten teachers and patients. The soundproof walls of consultation rooms and wards should be built to the top of the structural board to avoid cross-interference of sound. The designer should carry out the acoustic design in the kindergarten in accordance with or higher than the requirements of the national design standards, so as to ensure that the quality of the kindergarten’s spatial acoustic environment is truly comfortable and pleasant.

The public space of the kindergarten is the first functional space for patients to enter the kindergarten. It has functions such as information prompts, guiding and diversion, and guiding everyone to quickly enter the traffic space or other functional spaces. It is a place with a large flow of people and the key to the image of the kindergarten. Therefore, when designing space, it is often given priority, including capital investment. For example, when designing the lobby of the outpatient and emergency building of Peking Union Medical College Kindergarten, the original design plan of the entrance hall was 4 floors. Later, according to the requirements of the construction party, the height of the hall was increased to 9 floors, and a glass roof was added, which achieved good results. Space visual effects and sense of experience. At the same time, a great price was paid for this. First of all, in terms of traffic, the increase in the number of elevators and the height of the building changed the building structure. The increase in floor height forces the indoor fire protection system to make corresponding adjustments, leading to chain adjustments in various electromechanical designs such as air conditioning and lighting. After the kindergarten was put into use, it was discovered that due to the change of space, the noise in the hall was high, and noisy sounds filled the entire indoor hall every day, greatly reducing the quality of the indoor environment. After on-site testing, it was found that the problem lies in the large area of smooth and hard glass, stone and other materials used on the roof and façade of the entire hall, while the area of the textured and sound-absorbing materials is too small to meet the sound absorption needs. The adjustment of acoustic design experts has improved this problem. It can be seen that in order to obtain the ideal spatial effect and experience, not only professional interior space design is required, but also the coordination of complicated related professional field design and the increase of engineering capital investment. “Therefore, the design of the indoor space of the kindergarten must adhere to the principle of moderation, and ensure the space effect under the premise of first meeting the use function.

Computer-aided technology can be defined as

$$\psi_{\alpha,s,t}(x) = a^{-3/4} \psi \left( S_s^{-1} A_\alpha^{-1} (x - t) \right) \tag{1}$$

In the formula,  $A_a = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}$  and  $S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$  represent the anisotropic stretching and translation matrix and shearing matrix respectively, and  $\psi_{\alpha,s,t}(x)$  is the computer-aided technology.

Assumed

$$\sum_{j \geq 0} \left| \widehat{\psi}_1 \left( 2^{-2j} \omega \right) \right|^2 = 1, \left( |\omega| \geq \frac{1}{8} \right) \tag{2}$$

For every  $j \geq 0$ , there is

$$\sum_{l=-2^j}^{2^j-1} \left| \widehat{\psi}_2(2^j\omega - \lambda) \right|^2 = 1, (|\omega| \leq 1) \quad (3)$$

It is easy to see from the support conditions of  $\widehat{\psi}_1$  and  $\widehat{\psi}_2$  that the function  $\psi_{j,l,k}$  has frequency support, ie.

The design of the hall and atrium of the kindergarten should not only solve the use function of the kindergarten space, but also pay attention to the psychological feelings of the kindergarten teachers and patients. The addition of appropriate natural lighting and green landscape in the space will bring comfort and relaxation to the people in it. Psychological feelings. Some kindergartens lacked activity space close to nature due to the limited site area. Therefore, when designing the interior space of their high-end ward building, they considered adding multiple three-dimensional indoor atrium gardens and giving different natural themes, making it feel like being in it. In nature. In addition, the design of the atrium garden also plays a good role in isolation. They divide the entire nursing unit into several parts, thereby effectively partitioning the nursing ward for easy management.

In recent years, due to the gradual improvement of Internet technology and kindergarten's intelligent information system, many patients will make appointments for registration online, which effectively reduces the flow of unnecessary people in kindergartens. At the same time, the information system also effectively reduces the number of registration, medicine, and medicine. The waiting time of the waiting procedure reduces the gathering of personnel. Therefore, it is foreseeable that the public space of large-scale kindergartens in the future will gradually decrease and the functions will be dispersed, directly connecting departments, reducing intermediate traffic links, effectively saving time and space, and reducing long-term queues and crowd waiting for patients. During the new crown pneumonia epidemic, human-to-human transmission occurred in many kindergartens, which exposed the inadequate design and management and operation of the infection zoning, hospital feeling control and other links that were emphasized in the previous article. This will affect the kindergarten buildings, indoors, etc. New requirements are put forward in the design process.

### 3 Interior Space Design Under Computer-Aided Technology

The traffic space that connects the kindergarten functional areas in series is equivalent to the kindergarten's "texture", including the kindergarten main street, the kindergarten main road, various consulting room passages, corridors and other spaces, which can be divided into horizontal traffic space and vertical traffic space. Optimizing the internal traffic organization of the kindergarten design is a link that designers should focus on. When designing many large-scale comprehensive kindergartens, they are accustomed to maximizing the volume of space. The Kindergarten in Chenggong New District, the First Affiliated Kindergarten of Kunming Medical College, built a nearly 100-m-long main street of the kindergarten during the design, connecting the main functional areas

such as outpatient clinics, emergency rooms, and ward buildings. The main street is 3 stories high and has a width of about 24. Meters, more than 30 m at the widest part. In consideration of the local climate, natural ventilation is adopted in the main street space of the kindergarten, and plants are planted to adjust the indoor microclimate. On the nearly 100-m-long main street, there are multiple sets of vertical traffic core tubes, which quickly divert and disperse patients entering from various halls. Multiple groups of leisure areas are set up in the main street, including coffee, water bars, etc. Facilities make it convenient for patients and their families to wait. The main street of the kindergarten is the “main artery” of the kindergarten traffic organization. It is extremely important in the design of the kindergarten traffic space. Because of the comprehensive and diverse functions, it can be appropriately segmented. Natural light, green plants or artworks are often used for design decoration, which is rich, narrow and dull Space (Fig. 1).



**Fig. 1.** Interior atrium design of a kindergarten

In recent years, with the innovation of the kindergarten’s architectural design and the continuous upgrading of the construction of the smart kindergarten, the kindergarten covers a large and important space, and its design is gradually evolving. Compress the length of the main street or traffic space and combine it with halls, halls and other functional spaces, blur the clear space concept, and even evolve into a complex traffic hall, combining with public space and large-scale traffic space, this intensive The type space can effectively shorten the traffic distance, save the space area, and save the investment cost. The Hebei Integrated Traditional Chinese and Western Medicine Children’s Kindergarten, which was put into use in 2018, was renovated to reduce the plane scale of the traffic hall, with three function halls connected in an arc shape. The project was designed according to local conditions, combined with the current status of the spatial

plan function, and combined the entrance hall with the traffic space. The arc structure integrates the three inherent buildings together to form a unique comprehensive hall. The facade of the hall is a stained glass curtain wall, which not only is visually beautiful, but also has good lighting, creating a lively and relaxed atmosphere for young patients and their families. The partial facades of the three inherent buildings become the walls of the indoor hall. The traffic core area of each building adopts the same colored glass design as the outer curtain wall, forming a visual effect of the indoor and outdoor spaces. After the kindergarten is put into use, its functions are reasonable and compact, and it runs well (Fig. 2).



**Fig. 2.** Kindergarten's main street of the kindergarten in Chenggong New District

The consultation room passages, corridors and other spaces in the traffic space of the kindergarten are also very important spatial components of the kindergarten building. In the design, not only the use function and the sense of hospitality control must be considered, but also the humanized care must be fully considered. Human behavior and psychology. For example, in long and narrow passages and corridors, there should be a segmented design, and a rest area should be reserved. Seats should be arranged at intervals of 20 to 30 m to facilitate patients to rest and buffer at any time. Due to the large flow of people, it is recommended to use durable and corrosion-resistant materials in this space. In addition, this type of space is located between various functional areas, and the daylighting is generally poor. Therefore, designers should make full use of artificial lighting and light-transmitting partition walls to create a transparent space. In the traffic space, attention should also be paid to the traffic sign guidance system, especially the design of the barrier-free guidance system to facilitate the convenience of medical treatment for all kinds of people, so as to achieve the function of efficiently connecting the waiting, lobby, hall and other functional spaces.

## 4 Kindergarten Interior Teaching Space Design

The teaching space includes the medical technology department and the clinics of the outpatient department. They and the nursing unit area (including the ward and the nursing station, etc.) occupy a relatively high proportion of the overall building space of the kindergarten. They are an important space for doctors and patients to interact with each other. And standardized features. There are detailed design specifications and specific requirements for the plan of this kind of space, and sometimes a professional kindergarten process design company is used as the content of the three-level process to design the layout. Due to the large amount of teaching space and the possibility of standardized design, it is more suitable for the assembled design that is now vigorously promoted by the country. In this epidemic, the prefabricated building design has been vigorously promoted. The building interior materials are produced by the factory and then transported to the site for installation. The construction of Vulcan Mountain and the kindergarten are all prefabricated designs in the form of containers.

In recent years, the use of “fabricated” in interior decoration projects has become more mature. As people pay more and more attention to the health of kindergartens, the market demand for kindergarten construction will gradually increase, and the time requirements for construction will become higher and higher. The most effective design and construction method to save time and cost is assembly. With the continuous upgrading and improvement of various technologies and materials, the entire ward, consulting room or corridor can be completed by assembly construction. The prefabricated design has many advantages in construction. It is widely used in the indoor renovation of new buildings, especially the existing kindergarten buildings. The combination with mechanical and electrical equipment, surface materials, etc. is also optimized, which saves the construction period and also saves costs.. In the future, prefabrication will be popular in kindergarten interior space design, especially in standardized interior space design such as nursery consultation rooms and ward nursing units. This is also a revolutionary advancement in interior space design and construction and will gradually change the design industry. The pattern and way (Fig. 3).



**Fig. 3.** The teaching space of a kindergarten

## 5 Conclusion

In recent years, the construction of kindergartens in China has been continuously developing at a high speed. The country's continued reform of the kindergarten system and the people's demand for kindergarten health services have promoted the kindergarten health market to continue to heat up. In addition, social capital has gradually intervened in the kindergarten industry, which has continuously improved the quality of China's medical care environment. Throughout the history of human survival and development, every major epidemic will usher in a progress in science, technology and ideology. With the continuous development of science and technology, kindergartens may undergo tremendous changes. The future development of kindergarten design changes in the way, and what remains unchanged is the advanced concept and forward-looking vision. The interior space design of the kindergarten is an important part of the construction of the kindergarten. There will be new developments.

**Acknowledgments.** Study on Kindergarten Interior Space Design Based on COVID-19 Epidemics.

## References

1. Shan, X., Wan, M.: Spatial color analysis of kindergarten interior design based on children's psychological activities in digital environment. *E3S Web of Conferences* **5**(9), 170–173 (2021)
2. Kreskey, D.D., Truscott, S.D.: Is computer-aided instruction an effective tier-one intervention for kindergarten students at risk for reading failure in an applied setting? *Contemp. Sch. Psychol.* **20**(2), 142–151 (2015). <https://doi.org/10.1007/s40688-015-0056-8>
3. Yan, L.: Research on the application of bim technology in computer aided architectural design. *Agro Food Industry Hi Tech* **28**(1), 1132–1136 (2017)
4. Goumeidane, A.B., Nacereddine, N., Khamadja, M.: Computer aided weld defect delineation using statistical parametric active contours in radiographic inspection. *J. Xray Sci. Technol.* **23**(3), 289–310 (2015)
5. Lin, H.-H., Chang, H.-W., Lo, L.-J.: Development of customized positioning guides using computer-aided design and manufacturing technology for orthognathic surgery. *Int. J. Comput. Assist. Radiol. Surg.* **10**(12), 2021–2033 (2015). <https://doi.org/10.1007/s11548-015-1223-0>
6. Ng, L.Y., Chong, F.K., Chemmangattuvalappil, N.G.: Challenges and opportunities in computer-aided molecular design. *Comput. Chem. Eng.* **228**(33), 14–19 (2015)
7. Celani, G., Sperling, D.M., Franco, J.M.S. (eds.): *CCIS*, vol. 527. Springer, Heidelberg (2015). <https://doi.org/10.1007/978-3-662-47386-3>
8. Zhao, W.: An application of bim technology in computer-aided building energy saving design. *Computer-Aided Design and Applications* **18**(S1), 133–143 (2020)
9. Akca, E.: Development of computer-aided industrial design technology. *Periodicals of Engineering and Natural Sciences (PEN)* **5**(2), 124–127 (2017)
10. Webster, R., Ottway, R.: Computer-aided design (cad) certifications: are they valuable to undergraduate engineering and engineering technology students? *J. Eng. Technol.* **35**(2), 22–32 (2018)





# Application of Data Mining Algorithm in Computer Mathematical Modeling

Yanqin Yang<sup>(✉)</sup>

College of Computer Information and Engineering, Shanxi Technology and Business College,  
Taiyuan 030006, Shanxi, China  
sxgsyyq@126.com

**Abstract.** The application of data mining algorithm in computer mathematical modeling is the process of applying data mining algorithm to problems that can be modeled by mathematical models. The application of data mining algorithm in mathematical model problems is called “data mining application” or “data mining implementation”. The application of data mining algorithm in computer mathematical modeling is the process of using data mining algorithm to transform, summarize and model information. This is an important aspect of data analysis, because it enables us to extract useful knowledge from a large amount of data. The goal here is not only to discover patterns, but also to determine what can be done once these patterns are discovered. Data mining applications are used to solve many practical problems that do not have direct access to basic mathematics, such as automated decision-making processes (e.g., recommendation engines), generating new hypotheses based on existing hypotheses (e.g., hypothesis testing), and improving existing theories and models (e.g., improving classification).

**Keywords:** Data mining algorithm · Mathematical modeling · Data extraction

## 1 Introduction

Mathematical modeling: it is a branch of mathematics. Theoretically, it establishes a mathematical model according to practical problems, solves the mathematical model, and then solves practical problems according to the results. The abstract and concise description of the essential attributes of a practical subject with mathematical symbols, mathematical formulas, programs, graphics, etc. can either explain some objective phenomena, or predict the future development law, or provide the optimal strategy or better strategy in a certain sense for controlling the development of a phenomenon. Mathematical model is not a direct copy of real problems [1]. Its establishment often requires people not only to deeply observe and analyze real problems, but also to flexibly and skillfully use various mathematical knowledge. This process of abstracting and extracting mathematical models from practical subjects is called mathematical modeling.

If you want to summarize and explain in one sentence, it is the process of summarizing similarities and differences from specific affairs and abstracting features. Take a

particularly easy to understand example: for example, when teaching children to understand three-dimensional graphics: only spherical three-dimensional graphics can only scroll; Called -- sphere, football, basketball, etc.; A solid figure with a sphere and a plane, which can be scrolled and moved; Called -- vertebral body, triangular cone, Christmas tree, etc.; A solid figure with a sphere and two planes, which can be rolled, moved and stacked; Called -- cylinder, thermos cup, foam roller, etc.; There is no sphere but only a plane solid figure, which can be moved and stacked; Called – cube [2] R, magic cube, pencil box, etc.; When children learn these three-dimensional figures, they summarize the characteristics: rolling, moving and stacking. What conditions they have, they will have what characteristics. Such a simple classification and statistics is actually the simplest mathematical modeling process. Such thinking training is to train logical thinking.

## 2 Related Work

### 2.1 Research on Computer Mathematical Modeling Contest

When applying mathematics to solve various practical problems, the establishment of mathematical model is a very critical and difficult step. The process of establishing mathematical model is to simplify and abstract complicated practical problems into a reasonable mathematical structure. Through investigation and data collection, we should observe and study the inherent characteristics and internal laws of the actual objects, grasp the main contradictions of the problems, establish the quantitative relationship that reflects the actual problems, and then use mathematical theories and methods to analyze and solve the problems.

Now more and more mathematical modeling competitions have entered the campus, offering various forms of mathematical modeling courses and lectures, which has opened up an effective way to cultivate students' ability to use mathematical methods to analyze and solve practical problems [3]. The National Undergraduate Mathematical Modeling Competition ° is co sponsored by the Higher Education Department of the Ministry of education and the Chinese society of industrial and applied mathematics °. The main criteria for competition and award are the rationality of assumptions, the creativity of modeling, the correctness of results and the clarity of written expression.

Data mining generally refers to the process of searching hidden information from a large amount of data by algorithm. Data mining is usually related to computer science, and achieves the above goals through many methods, such as statistics, online analysis and processing, information retrieval, machine learning, expert system (relying on past experience rules) and pattern recognition.

Mathematical modeling is highly valued in the school, and everyone feels it is necessary. The school has a high participation rate in both the national and American competitions. General competition team, computer professional students are dispensable Because there are MATLAB, lingo and other related courses, and it is easy to learn how to use them well (I feel that the computer major has no great advantage in the programming of modeling). One person can complete the modeling and programming work with clear ideas, and it is also good to discuss the algorithm model of mathematics together. Of course, it is very popular to learn computers and have a good command of

mathematics [4]. It is very important to be good at English in the American competition, because it doesn't matter whether you are studying computer or mathematics to write a thesis. It's ok if you can clarify your ideas and write a thesis.

I asked some students from the College of mechanical and electrical engineering. Not many people are interested in participating in mathematical modeling. Because there are competitions such as ACM, they are closer to their majors and have no idea about mathematical modeling.

Second, let's talk about the use of participating in mathematical modeling (simply speaking of this competition). At first, I felt that I had to do it because everyone did it. The teacher also said that it had a lot to do with the guarantee of postgraduate studies and the bonus of postgraduate entrance examination. The results were not top-notch, but winning the prize would guarantee a very good school. How So at the beginning of the University, it was a big goal to win the prize for modeling [5]. However, it has only recently come to realize that this amount of "welfare" is not as good as previously thought.

### 2.2 Data Correlation Analysis

When reviewing the association analysis algorithm before the exam, I read many blogs. Most blogs only introduced the Apriori algorithm of association analysis, but there are other methods rarely mentioned, such as the maximum frequent itemset method. Therefore, this article will introduce the association analysis algorithm in detail and attach the corresponding r code. After reading the article, you will have a try with the data, and you will have a clearer understanding of the association analysis algorithm [6].

Many people may have heard of a well-known book "beer and diapers". In the classic case, "beer" and "diapers", two seemingly unrelated commodities, were put together for sale and obtained good sales revenue. This phenomenon is the correlation between commodities in the store. There is also a case of "cigarettes" and "milk powder", which seems to have no connection, but in fact there is a certain connection.

$$Info(S) = - \sum_{i=1}^m p_i \log_2(p_i) \tag{1}$$

$$J_c = \sum_{i=1}^k \sum_{p \in C_1} ||p - M_i||^2 \tag{2}$$

To use Apriori algorithm, you first need to understand a concept:

Frequent itemsets: a collection of items that often appear together. When we use the Apriori algorithm, we will have a preset support, such as 3/5. The support of all items in this itemset is greater than the one we set.

1. Extract all frequent itemsets and their support from the data
2. Generate all effective association rules (confidence > minconf)

Then the problem arises. How to extract all frequent item sets and their support from the data? First, start with an example. After understanding this example, you will

understand how the algorithm is implemented: the left is the original data set of the transaction, and the right is the matrix composed of the number of occurrences of each item, as shown in Fig. 1 below.

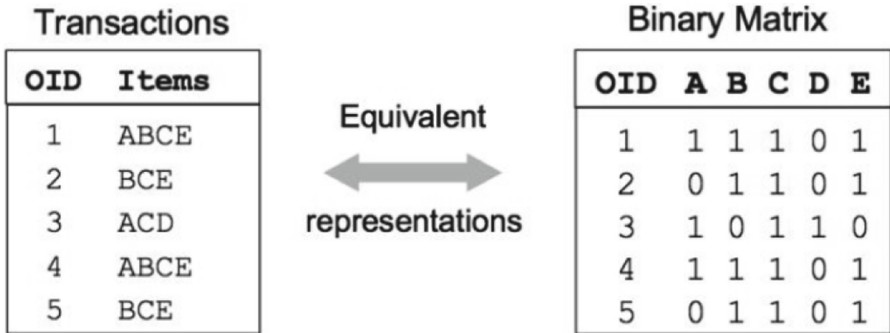


Fig. 1. Data set

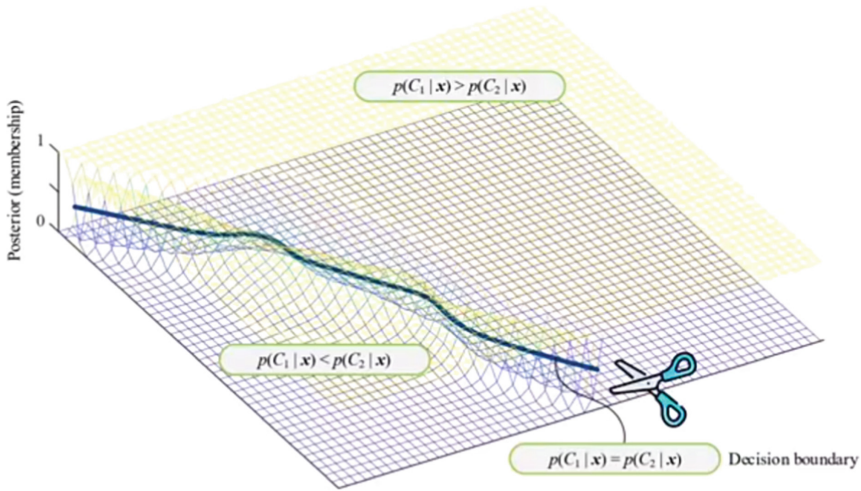
### 3 Research on the Application of Data Mining Algorithm in Computer Mathematical Modeling

From the common models of mathematical modeling, the models here can be used for many different types of real models. For example, the raindrop model just now is a simple model. It does not need a lot of data. It only needs to simplify the modeling of reality. However, with the development of the Internet, a very magical phenomenon has emerged, that is, big data [7]. This has resulted in a lot of work related to big data. So data analysis and data mining came into being.

What kind of relationship the data forms with each other, what kind of structure it is stored in, and how to facilitate query and retrieval, all of which are related to the data model.

In the process of data development, there have been three basic data models, which are hierarchical model, mesh model and relational model. These three models are named according to their data structures. The basic structure of hierarchical model is tree structure; The basic structure of the mesh model is an undirected graph without any restrictions. The relational model is an unformatted structure, which uses a single two-dimensional table structure to represent entities and the relationships between entities. One of the most widely used is the relational model [8].

Hierarchical model and mesh model are difficult to modify, retrieve and locate due to inconvenient data reading, which also restricts the volume of data to a certain extent. Now it is widely used in more relational data structures. Figure 2 below shows the binary four classification model of data mining.



**Fig. 2.** Data mining binary four classification model

Relational data structure Organizing data in the form of record groups or data tables is a very effective data organization method to establish the relationship between spatial data and attribute data, so as to facilitate the storage and transformation by using the relationship between various geographical entities and attributes, without layering or pointer [9]. The advantages are that the structure is particularly flexible and the concept is single, which meets the query requirements of all Boolean logic operations and mathematical operation rules; Be able to search, combine and compare different types of data; It is very convenient to add and delete data; It has high data independence and good security. The disadvantage is that when the database is large, it takes time to find the data that meets the specific relationship; The spatial relationship cannot be satisfied.

The process of analyzing the structure and laws of sample data sets is data exploration by means of testing the data quality of data sets, drawing charts, calculating some characteristic quantities and so on [10]. Data exploration can help to select appropriate data preprocessing and modeling methods, and even complete some problems usually solved by data mining.

## 4 Conclusion

The application of data mining algorithm in computer mathematical modeling is a process of extracting useful information from a large amount of data. The extraction process involves a series of steps, including: (I) data collection, (II) data cleaning, and (III) data preparation. In this article, we will discuss the methods used for these three steps. Data collection: this step refers to collecting or collecting relevant information from different sources such as the Internet, books, periodicals, etc., and then storing it in the database. The data can be organized into various categories, such as customer information, product functions and other similar types that may help in decision-making.

## References

1. Sun, X., Zeng, Y.: On the Application of Data Mining Algorithm in College Student Management (2021)
2. Yang, Y., Guo, C., Huang, J., et al.: Application of data mining technology in acupuncture prescription compatibility. In: International Conference on Application of Intelligent Systems in Multi-modal Information Analytics. Springer, Cham (2021). [https://doi.org/10.1007/978-3-030-74814-2\\_62](https://doi.org/10.1007/978-3-030-74814-2_62)
3. Zhang, J.: Application of data mining based on computer algorithm in personalized recommendation service of university smart library. *J. Physics: Conference Series*, **1955**(1), 012008 (2021)
4. Guo, C.: Application of computer technology in optimal design of overall structure of special machinery. *Mathematical Problems in Engineering* (2021)
5. Zhang, S.: Application of data mining technology in the analysis of e-commerce emotional law. *J. Phys: Conf. Ser.* **1852**(2), 022044 (2021)
6. Veresnikov, G., Skryabin, A.: The development of electromechanical actuator mathematical model for fault identification using data mining methods. In: 2020 13th International Conference Management of large-scale system development (MLSD) (2020)
7. Chen, Q.: Practical Application of Improved Algorithm of Association Rules Based on Computer Technology in Teaching Evaluation (2021)
8. Shang, Y.: Application of Mathematical Signature Technology in Computer Information Security Design. Clausius Scientific Press (4) (2021)
9. Saad H . The Application of Data Mining in the Production Processes[J]. Papers, 2020
10. A mathematical algorithm of the facial symmetry plane: Application to mandibular deformity 3D facial data. *Journal of Anatomy* (2021)



# Application of Hybrid Recommendation Algorithm in College Life Safety Education Management System

Shuchao Liu<sup>(✉)</sup> and Yuanyuan Xu

Chongqing Jiaotong University, Chongqing 400074, China  
lsc20120229@163.com

**Abstract.** How to establish a life safety education system and promote the development of life safety education is an important guarantee to ensure the growth and success of college students, as well as the practical and objective needs of cultivating qualified builders and reliable successors of socialism with Chinese characteristics. Life safety education is the value embodiment of “people-oriented” and an important content of “quality education”. Taking the essential attributes and core elements of life safety education as the logical starting point, the research examines the value orientation and functional orientation of life safety education curriculum from the multiple viewpoints and perspectives of life safety education, in order to explore the theoretical system of life safety education curriculum and its dynamic process of implementation in school physical education, Explore the security needs of today’s society and modern education for the individual life growth of young students. With the rapid development of Internet technology, people are constantly creating new information while obtaining information. In such an era full of massive data, users are easy to get lost in the ocean of information. The recommendation system is produced because of this actual demand. It analyzes the historical behavior data generated by users in the system, predicts users’ interests and preferences, and then pushes appropriate items to users. Based on the hybrid recommendation algorithm, this paper analyzes and studies the life safety education management system in Colleges and universities.

**Keywords:** Life safety · Education management · Hybrid recommendation algorithm

## 1 Introduction

As the carrier of human beings, life constitutes the foundation of the existence of the human world and the necessary condition for the survival and development of human society. Without the basis of human life, everything in the material world will become meaningless. “Without skin, how can hair be attached”? For everyone, all achievements in life are based on the cornerstone of life safety. If you lose your life, you lose everything. Psychologist Abraham h. Maslow’s hierarchy of human needs reveals the hierarchical

structure and progressive relationship of life, safety, belonging, respect and self realization, and expounds the concept of survival and development of human society, that is, after meeting the “physiological and security needs” of the basic level, Will pursue a higher level of “belonging needs, respect needs and self realization needs” [1]. Therefore, people-oriented, paying attention to the ontological needs of human life, ensuring life safety, obtaining life value and improving life quality are the call and needs of everyone’s soul in the process of modern social development. Similarly, life is the eternal theme of education. Education is conceived and formed by people’s self generation and self-improvement. The essence of education is for people’s development, care about people’s growth of a complete life, and pay attention to strengthening people’s spiritual belief, personality construction and personality cultivation.

The current value orientation of life education is to cherish life, nourish life, enrich the connotation of life, realize the desire of healthy growth of individual life and show the needs expressed by individual life [2]. Life safety education undertakes the function of linking life, safety and education, and injects the survival concept of human health and safety into the theory and practice of life safety education. Therefore, life safety education is a practical course with important value and indispensable in social development and modern education system [3].

In fact, life safety education is an ancient and emerging social practice education activity. Its antiquity lies in the history of life safety education as long as human history. In the ancient times when mankind evolved from ape man to man, with the formation of family and ethnic group society, people passed on survival and safety skills to the next generation through the natural method of teaching by example, forming the original form of life safety education; It is a new educational activity, because in the long process of human social development, there has never been a systematic and complete theory and method system involving personal life safety education [4]. Although since ancient times, a variety of personal safety protection skills have existed in martial arts, boxing and other confrontation sports in the East and West in the form of attack and defense fighting technology, for a long time, people have not improved and transformed personal safety skills into a scientific theoretical system and applied them in the field of school education. It was not until the 1970s that people awakened to life education and re recognized the value of life [5]. As an important research object, life safety education was concerned by scholars, and its educational theoretical research and practical activities were gradually carried out. Especially in recent years, under the modern educational concept of “people-oriented” and the value pursuit of “quality education”, Maintaining life dignity, improving life quality and ensuring life safety have become the consensus of modern society. The role and function of life safety education are increasingly recognized and valued by people, and has become a curriculum research hotspot of education reform all over the world in recent years.

## **2 Related Work**

### **2.1 Collaborative Filtering Recommendation**

Collaborative filtering is the most famous and mature recommendation algorithm in the recommendation system. The idea of the algorithm is to find the items of interest for users.



First, find other users with the same interest as the target user through historical behavior, and infer that users have similar interests in the future according to their similar hobbies in the past, Then recommend their favorite projects to target users. Recommendation algorithms based on collaborative filtering are mainly composed of two types: user based collaborative filtering (usercf) and item based collaborative filtering (itemcf) [6]. Therefore, based on this principle, the user based collaborative filtering recommendation algorithm is mainly divided into three steps: one is to calculate the interest similarity between users according to a certain algorithm, the other is to find the user set similar to users' interests and hobbies; The third is to get the items that the target user may like from the user set with similar interests to the target user, and then generate recommendations. Here, it is divided into five small steps:

Step 1: establish a user item scoring matrix;

Step 2: select or design a certain algorithm to calculate the similarity of interests between users;

Step 3: according to the similarity calculated in the previous step, sort the similar users of the target users according to the similarity value, and select the first n users as the interest nearest neighbor set of the target users;

Step 4: according to the similarity between users and the degree of interest in the project, predict the score value of the target user for the project not involved;

Step 5: select TOPM items according to the scoring value in the previous step, so as to recommend them to users.

User based collaborative filtering (usercf) has the following advantages: (1) it does not need to consider the content factors of recommended items and can process unstructured data such as video and music; (2) Without considering the implicit information of the user, only the user's display score of the item is used as the core data of the recommendation, and the feature vector in the item content is not considered, which makes the recommendation easier; (3) The score is used to calculate the similarity between users. Among the items loved by neighbor users, there are items that the target users are potentially interested in, which makes the recommendation system diverse. However, the user based collaborative filtering recommendation algorithm also has some problems: (1) in the current era of information explosion, the number of users and items increases rapidly, resulting in the construction of user item one scoring matrix becoming very sparse, and the proportion of items scored by users will also be very low, resulting in inaccurate recommendation results: (2) when new users join the recommendation system, Because there is no behavior, the recommendation cannot be obtained, so there is the problem of "cold start". (3) Because users and projects are growing in geometric form, the computational efficiency of the algorithm is low.

## 2.2 Content Based Recommendation

Content based recommendation is a hot research topic in the field of information retrieval. Its application time is also the earliest. The recommendation process is mainly based on the attributes and characteristics of the project. The dependent assumption is that the characteristics of the item can be extracted through the content of the text itself, so the core of the algorithm is to find the item in which the user has generated behavior, then analyze its content and find the key information in which the user is interested. Use these

marks to model the interest of the target user, then calculate the similarity between the item and the interest preference modeled by the target user, and finally give the predicted score value topn items to the user to form a recommendation[7]. At present, content-based recommendation algorithms use text extraction to extract the characteristics of items, and mainly use some algorithms in data mining, such as Bayesian classification, decision tree, cluster analysis and so on. The formula of the final predicted value is

$$u(c, s) = score(Content(c), Content(s)) \quad (1)$$

The advantages of content-based recommendation algorithm are; (1) There is no need to establish a user item scoring matrix, which is not constrained by the number of user scores, so there is no problem of data sparsity; (2) because the similarity is calculated by calculating the interest preferences modeled by the project and the target user, and there is no need for the user's scoring data, recommendations can still be generated when new users join the recommendation system, so there is no problem of cold start; (3) The recommended method can be explained and understood simply, which only depends on the characteristics of the project content itself; (4) The technology of text content extraction is mature, so it is easy to implement. However, the content-based recommendation algorithm also has some disadvantages: (1) as shown by its advantages, the algorithm relies heavily on extracting the features of items, while it is difficult to extract the feature attributes for some music and video data, so it relies heavily on structured item data; (2) User preferences also need to be easy to extract and can be described in the form of project features; (3) The user's scoring data of the project is not used, resulting in serious waste of data resources and loss of recommendation accuracy.

### 2.3 Mixed Recommendation

Based on the previous several recommendation algorithms, it can be seen that each recommendation algorithm has its advantages in a specific field, but also has some defects in some aspects. Therefore, some scholars proposed to combine a variety of recommendation algorithms into new algorithms to learn from each other, so as to get an algorithm with better recommendation effect. For example, the content-based recommendation algorithm does not have the problem of cold start, while the recommendation algorithm based on collaborative filtering has the problem of cold start. Therefore, mixing the two can give play to their respective advantages [8]. Common hybrid recommendation algorithms include algorithm switching, weight change, feature combination, algorithm layering, feature expansion, meta level and so on. Switching refers to selecting the appropriate recommendation algorithm for different problems according to the scene of the recommendation system and other environmental factors, using the content-based recommendation algorithm to solve the cold start problem, and then switching to the recommendation algorithm based on collaborative filtering when the scoring data is to be used; Weight change refers to the final addition of different weights accounted for by multiple recommendation algorithms, and the recommendation is made according to the final score; Algorithm layering refers to the serial execution of the recommendation process. After the execution of one recommendation algorithm, the result is generated, and then the result is used as the input data of the next algorithm to recommend according

to the final result. In addition, feature combination is to combine the features of different data sets according to a certain algorithm, and then use the combined features as the required features of the algorithm, Recommend the final generated results; Feature expansion refers to considering the characteristics of multiple recommendation algorithms at the same time, not just the characteristics of a single technology; Meta level refers to the input of recommendation algorithm, which can also be modeled by other recommendation algorithms [9].

### 3 Current Situation of Life Safety Education for College Students

College students are the pillars of the country in the future. The increasingly complex and rapidly changing society makes it difficult for college students to bear the fierce competition and high pressure life in the real society. In the face of setbacks and difficulties, they often choose to escape and dodge, resulting in psychological distortion, psychological fear, and even contempt for life, self mutilation of life and injury to the lives of others. College students lack life safety awareness, despise their own and others' lives, and ignore the social significance and value of life.

#### (1) Lack of life safety awareness

Life safety consciousness means that college students should understand what life safety is and the importance of life safety. The most precious thing is life. Life is only once for everyone. When it comes to life, we cannot fail to associate it with safety. Nowadays, college students have a weak sense of life, and suicide occurs frequently in Colleges and universities. It is not only increasing year by year, but also has an obvious "chain effect". That is, after learning that others commit suicide, some students with original suicidal tendency will follow its way to end their lives. The "Liu Haiyang incident", which once caused a sensation in the country, and the "cat abuse incident" of master students of Fudan University strongly show that some college students not only lack the awareness of life safety, but also despise their own and other people's lives. The questionnaire investigates whether the people around college students have life safety awareness. The data show that 47.1% of the students think they lack life safety awareness, 37.7% of the students think their parents lack life safety awareness, 39.5% of the students think the school lacks life safety awareness, and 23.7% think they, their parents and the school lack life safety awareness.

China vigorously advocates people-oriented quality education and pursues the all-round and coordinated development of students' morality, intelligence, physique, beauty and labor. However, facing the realistic pressure of entering a higher school, quality education has gradually become a mere formality. At present, due to the imperfect school system, poor infrastructure and high teaching pressure, colleges and universities in China have caused the lack of life safety knowledge, weak life safety awareness and poor escape ability of college students, so they can not deal with sudden safety events. The lack of understanding of life safety education in schools makes college students lack correct life values, resulting in some college students' lack of passion for life, loss of sense of responsibility and

indifference to life. In the face of crisis situations, it is easy to be psychologically distorted, so as to endanger the life safety of oneself and others.

(2) Imperfect life safety education system

A good life safety education system is the institutional guarantee for the operation of life education. Without a sound system, implementation will be restricted. At present, the system of life safety education in Colleges and universities is almost zero. The questionnaire shows that only 1.3% think the school life safety education system is perfect. Most colleges and universities not only do not include life safety education in the teaching plan, but also have little life safety education for college students in their daily study and life. The imperfect life safety education system directly leads to the lack of time guarantee, process implementation and effect evaluation of life safety education in Colleges and universities. In addition, the lack of life safety education system makes teachers fundamentally do not pay attention to life safety education[10]. Even if some teachers occasionally carry out life safety education, it can not be carried out systematically.

(3) Life safety education methods are not comprehensive

At present, the methods of life safety education for college students are very old and backward. The existing life safety education methods are mainly teachers' teaching according to the book, the posting of some life safety knowledge in the school bulletin board and the sudden learning after the life safety threatening events around. This undoubtedly leads to the passive situation of life safety education. The old and limited educational means make it impossible to carry out comprehensive, centralized and in-depth life safety education for all college students. After investigating 500 college students, the results show that most schools carry out life safety education for students by means of daily indoctrination, teacher teaching and poster posting, and only 13.1% of students have received field exercise education, as shown in Fig. 1.

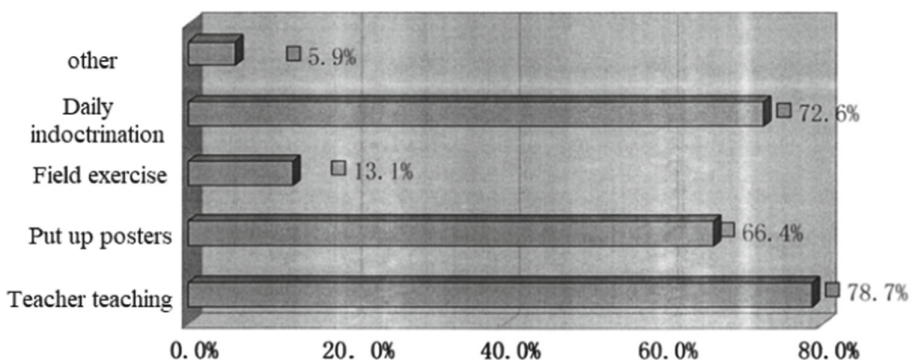


Fig. 1. Current situation of life safety education methods

(4) The content of life safety education is not rich

At present, life safety education in Colleges and universities usually includes fire prevention education, traffic safety education, dormitory electricity safety education, etc., and there is an extreme lack of education in disaster prevention education and mental health education. The reason why the Japanese people can evacuate in an orderly and rapid manner in the face of danger in the event of a major earthquake is that they have been regularly trained in disaster prevention since childhood. After the Wenchuan earthquake, China called on everyone to carry out disaster prevention exercises, but it was a flash in the pan. Disaster prevention education requires us to take preventive measures, rather than learning to make amends after disasters and injuries. At the same time, college students' mental health and personal safety education is also very insufficient. Many colleges and universities have set up institutions for mental health counseling for teachers and students, and set up a psychological education committee in each class. It seems to be in full swing, but in fact, most colleges and universities are mere formality and do not really carry out regular mental health education for college students. Due to the lack of basic knowledge of life safety, college students will not protect themselves in the face of injury, which makes life injury events occur frequently. The questionnaire investigates the current situation of the content of life safety education. The results show that at this stage, the content of life safety education in Colleges and universities is mainly fire education, traffic safety education and personal safety education, but there is a lack of disaster prevention education and mental health education.

#### **4 Countermeasures and Suggestions for Life Safety Education in College Teaching**

(1) Introducing the concept of safety education in the teaching process of colleges and Universities

College students often do not have a good understanding and attention to the value of life. They will commit suicide or even endanger the life and health of others due to their poor study, failure in emotional life and disharmony with their classmates, causing irreparable harm to their own and other people's life safety. Therefore, college teaching should pay attention to cultivating students' positive attitude, Educate them to face life with an optimistic and positive attitude. Teachers should fully show the vitality and brilliance of life to college students. In terms of teaching content, they should not only exercise the body of college students, but also stimulate their own interest in sports, so that they can exercise independently and actively participate in activities, so that they can realize that the existence of life is a beautiful thing, Fully enjoy the wonderful life and the fun of sports. Teaching should make use of natural resources and properly implement outdoor sports, so that college students can exercise their sports ability in nature and protect themselves from injury.

(2) In the process of teaching, we should create life classroom situation

Knowledge comes from practice and serves practice. It is closely related to the reality of life. As an effective teaching method, affective teaching method emphasizes the emotional field of college students, which is often ignored by traditional

teaching. Teachers should start from the reality of college students and creatively create some life-related and life situations. These situations should be suitable for students' knowledge base, understanding level, life reality and age characteristics, so that college students can understand, want and learn. The use of situational teaching, according to the curriculum objectives and the needs of life education, leads to the life problems in the problems, so as to make college students feel the existence of the problems, cause a thirst for knowledge and Reflection on life, and make college students actively and willingly invest in learning and exploration. Because this situation is very close to the reality of students' life, college students are relatively easy to accept, so that on the basis of students' initiative, it promotes the acquisition of College Students' knowledge, the practice of relevant life education in the curriculum and the realization of teaching objectives.

(3) Teachers pay attention to life with development vision and inclusive attitude

Teachers pay attention to life with development vision and inclusive attitude. In teaching, teachers' task is to find and develop students' strengths. In the process of teaching, teachers should abandon the differential treatment and attitude towards "poor students" and "excellent students". There are no absolute norms for technology and skills in teaching, and the evaluation of skills is of little significance to ordinary college students. Physical education teachers should realize that the learning of skills is to serve the development of life. At the same time, in the learning process of the same sports skill, with the same teaching method and the same time, the results are different for college students with different sports foundation and different comprehension ability. Therefore, teachers should pay attention to life and students' feelings with the vision of development and inclusive attitude in the teaching process. Teachers should make college students experience the joy of life, make them understand the connotation of appreciating sports, avoid verbal criticism, and create a classroom atmosphere of praise and appreciation as much as possible, which is conducive to the development of College Students' physical and mental health.

(4) Construction of life safety curriculum system

To fear life, we must improve safety awareness. Life safety education can be incorporated into the teaching content of physical education to build a life safety curriculum system. Because physical education is not only for students' physical health and cultivating lifelong physical education. In the process of curriculum implementation, we should pay more attention to the combination of theoretical teaching and practical teaching. On the one hand, we should educate college students to cherish life and pay attention to protecting the safety of individuals and others. They should be people-oriented whether they participate in classroom or extracurricular activities, First, consider safety factors, and observe and think about whether there are accidents and dangers. On the other hand, we can carry out life safety education for college students by simulating the real scene, and guide college students to learn self-protection. For example, in gymnastics teaching, we can organically combine pull-up and climbing, and combine high jump and long jump with acute jump; In Wushu teaching, we can learn some knowledge and skills of self-defense and escape; Only by carrying out simulation teaching in different situations can we attract more college students' attention and interest, consolidate

and apply the life safety knowledge and skills of the University, not only improve college students' enthusiasm for physical education, but also improve the skills of life safety protection and defense in practice.

## 5 Conclusion

At present, under the influence of the trend of "rational education", the transmission of knowledge is more important than the realization of the educational value of knowledge, rational training is more important than the cultivation of rich life safety, strict procedures and order are more important than the attention to creativity, and students' test scores are more important than the improvement of students' individual quality of life, The preparation for students' future life is more important than the care for students' real life. Higher education should always adhere to "people-oriented" "To protect students'" life safety and improve students' quality of life. We should strongly educate and guide students to pay attention to life, how to ensure their own safety in danger and maintain the life safety of others, shape good psychological quality, resist pressure. College physical education, as an important stage for students to systematically learn scientific physical education knowledge, should improve the new model of physical safety education and promote students' development Improve the overall quality.

**Acknowledgement.** Education Reform Project of Chongqing Jiaotong University: Construction of life education Curriculum system for college students in the new era (Number:2020321073).

## References

1. Gao, Y., Liang, H., Sun, B.: Dynamic network intelligent hybrid recommendation algorithm and its application in online shopping platform. *J. Intelligent Fuzzy Syst.* **1**, 1–13 (2021)
2. Kurt, Z., Gerek, M.N., Bilge, A., et al.: A Multi Source Graph-Based Hybrid Recommendation Algorithm (2021)
3. Cai, B., Zhu, X., Qin, Y.: Parameters optimization of hybrid strategy recommendation based on particle swarm algorithm. *Expert Syst. Appl.* **168**(12), 114388 (2021)
4. Dwa, B., Yan, C.A.: A novel cascade hybrid many-objective recommendation algorithm incorporating multistakeholder concerns - ScienceDirect. *Information Sciences* (2021)
5. Cui, Y.: Intelligent recommendation system based on mathematical modeling in personalized data mining. *Math. Probl. Eng.* **2021**(3), 1–11 (2021)
6. Li, Y., Wang, M., Tan, X., et al.: Application of hybrid silicate as a film-forming agent in high-temperature water-based drilling fluids. *ACS Omega* (2021)
7. Duan, C.: A deep Hybrid Recommendation Algorithm Incorporating an Attention Mechanism (2021)
8. Kk, A., Nv, A., In, A.: Hybrid Image Recommendation Algorithm Combining Content and Collaborative Filtering Approaches (2021)
9. Wang, X., Dai, Z., Li, H., et al.: Research on hybrid collaborative filtering recommendation algorithm based on the time effect and sentiment analysis. *Complexity* **2021**(2), 1–11 (2021)
10. Dong, Z., Leng, C., Zheng, H.: Employment Service System Based on Hybrid Recommendation Algorithm (2021)
11. Yan, C., Men, Y., Liu, W., et al.: Research on Hybrid Recommendation Algorithm of Insurance Products Based on Time Weighted Optimization (2021)



# Application of Improved Genetic Algorithm in Distribution Network Planning with Distributed Generation

Yuanhong Hou, Chao Feng<sup>(✉)</sup>, Yougui Aer, Shunhai Xue, Yuqiang Zhou,  
and Hongchao Wang

State Grid Qinghai Electric Power Company Huanghua Power Supply Company,  
Jianzha 811200, Qinghai, China  
qh Fengchao@163.com

**Abstract.** Distributed generation (DG) technology, especially the combination of renewable energy generation and conventional large-scale power supply, is an important growth point of modern power technology, which can realize the efficient use of clean energy. In this project, we will use genetic algorithm (GA) for distribution network planning. We will use distributed generation in the model, so GA will be applied to it. The main idea of this project is to find a good solution with the lowest cost and the largest profit by applying genetic algorithm on the distributed generation model. In other words, we hope to find an optimal solution to maximize profits and minimize costs while keeping energy consumption at a minimum level.

**Keywords:** Improved genetic algorithm · Distribution network · Distributed · Power Supply

## 1 Introduction

Distribution network is an important part of power network, and also an important part of urban and rural production and life. Its planning results will directly affect the economy of power grid operation and the reliability of load power supply. After entering the 21st century, it is of great practical significance in improving the economy of the power system to carry out reasonable and scientific distribution network planning with the planning first electric power construction method, thereby greatly improving the power quality and power supply reliability, saving energy and reducing energy consumption. The rapid development of economy and the improvement of living standard make people's demand for electric energy more and more large. At the same time, the requirements for power quality and power supply reliability are also getting higher and higher, which also makes the smart grid develop rapidly, especially in terms of expanding the scale of distribution network and constantly updating technology [1].

With the rapid expansion of cities in China, it is common that urban power grids cannot fully meet the requirements of reliability and power quality, especially the rapid growth of load power consumption, and the requirements for power supply quality and



load reliability are increasing. In recent years, the medium and low voltage power grids, which are the weakest and most difficult to reconstruct in almost all large and medium-sized cities in China, are under great pressure, and the peak load in summer has increased year after year [2]. It can be seen that in areas where the urban distribution network construction is backward, it is urgent to accelerate the construction of distribution network and expand the distribution capacity. Traditional distribution network planning, which mostly adopts manual calculation, mainly focuses on scheme comparison. The main idea is to select the optimal scheme from the given feasible schemes through technical and economic comparison of several schemes. Due to limited conditions, the schemes participating in the selection do not necessarily include the best scheme quoted, but are proposed by planners based on experience [3]. Therefore, the final recommended scheme can no longer meet the requirements of modern power system, and it has certain subjectivity and limitations.

## 2 Related Work

### 2.1 Research Status of Distribution Network Planning

In terms of voltage level, the power system mainly includes 500 kV, 330 kV, 220 kV, 110 kV, 35 kV, 10 kV and 380 V. According to the provisions of the Guidelines for Urban Power Network Planning and Design, the ultra-high voltage and ultra-high voltage of 500 kV and above are transmission voltage, the high-voltage distribution includes 35 kV and 110 kV voltage levels, 10 kV and the recently studied 20 kV are medium voltage distribution voltage levels, and 380 V is low-voltage distribution voltage levels. With the continuous expansion of urban distribution capacity and power supply scope, some mega cities such as Beijing and Shanghai have introduced 220 kV voltage into urban areas for power distribution, and 500 kV ultra-high voltage substations have gone deep into urban areas for power supply [4].

Under the condition of meeting the power supply capacity and various power operation technical indicators required in the planning period, the distribution network planning is to determine when and where to build what type of lines and the number of circuits under the condition of minimizing the construction and operation costs of the system [5]. Mathematically, distribution network planning is an optimization problem with the following five characteristics:

- (1) Discreteness. The number of lines is the decision variable based on the number of lines erected, and the value of the decision variable must be discrete and integer.
- (2) Dynamic. On the basis of meeting the technical and economic indicators in the target year, the distribution network planning should also consider the realization of performance indicators in the continuous development of the distribution network.
- (3) Non linear. The electrical parameters, network power flow, line power, active power loss and other electrical and economic indicators of the line are nonlinear.
- (4) Multi objective. In the distribution network planning, it is necessary to consider not only the impact of natural factors such as society and environment, but also the economic and technical indicators that are contradictory to each other.

- (5) Uncertainty. The equipment availability and load forecasting in distribution network planning are uncertain.

It can be seen from the above analysis that the essence of distribution network planning is a nonlinear dynamic multiple mixed integer programming problem with uncertain objectives. Some conditions must be assumed and simplified, and many corresponding planning models have been formed according to different simplification methods to solve this complex problem.

## 2.2 Distributed Generation and its Impact on Power Grid

Distributed generation (DG) can be directly connected to the distribution network system for grid connected operation or independent operation. The name of DG is derived from the centralized large power grid, which emphasizes that it is different from the centralized large power grid. At present, although there is still no specific, complete and accurate definition of the specific meaning and content of distributed generation, and even the terms at home and abroad are quite different, the main characteristics of distributed generation are the local digestion capacity, which is independent of the scale of the power station and the type of power generation. It can be used as power system load peak shaving, power supply for remote areas or important commercial and residential areas, which can greatly save the investment cost of power transmission and transformation, improve the reliability of load power supply, etc. [6]. Conventional distributed micro power sources include standby small diesel generators, small hydro generators, solar photovoltaic power generation, power generation equipment installed in the power grid to provide voltage support or improve power supply reliability, small power generation equipment configured for users and their surroundings, and power generation facilities installed in or near the load center [7]. The connection diagram of a simple grid connected distributed generation system is shown in Fig. 1.

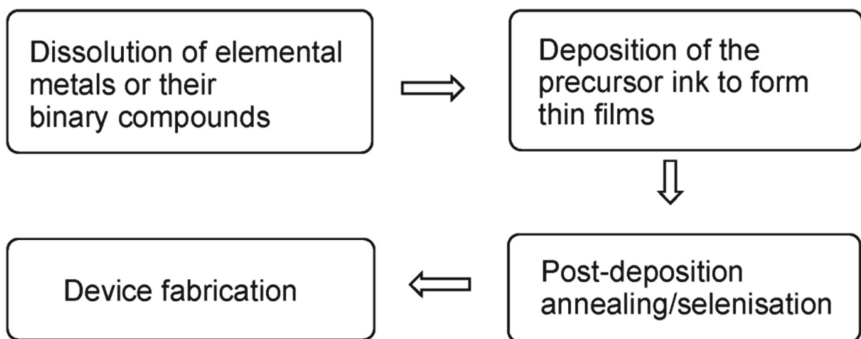


Fig. 1. Schematic Diagram of Distributed Generation

Distributed power generation is not a new kind of power generation. In the past, in some hospitals, mines and other important departments or places, users often installed small diesel generators as emergency backup power. In the early days of China, self

owned power plants and small thermal power plants, which mainly used coal, were also generally included in the scope of distributed power generation, but were limited by poor technical performance or low efficiency or greater impact on the environment, Such power supply is gradually eliminated or replaced [8].

### 3 Mathematical Model of Distribution Network Planning

The main contents of distribution network planning include substation address selection, distribution line connection design and line type selection. The selection of substation site refers to making decisions on the substations that may be arranged in the planned power grid, so that there are sufficient technical data for comprehensive economic and network loss comparison in the process of distribution network planning, so as to select the substation site and capacity. Understand the role and function of substation in the system, that is, system hub substation, regional important substation or intermediate substation and terminal substation of general substation. Generally, the substation shall be close to the load center or power supply area, so as to facilitate transportation and meet the needs of recent construction and development [9]. In the actual planning of distribution network, because the selection of substation site is easily restricted by factors such as land approval, residential relocation, etc., there is little room for selection and optimization.

The “distribution point” problem of distribution network planning is solved after the substation address is determined. The remaining problem is the connection between substations and power plants. According to the different objective functions of distribution network planning, the objective mathematical models of distribution network planning are discussed respectively.

The distribution network planning with the lowest cost is to minimize the annual operation and investment costs by deciding when and where to set up the number of distribution lines in the planning period when and where the power generation status, load demand and substation establishment time and location have been determined. The objective of planning here is to minimize the investment cost and operation cost of the line. In the model, the capital value of time is considered, that is, the annual cost of discount is considered to be the minimum, that is, the objective function is shown in Formula (1).

$$\min F = \frac{i(1+i)}{(1+i)^n - 1} \quad (1)$$

For the distribution network planning considering distributed generation, the first consideration is the impact of economic and environmental factors on the load surroundings. While meeting the above two basic conditions, further discuss the relevant planning of distribution network. In traditional power grid planning, three steps are generally followed: load power prediction, power source design and network planning. After the connection of distributed generation phases, not only the predicted load but also the load distribution should be clear. According to different objectives, distribution network planning including distributed generation can be roughly divided into two parts:

distribution point planning of distributed generation in the power network and distribution network expansion planning based on distributed generation [10]. The former takes planned power supply as the starting point, and the latter takes network structure planning as the goal.

#### 4 Application of Improved Genetic Algorithm in Distribution Network Planning with Distributed Generation

When calculating the fitness of the traditional 0–1 vector coding mode, we need to decode according to certain rules. If the number scale is larger, the decoding process will be more computationally expensive and will waste more time. We use matrix coding to make the calculation more simple and fast, as shown in Fig. 2.

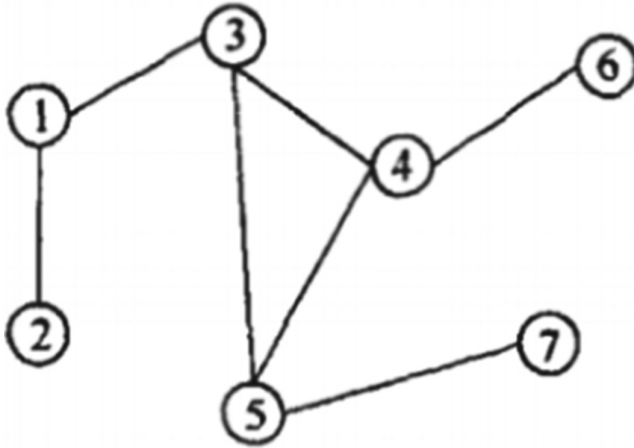


Fig. 2. Improved genetic algorithm matrix coding

First, define: the column of the node where the power or current flows out. Then, the column refers to the position of the corresponding node in which the current from the node or the power flows in. The relevant equivalent column refers to: if one node is listed in another node, and the two nodes are directly connected, then the two nodes are equivalent to the other. Specific operation method: write the column from small to large according to the size order of the node number. At the same time, search the occurrence times of the listed node. When it is greater than 1, we will use the equivalent node whose occurrence times are 0 to represent it. We can construct a matrix B. Then, according to matrix B, we can get that the new node can only be connected to one of the connected networks, which will form a ring network. If there is no connection, it will form an isolated network, which can be expressed by a matrix: if the requirements for network connectivity and radiation are met, then each node is out of the column if and only if there is one listed node. Then we repair B, we search the corresponding node according to the number from small to large, and then repair the network. The following

describes the calculation process. The improvement of the genetic algorithm in this article includes the following points: (1) To prevent the cross of close relatives, that is, in each iteration process, when we increase the matrix comparison, when calculating the Euclidean distance between them, we need to remove the value corresponding to the value with too small distance. (2) We need to increase the brother comparison link in the comparison link. It is necessary to make a corresponding comparison between the chromosomes of the offspring generated by the chromosomes of each parent generation, so as to further remove unreasonable values to reduce the search scope. (3) We should try to avoid repeated comparison process. We can add a bad cluster to count the nodes that have been eliminated, avoid duplication, increase the amount of computation, and increase the efficiency of computation.

## 5 Conclusion

Through the above analysis, we can draw a conclusion that the coordination planning framework obtained by adding DG in the distribution network planning process can effectively solve their interaction in the planning process and calculate its feasibility. At the same time, when environmental benefits are added, we carry out an environmental incentive mechanism for more environmentally friendly DG power generation, which saves resources and conforms to the strategy of sustainable development of recycling courses that we have always advocated. In this paper, we use the improved genetic algorithm to solve a series of problems in the coding of computing nodes efficiently and quickly. Compared with the traditional coding algorithm, it has the advantages of less computation, shorter computing time and higher computing efficiency.

## References

1. Hu, Y.: Energy storage expansion planning method for active distribution network based on improved particle swarm optimization algorithm. IOP Publishing Ltd (2022)
2. Song, H., Cai, M., Cen, J., et al.: Application of improved adaptive genetic algorithm in energy consumption optimization of electric refrigerated vehicles. In: Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. SPIE (2021)
3. Gao, C., Chen, Y., Tang, X.: Research on distribution route planning model and algorithm of unmanned aerial vehicle (UAV) based on Improved Multi-objective Genetic Algorithm. (2021)
4. Wu, F., Xin, B., Liu, T.: Urban green economic planning based on improved genetic algorithm and machine learning. *J. Intell. Fuzzy Syst.: Appl. Eng. Technol.* **4**, 40 (2021)
5. Yang, D., Yu, Z., Yuan, H., et al.: An Improved Genetic Algorithm and Its Application in Neural Network Adversarial Attack. arXiv e-prints (2021)
6. Tw, A., Zy, A., Xw, A., et al.: Improved Distributed Optimization Algorithm and its Application in Energy Saving of Ethylene Plant. (2022)
7. Wang, H., Wang, Y., Lv, X., et al.: Genetic algorithm with local search for the multi-target scheduling in flexible manufacturing system. *J. Circuits, Syst. Comput.* **31**(16), 2250279 (2022)
8. Sun, F., Shi, G.: Study on the application of big data techniques for the third-party logistics using novel support vector machine algorithm. *J. Enterp. Inf. Manag.* **35**(4/5), 1168–1184 (2022)

9. Solat, S., Aminifar, F., Shayanfar, H.: Distributed Generation Hosting Capacity in Electric Distribution Network in the Presence of Correlated Uncertainties (2021)
10. Andoni, S., Moore, K.D., Bonab, E.M., et al.: Execution of a Genetic Algorithm with Variable Evolutionary Weights of Topological Parameters for Neural Network Generation and Training., US11106978B2[P]. (2021)



# Application of Mathematical Simulation Method in Phosphate Ore Dressing

Lingpan Du<sup>(✉)</sup> and Yongjie Guo

National Engineering Research Center of Phosphate Resources Development and Utilization,  
Yunnan Phosphate Chemical Group Co., Ltd., Kunming 650600, China  
20832814@qq.com

**Abstract.** The influence of different factors on phosphorus content and quality of phosphate rock is calculated by mathematical simulation method in phosphate ore dressing. The main purpose is to find out how much phosphoric acid can be produced from a given amount of phosphoric acid ore. It also helps to find the best way to treat mine waste so that it can be processed into phosphoric acid cheaply and efficiently. In addition, it helps to determine whether the existing mine waste treatment plants are sufficiently effective in treating mine waste into phosphoric acid. The simulation method is used to determine the concentration of phosphate in the solution and can be applied to all types of mineral processing. The simulation method is based on the fact that phosphate ore will dissolve into solution at a certain rate, which depends on its chemical composition. Application: it is widely used to determine the concentration of phosphate in solution by using mathematical simulation methods such as empirical formula calculation and stoichiometric analysis.

**Keywords:** Digital analog · Phosphate rock selection · Space mineral deposits

## 1 Introduction

With the social and economic progress, the demand for resources continues to increase, resulting in the continuous reduction of easy to mine resources, resulting in the increase of underground resource mining, and many open-pit mines have been converted to underground mining. With this trend, the related research of underground engineering is gradually increasing, and the research content is more and more extensive, such as the stability analysis of the stope roof, the study of the movement and deformation law of the overburden, the role of groundwater, and the impact of key layers on underground engineering [1]. The movement of the stope roof and its overlying rock is a continuous and interactive process in time and space, but the deformation is nonlinear due to the different properties of the rock strata. The water conducting fracture zone formed by overburden failure is the flow channel of groundwater, and its development height is affected by the key layer.

With the increase of underground mining, the frequent occurrence of mining accidents and the instability of the roof, great threats have been posed to personnel and

equipment, which not only affect normal production but also cause great losses. Therefore, the research and analysis of roof stability has gradually become an important research topic and direction of many experts and scholars at home and abroad [1]. With the development of science and technology, more and more new ideas and methods have been applied to the stability analysis of stope roof, such as damage mechanics, plastic mechanics, energy theory, FLAC and ANSYS simulation software, and many achievements have been made. The stability analysis of roof is to study the influence of internal and external factors on the stability of roof, including its own nature, existing unstable structures, etc.; external factors include the impact of blasting vibration and mine room Width, water content, etc. internal factors are uncontrollable and can only be reduced or avoided in the design and production process by controlling external factors. Therefore, the research on roof stability is of great significance, including guiding the early mining design, production safety in the production process, gob treatment after mining, etc. the whole production process involves stope stability. When the boundary conditions of roof are different, they can be regarded as different models, and their span values are also different [2]. The study on the span of different models is of great significance for the selection of mining schemes and the optimization of design parameters. Based on this, this paper studies the application of mathematical simulation method in phosphate ore dressing.

## 2 Related Work

### 2.1 Phosphate ore Dressing Method

At present, the dominant phosphate ore dressing methods at home and abroad include flotation, scrubbing desliming and roasting digestion. In recent years, the processes of chemical leaching, photoelectric beneficiation, magnetic separation, heavy medium beneficiation and gravity flotation combination of phosphate rock have been gradually paid attention. Flotation has always been considered as the most effective of all research methods. It is used to separate sedimentary ore and endogenic apatite with siliceous gangue. As early as 1982, people have confirmed the possibility of recovering PSO with fatty acid. The study of apatite flotation also includes the study of flotation process, flotation equipment and flotation reagents.

#### (1) Direct flotation process

This technology is suitable for embedding silicon calcium phosphate rock with very fine particle size. The phosphate rock is ground once until the monomer dissociates, effective gangue mineral inhibitor is added, and then effective collector is added to float the phosphate mineral. The process has the advantages of simple flow and high impurity separation efficiency [3]. However, the floatability of carbonate and colophonite is close. Moreover, the poor selectivity and separation of fatty acids lead to the low grade of the products, and the pulp often needs to be heated, resulting in high production costs. Moreover, the direct flotation process is only applicable to ores with low MgO content.



(2) Reverse flotation process

This technology is applicable to phosphate ores with high carbonate and low silicon content, and is used to inhibit phosphate minerals in weak acidic medium. Carbonate minerals are floated with highly selective collectors. The process flow is simple and the carbonate separation efficiency is high. The reverse flotation technology has been successfully studied in the laboratory and applied to Wengfu Phosphate Mine in Guizhou Province. However, it can not separate siliceous gangue. In addition, it also has the disadvantages of fine particle size of foam products, difficult to transport and handle, and it is also necessary to find excellent collectors and inhibitors [4].

(3) Forward reverse, reverse reverse and reverse positive flotation processes

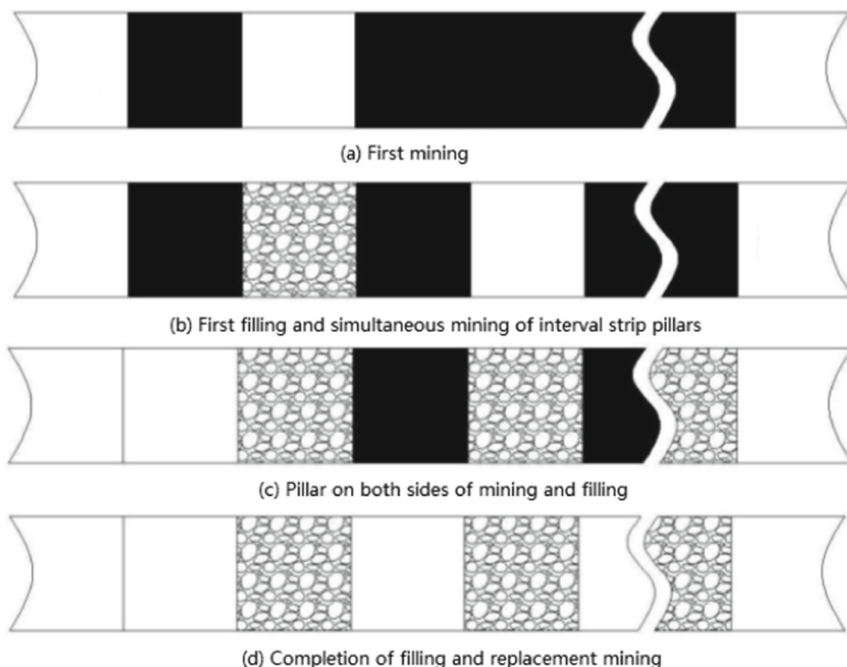
In order to overcome the disadvantages of direct flotation and reverse flotation, the process of direct reverse, reverse reverse and reverse positive flotation can be used to treat silicon calcium (calcium silicon) phosphate ore, which is essentially an organic combination of “carbonate flotation” and “silicate flotation”, i.e., “two-step flotation” is used to remove carbonate and silicate impurities in phosphate rock. According to the nature of the ore to be treated, the process of “reverse positive” flotation “Positive reverse” flotation or “double reverse” flotation technology. The process is characterized by strong adaptability to ore properties.

## 2.2 Digital Analog Method

In today's GIS systems, grid and tin data formats are basically supported, with tin as the main format and grid as the auxiliary. Nowadays, many algorithms of tin have been mature. Among the tin generation algorithms, there are three generally accepted and adopted methods, namely, divide and conquer method, data point successive insertion method and triangulation network growth method.

In digital terrain modeling, tin approximates the terrain surface by continuous triangulation generated from irregular data points. In terms of expressing terrain information, the advantage of tin model is that it can describe the terrain surface with different levels of resolution. Compared with grid model, tin model can express complex surface more accurately in less space and time at a specific resolution. Especially when the terrain contains a large number of features such as fault lines and tectonic lines, the tin model can better take these features into account and can more accurately express the surface morphology [5].

Interpolation of spatial data can be described as deriving arbitrary point or partition data through finite known point or partition data, so as to reconstruct a continuous feature change on the plane or in 3D space. Its objectives can be summarized as follows: ① missing value estimation: estimating the missing observation data at a certain point to improve the data density; ② Interpolation isoline: display the spatial distribution of data intuitively in the form of isoline; ③ Data grid: interpolation of irregular spatial data into regular spatial data sets, such as regular rectangular grids. The purpose of interpolation in this paper is to grid the data. The following Fig. 1 shows the phosphate rock simulation hierarchy.



**Fig. 1.** Phosphate rock simulation hierarchy

When the geological 3D model is established, because the sampling rate of geological 3D data is very low, there are only 116 sampling points in this paper. It is very difficult to fully and accurately express the real situation of geological phenomena according to these limited 3D sampling data. Therefore, to form a reasonable three-dimensional model using discrete spatial data, data interpolation must be carried out first, and then visualization and other operations of data can be carried out. Therefore, the interpolation of spatial data is a very important link in the process of 3D simulation and visualization. However, geological phenomena are ever-changing, and various interpolation methods have their own advantages and disadvantages [6]. Therefore, the choice of interpolation method is also a problem worth studying.

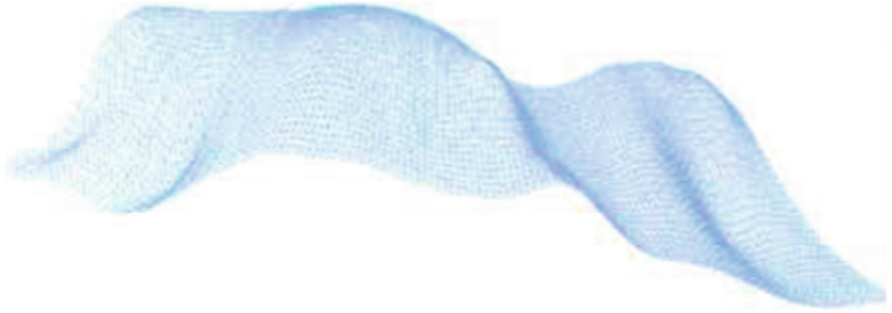
### 3 Application of Mathematical Simulation Method in Phosphate Ore Dressing

Geological data has its special characteristics. In spatial data interpolation, we can not simply apply the existing automatic interpolation method. We must consider many constraints and related geological principles. First of all, different interpolation methods have their own advantages, while different geological phenomena have different characteristics. It is necessary to select appropriate methods to simulate in order to form accurate and reliable models. Therefore, the key of spatial interpolation is to select a suitable interpolation method which is suitable for the spatial distribution of data through comparison.

For many spatial interpolation methods, there is no absolute optimal spatial interpolation method, only the optimal method under specific conditions. Therefore, it is necessary to select the optimal spatial interpolation method according to the internal characteristics of the data, based on the spatial exploration and analysis of the data, and through repeated experiments [7]. At the same time, the interpolation results should be strictly checked. At the same time, these special conditions and requirements must be taken into account in the processing and interpolation of geological data, and conventional automatic interpolation methods cannot be directly applied. It is necessary for geological data interpolation to interactively add geological principles and relevant constraints. Although the data distribution is often very uneven and the number is limited, there are still some connections and regularity between them. In data processing and interpolation, the distribution of data, the relationship between data and geological laws should be considered and utilized to obtain reasonable results. In many cases, the strata to be interpolated are not only one layer, but multiple layers that are related to each other, and these layers are very similar to each other in morphology. This similarity can be used in the interpolation process. For example, due to the influence of some factors, there may be few data at a certain level, and it is often difficult to form a correct interpretation only by interpolating these data. If there are many data in adjacent layers, the difference between two planes can be calculated at the position where there are data in both layers, and these differences can be extrapolated to correct the unreasonable layer shape. This can make up for the interpolation distortion caused by the scarcity of data in some layers and form a true and reliable layer shape. When there is little data, it is difficult to form reasonable results. It is allowed to add some control points automatically or manually based on a reliable geological interpretation or principle. The automatic method is suitable for those projects with increasing new data and slow manual addition: Although manual addition is time-consuming, it is often used to control those small and important geological features in areas with scarce data, which do not have enough data to establish grids or conduct automatic processing [8]. When the data points are extremely scattered, the linear interpolation method cannot accurately calculate the maximum and minimum values in the region. In this case, some high-order interpolation methods, such as Kriging method, can obtain good results.

The grid data is obtained above, and the geological layer will be drawn using Visual C++ and OpenGL. Here, the method of generating triangular network from regular grid data is used to reconstruct the geological layer. When the program is implemented, the data structure used is triangle, which is to divide the grid into two triangles and use small triangles to approach the three-dimensional geological layer. The results are shown in Figs. 2:

The three-dimensional visualization of geological information refers to the establishment of a mathematical model of geological features with appropriate data structure, and the use of computer graphics technology to express the mathematical description in the form of three-dimensional realistic images. 3D visualization technology is very important for the study of geological structure [9]. The three-dimensional visualization model can vividly express the “real” morphological characteristics of geological structure and the spatial relationship of structural elements. Combined with the information



**Fig. 2.** Three dimensional visualization of geological layer of phosphate rock

processing and spatial analysis functions of three-dimensional GIS, it can make the geological structure analysis more intuitive and accurate, and provide a realistic way for the quantitative development of geological structure research.

## 4 Conclusion

The application of mathematical simulation method in phosphate ore dressing is to simulate the behavior of biological system under different conditions. Mathematical modeling can be used to predict the behavior of biological systems when the environment or internal state changes. The most common applications of mathematical models are ecology and evolution, which are used to predict population growth, distribution and extinction patterns. In other fields such as engineering and medicine, they are often used to simulate the effects of various factors on the system (such as fluid flow). The mathematical simulation method can also be applied to problems involving chemical reactions.

**Acknowledgements.** Research and development of key technologies for resource utilization of flotation phosphorus tailings (Key R & D projects of Yunnan Provincial Department of science and technology: 202103AA080007) .

## References

1. Gao, K., Hu, D., Wang, S., et al.: Application of cerium phosphate in preparing anti-ultraviolet PET fibers with masterbatch method. *J. Polym. Res.* **27**, 1–12 (2020)
2. Hu, R., Zhou, X., Yu, Y.: Application Research on Real-Time Interactive Simulation System of Container Yard Based on Unity3D **10**(3), 11 (2022)
3. Huang, Z.: Application of an estimation method in the lure system. *Math. Probl. Eng.* **2020**, 1–7 (2020)
4. Rabe, M., Bilan, Y., Widera, K., et al.: Application of the linear programming method in the construction of a mathematical model of optimization distributed energy. *Energies* **15**(5), 1872 (2022)
5. Xue, L.: *Application of Mathematical Methods in Computer Algorithms Under the Background of Internet*. Francis Academic Press, (1) (2021)

6. Shmigirilov, R.V., Ryabinin, A.N.: Physical and mathematical simulation of bridge segment oscillations. *J. Phys: Conf. Ser.* **1959**(1), 012044 (2021)
7. Shurina, E.P., Itkina, N.B., Kutishcheva, A.Y., et al.: Mathematical simulation of coupled elastic deformation and fluid dynamics in heterogeneous media. In: *International Conference on High-Performance Computing Systems and Technologies in Scientific Research, Automation of Control and Production*. Springer, Cham (2022). [https://doi.org/10.1007/978-3-030-94141-3\\_11](https://doi.org/10.1007/978-3-030-94141-3_11)
8. Kunderák, J., Morgan, M., Mitsyk, A.V., et al.: Mathematical simulation of the vibration treatment of parts in a liquefied abrasive working medium. *Int. J. Adv. Manuf. Technol.* **120**(7/8), 53775398 (2022)
9. Application of Mathematical Method in Population Growth Model under New Situation. *Statistics and Appl.* **09**(4), 499–505 (2020)



# Application of New Light and High Strength Materials in Prefabricated Buildings

Jia Liu<sup>1</sup>, Mingfei Huang<sup>1</sup>(✉), Shanshan Li<sup>1</sup>, and Jinfeng Chen<sup>2</sup>

<sup>1</sup> Chongqing College of Architecture and Technology, Chongqing 400000, China  
570276792@qq.com

<sup>2</sup> Army Logistics Academy, Chongqing 516562, China

**Abstract.** The application of new lightweight high-strength materials in prefabricated buildings refers to the application of new lightweight high-strength materials such as aluminum, steel or other metal structures in building construction. These materials are used because they are lighter than traditional building materials and can help reduce the weight of the foundation structure by up to 20%. The main advantage of applying new lightweight high-strength materials in prefabricated buildings is that compared with traditional building methods, it will reduce energy consumption by up to 50%, which means less carbon dioxide emissions. It also reduces construction costs by reducing material costs, waste disposal problems and improving productivity. The use of these lightweight components also helps to reduce the energy required to build and maintain buildings. This will ultimately save money for owners, tenants and contractors.

**Keywords:** New light and high strength materials · Prefabricated building

## 1 Introduction

The concept of building industrialization in China was also put forward in the 1950s. Due to the backward economy and technology at that time and various reasons, the building industrialization at this stage was forced to run aground. It has been more than 70 years since China first proposed building industrialization. As China's economy enters a period of high-quality development, the high energy consumption and high pollution exposed by the development of traditional industries have become prominent problems hindering industrial transformation and upgrading [1]. Compared with the needs of industrial transformation and foreign mature experience, some scholars and experts put forward the concept of "new building industrialization", represented by prefabricated buildings, Adopting standardized components for production, transportation and installation not only improves the construction quality, but also saves the cost, and conforms to the theme of resource and environmental protection of the development of the times. Among them, the word "new type" mainly refers to the collision and combination of building industrialization and information technology, as well as the new technologies, new materials, new equipment and new processes adopted in the construction process [2]. Therefore,

the development of building industrialization in the future needs more research and exploration.

If a construction enterprise wants to survive in the fierce market competition environment, it needs to strictly control the buildings from the aspects of quality, cost and construction period, so as to increase the popularity and reputation of the enterprise and increase the profit rate of the enterprise [3]. Therefore, it is urgent to manage the construction project from the aspects of quality, cost and construction period. Its main advantages are highlighted in the following aspects: the cost management of construction projects can have a great influence on improving the profits of enterprises. To achieve the cost objectives of enterprises, effective cost control measures need to be taken to have a certain impact on reducing the costs of enterprises [4]. The results of cost management of construction enterprises can grasp the management ability and management level of construction projects in the most convenient and rapid way, and more intuitively reflect the results of construction project management. Construction project quality management can play a certain role in improving the popularity of enterprises. The construction period management of the construction project has a favorable impact on increasing the reputation of the enterprise.

## **2 Related Work**

### **2.1 Research Status of Prefabricated Buildings**

At home, although the construction industry in China has developed rapidly since the reform and opening up, the prefabricated buildings are still in the primary stage of the development of the construction industry. With the advancement of modern economic industry and industrialization, the advantages of the prefabricated construction industry are the direction of the development of the present era. However, the development status of the construction industry in China is still in the traditional way of on-site construction. The advantages of the prefabricated construction project have not been shown in China and need to be paid attention to. Scholars have conducted extensive analysis and Research on the development of prefabricated buildings in the domestic construction industry [5].

Zhang Chengcheng and others explained the reasons for the high cost difference between prefabricated houses and cast-in-place houses in terms of technology and cost through cases, and pointed out that there are few prefabricated component manufacturers in China that need to be imported from abroad, which leads to high production costs of prefabricated components [6]. This requires strong support from the national economy and policies, reducing and controlling the cost of prefabricated buildings in some aspects, and mobilizing the enthusiasm of enterprises to develop prefabricated buildings in a large scale.

Mai Junming and Yang Bao analyzed the differences between the structural forms of prefabricated concrete buildings and traditional residential buildings, clarified the advantages of the sustainable development of prefabricated concrete buildings, and built green houses. The strategic goal of China's future sustainable development is to promote the prefabricated development of the construction industry.

Ye Haowen pointed out that China is now facing a critical period of strategic development of building industrialization. However, there are many shortcomings in the development process of new building industrialization, and he put forward several suggestions for these problems. As an important method to develop new building industrialization, many problems should be corrected and solved. Follow the pace of the times, seize strategic opportunities, create new development together, and improve the industrial level of new building industrialization [7].

Yu Longfei pointed out in the article that the restricted factors generated in the development of prefabricated buildings in China. For the development trend and demand of the current society, China should study and optimize the laws, formulate normative policies, design and construction technology, etc. the whole life cycle management of construction projects should promote the development of the whole industrial chain of prefabricated buildings in China by combining BIM Technology.

## 2.2 Research Status of Light and High Strength Materials

As one of the most widely used artificial building materials in the civil engineering field, concrete material has the inherent disadvantages of excessive self weight and prominent brittleness. In consideration of safety in engineering application, it is often designed as a large section and high reinforcement. Therefore, lightweight and high reinforcement is an important development direction of future engineering building materials. Reducing the dead weight and improving the mechanical properties of materials are not only conducive to improving the use efficiency of materials and reducing the additional stress of structures, but also can save the resources and energy of building materials [8]. At present, scholars at home and abroad have done a lot of research on lightweight and high-strength materials widely used in engineering, and applied them to a variety of engineering disease treatment.

### (1) Light weight high strength modified concrete

Lightweight high-strength concrete material is a branch of high-performance concrete, which can greatly reduce its own quality and reduce the structural load on the basis of ensuring the strength, durability and other properties. Lightweight high-strength concrete materials originated in the early 20th century. In the late 1960s, the United States built a shell square tower with a height of 218 m. After the 1980s, it was widely used in the load-bearing components of bridges, super high-rise buildings and long-span buildings [9]. After the 1990s, domestic lightweight high-strength lightweight aggregate came out. At the same time, the state issued the technical specification for lightweight aggregate concrete, Since then, China's lightweight high-strength concrete has entered a period of rapid development and has been widely used in the fields of architecture, bridges, aerospace, cruise ships, offshore platforms and so on.

### (2) Polystyrene foam (EPS) and light mixed soil (SLS)

Expanded polystyrene foam (EPS) is made of polystyrene particles, which has the advantages of ultra light weight, thermal insulation, good mechanical properties and easy construction. In 1972, the Norwegian Highway Research Institute applied EPS to embankment engineering for the first time to solve the problem of soft

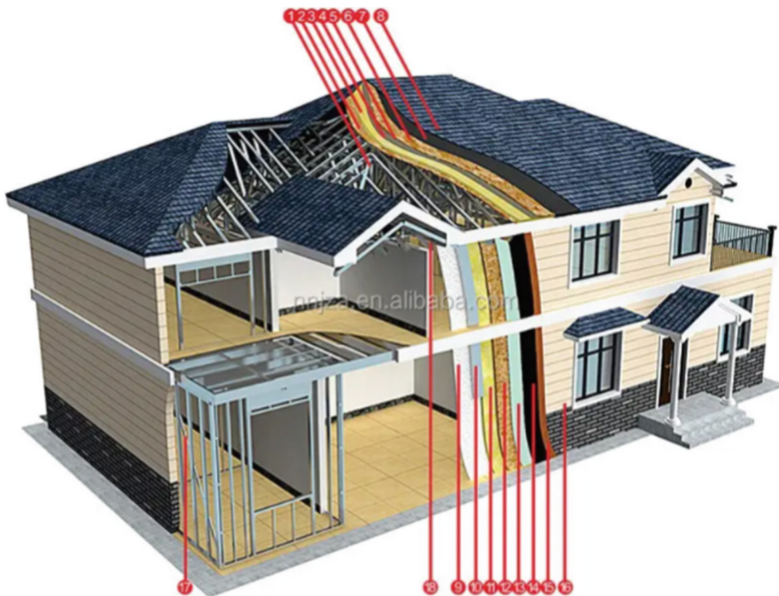


foundation, which attracted the attention of road workers around the world. After that, it was widely used in the Netherlands, France, Japan and other countries. The density grade of EPS is generally 200400 kg/m<sup>3</sup>, which is expensive, with low strength (80 kPa when the strain is 5%), high creep, easy to reduce the long-term service strength, and not widely used in engineering.

### 3 Relevant Contents of Prefabricated Building Project

#### 3.1 Concept of Prefabricated Building

Prefabricated building (also known as prefabricated construction, PC for short) is a building assembled on the site using prefabricated components, as shown in Fig. 1. During the implementation of prefabrication construction project, the traditional “construction” of the building has become a “manufacturing” project. During the construction process, the prefabricated components required by the building are produced in the enterprise workshop according to the production standardization, and then the prefabricated components of the primary structure are transported to the construction site. The prefabricated components are spliced, assembled and integrated by means of machinery and equipment through on-site hoisting, so as to form a building with use value. Since all prefabricated components are installed on site, there are also high standards for the production and assembly of prefabricated components.



**Fig. 1.** Prefabricated building

There are two clear indicators for prefabricated building engineering, namely, assembly rate and assembly rate.

- (1) Assembly rate refers to the speed of assembly of the volume used within the range of the volume ratio of prefabricated concrete of a single building. The assembly rate index reflects the industrial construction.
- (2) Assembly rate: the area of the total area where the rate requires the components to reach the proportion of the individual buildings of the building project.

### **3.2 Characteristics of Prefabricated Buildings**

Compared with traditional building engineering, prefabricated building engineering has the following characteristics due to its different planning, scheme design and construction methods:

#### (1) Advantages:

- 1) The primary structure and most of the prefabricated secondary structures in the construction project are produced by the factory, then transported to the construction site for assembly, and the factory production is realized in the workshop;
- 2) The production and prefabrication factory and data specification of materials shall meet strict requirements, precise production and high-quality effects to ensure the standardization level of prefabricated components;
- 3) The assembly operation at the construction site replaces the traditional cast-in-place operation in some aspects, improves the construction efficiency and the safety factor at the construction site, reduces the environmental noise pollution, and meets the assembly requirements;
- 4) Through on-site splicing and assembly, it is possible to achieve the goal of simultaneous civil engineering and decoration under the ideal state, integrated scheme design and construction methods, and implementation of integration.

#### (2) Disadvantages:

- 1) In the early stage of the development of prefabricated structures in China, the regulations and standards in various aspects are not comprehensive, and the production of prefabricated components meets the quality requirements;
- 2) Due to the complex site environment, the storage requirements of prefabricated components are unreasonable;
- 3) Prefabricated buildings have not been widely concerned, and the technology and construction level are not mature;
- 4) The product combination of prefabricated components requires the use of many large-scale mechanical equipment, which also has high requirements for the technical level of operators and the construction environment.

## **4 Application of New Lightweight and High-Strength Materials in Prefabricated Buildings**

### **4.1 Application of New Lightweight and High-Strength Materials in Block Buildings**

Block building can be divided into solid block and hollow block. In general, we use it to build multi-storey buildings. Because the whole wall is composed of block materials, we call it block building. This kind of block building construction material has a very simple production process, its cost is low, and it has good use flexibility. It can be widely used in the construction process of a variety of building materials, effectively promote the development of the process, and meet the rapid production and living needs of more people [10]. The application of new light-weight and high-strength materials to block buildings can ensure the use and definition of the wall. However, this material faces higher building requirements. In order to ensure that the wall can give full play to such functions, we need to strengthen and improve the strength of the new light-weight and high-strength materials, so as to better ensure the construction and building quality.

### **4.2 Application of New Lightweight and High-Strength Materials in Rising Slab and Rising Storey Buildings**

The so-called rising slab and rising storey buildings and modular buildings are important components of modern industrial prefabricated buildings. It mainly refers to the continuous and repeated pouring of the bottom layer of concrete. In this way, the corresponding concrete bottom layer can be effectively connected with the floor slab and roof slab of each layer, so as to effectively strengthen the corresponding design strength. In the process of production buildings, the relevant on-site construction personnel should be able to reasonably select the construction site according to the construction characteristics and construction requirements of the production buildings. In addition, we should use scientific methods to give a reasonable budget for the relevant construction speed according to the actual situation. By applying such lightweight and high-strength materials, the construction process can be greatly simplified and the actual construction time can be effectively shortened. Moreover, we should also fully consider the requirements of the actual characteristics of the equipment materials for the rising slab and the rising storey buildings, that is, we should be able to strengthen the demand for the construction strength and compressive strength of the materials, and this prefabricated building material also puts forward higher requirements for the construction and site environment. It has different pouring processes for the combination of external walls, brick walls and block walls, Sometimes external materials are needed to better meet the actual needs of the floor.

## **5 Conclusion**

The prefabricated building is an industrial building type that aims to speed up the construction of modern urban construction industry, draws lessons from the relative practices

of western developed countries, and further puts forward new requirements for current construction and credit managers on the basis of basic research. The in-depth analysis of the characteristics of prefabricated buildings and the manufacture and use of new building materials that can better match the prefabricated buildings will help to improve the quality standards of the prefabricated buildings in a series of industrial processes from design, production, sales and after-sales. The prefabricated building itself is an industrial basic field with very high requirements for technology integration. Therefore, the development and use of modern and mature building materials and technologies can be incorporated into the corresponding standards, which can better implement mandatory scientific and technological investment and quality improvement and better form a variety of equipment. It is the application of modern buildings in the actual urbanization construction.

## References

1. Geng, Z., Xin, M., Zhu, X., et al.: A new method for preparing photocatalytic cement-based materials and the investigation on properties and mechanism. *J. Building Eng.* **35**, 102080 (2021)
2. Shi, C., Zhang, Y.: The application of quality control circle to improve functional exercise execution rate of orthopaedic surgery patients: a SQUIRE-compliant quality-improving study. *Medicine* **100**(41), e27514 (2021)
3. Tay, L.T., Lee, Y.Y., Kueh, A., et al.: Flatwise and edgewise compression strengths of sandwich panel with silica aerogel mat. *IOP Conference Series: Materials Science and Eng.* **1101**(1), 012001 (2021)
4. Bahl, S., Singh, T., Kumar, V., et al.: A systematic review on recent progress in advanced joining techniques of the lightweight materials. *AIMS Materials Sci.* **8**(1), 62–81 (2021)
5. Lin, Y.: The Application of nucleic acids and nucleic acid materials in antimicrobial research. *Current Stem Cell Research & Therapy* (2021)
6. Wang, X.P., Guo, J.R., Liang, H.U.: Preparation and Application of Gallium-Based Conductive Materials in the Very Recent Years. (2021)
7. Chen, H.J., Lin, H.C., Tang, C.W.: Application of the taguchi method for optimizing the process parameters of producing controlled low-strength materials by using dimension stone sludge and lightweight aggregates. *Sustainability* **13** (2021)
8. Dang, H.M., Vo, C.T., Nguyen, V.D., et al.: A method for determining parameters of hyperelastic materials and its application in simulation of pneumatic soft actuator. *Int. J. Computational Materials Science and Eng.* (2021)
9. Chen, S., Xu, C., Ma, M., et al.: Application of solubility parameters in the preparation of PMMA with permanent antistatic, high toughness, and excellent optical properties. *Polymers for Advanced Technologies* (2021)
10. Sun, X.D., Yan, K., Tian, Z., et al.: New application of nano NaY zeolite in acrylic polymer pretanning agent. Microporous and mesoporous materials: The official journal of the International Zeolite Association (328-) (2021)



# Application of Virtual Simulation Technology in Microbial Image Segmentation in Sewage

Wenguang Xu<sup>(✉)</sup>

Wuhan Technology and Business University, Hubei 430065, China  
wenguangxu@aliyun.com

**Abstract.** The application of virtual simulation technology in microbial image segmentation is discussed. Microbial image segmentation is a process to determine the boundaries between different types of microorganisms. It has been widely used in medical imaging, biochemical analysis and food quality control. However, it is not applied to sewage treatment system because it has too many obstacles, such as high cost, low accuracy, complex data collection and treatment methods and so on. Virtual simulation is a method that can be used to create and simulate processes of interest by using computer software. In this paper, we propose an example of using virtual simulation to simulate bacterial behavior in sewage system. This method has been successfully applied to simulate water quality parameters such as turbidity, temperature and pH. The simulation results are then compared with the real data obtained from field measurements, which enables us to verify the model predictions of different samples collected from all over the world.

**Keywords:** Virtual simulation technology · Sewage treatment · Microbial image

## 1 Introduction

With the acceleration of China's urbanization process, the scale of the city is also expanding, facing a more and more severe situation of urban domestic sewage treatment. Data show that in the past decade, the average growth rate of urban domestic sewage discharge in China is 5.36%, and the total amount of sewage discharge ranks first in the world. However, the efficiency of sewage treatment lags behind that of other countries, which seriously affects the construction of urban ecological environment and the quality of residents' living standards. At present, the treatment of urban domestic sewage mainly adopts activated sludge process. It is a secondary biological treatment process most widely used in the world. It has the advantages of high treatment capacity and good effluent quality. The activated sludge treatment system is mainly composed of aeration tank, sedimentation tank, sludge return system and excess sludge removal system [1]. The first sewage treatment process is to use sewage treatment equipment to convert various substances that affect the environment contained in sewage into non-toxic and harmless substances, and then discharge them into the environment. Urban sewage is all the drainage collected through the sewer. It is the mixed water of various domestic

sewage, industrial wastewater and urban rainfall runoff discharged into the sewer system. Urban sewage includes domestic sewage, industrial wastewater and rainfall runoff.

Domestic sewage is the water discharged from people's daily life. It is water discharged from households, public facilities and living facilities such as kitchens, bathrooms, bathrooms and laundries in factories. The water quality of this kind of sewage is characterized by high organic matter, such as starch, protein, oil, nitrogen, phosphorus and other inorganic matter. In addition, it also contains pathogenic microorganisms and more suspended solids. The quality of domestic sewage is generally stable and the concentration is low.

This paper studies the application of virtual simulation technology in microbial image segmentation in sewage. The virtual simulation software for sewage treatment process is developed for the learning of water treatment in environmental specialty. The software is based on the whole process of sewage treatment process. The picture design is exquisite and real, which highly restores the real scene of sewage treatment plant, which is dynamic and realistic, and brings different experience effects for living learning.

## 2 Related Work

### 2.1 Virtual Reality Technology

At present, many models and improved algorithms have been proposed by scholars at home and abroad for microbial image processing in different scenes. However, because the morphology of sewage microorganisms is complex and changeable, and microbial images have different types, such as low contrast between foreground and background, blurred edge contour and single or multiple targets in a single image. This paper studies the application of virtual simulation technology in microbial image segmentation.

#### (1) Virtual reality technology

Computer simulation is used to generate a virtual world in three-dimensional space, build a virtual processing environment and processing objects of high microbial image simulation, and provide users with simulation of visual, auditory, tactile and other senses, so that users can rotate 360° in time and observe things in three-dimensional space without limitation, with friendly interface, interactive operation and lively form [2].

#### (2) The content of autonomous learning is rich

Explanation of knowledge points, including introduction of basic knowledge, explanation of software process, precautions during operation and video of process explanation;

#### (3) Process flow display

The real production process of sewage treatment process is displayed in 3D, and the comprehensive microbial image simulation simulates sewage treatment and supporting process principles. Students can roam in the microbial image simulation plant environment, learn the plant design requirements of sewage treatment process, and realize 360° rotation in the environment. Make the process flow into multiple processes, realize the function of plant microbial image simulation, display the process knowledge points in text form, and learn the theory of the process flow.

#### (4) Display of equipment working principle

Simulate the working principle and actual operation process of each equipment in sewage treatment in the form of text, animation and special effects, and intuitively show its internal and external form, mechanism and working principle.

## 2.2 Microbial Image Segmentation

The sewage treatment system requires high real-time performance. If the pixels in the input image are processed one by one, it will greatly increase the running time of the algorithm and affect the adjustment of process parameters and the stable operation of the system; Microbial images generally have the problems of unclear edge contour and low contrast between foreground and background. Therefore, it is difficult to achieve satisfactory segmentation effect by using classical segmentation algorithm. How to solve the problem of target extraction in the situation of low contrast and blurred edge contour is a difficulty; The gray distribution of some microbial images is uneven, and a single image contains multiple microorganisms of the same kind with rich morphology. Some multi-target microbial images also have the problems of blurred edge contour and low contrast of foreground and background, so it is difficult to extract targets quickly and accurately to meet the requirements of practical application [3]. For this kind of microbial image, it is necessary to re study the image segmentation algorithm for multi-target microbial prospect; The noise in the microbial image will have a great impact on the preprocessing process and the subsequent segmentation quality. However, in the process of removing the noise from the image, the linear filter will also smooth the important edge information. However, the nonlinear filter processes the image through the local structure of the image, which can not only remove the noise, but also highlight the edge characteristics of the image. Classical image segmentation algorithms use the texture information or edge information of the image to extract the region of interest. This kind of algorithm has a good segmentation effect for the image with clear edge contour, simple texture structure and high contrast between foreground and background.

$$\Delta w(i, y) = -\eta \frac{\partial e}{\partial w(i, y)} \quad (1)$$

$$e(w, b) = \frac{1}{2}(t - a)^2 = \frac{1}{2}(t - wp)^2 \quad (2)$$

However, some microbial images have edge blur and other problems, and the classical algorithm is difficult to achieve good segmentation effect, which seriously affects the subsequent recognition process. Grabcut algorithm is based on graph theory optimization and successfully combines texture and edge information. It is applied to this kind of microbial image and achieves better segmentation effect than the classical algorithm [4]. However, because the algorithm needs to manually frame the target area, there is still incomplete segmentation or over segmentation of the background area into the foreground. The increase of the number of iterations caused by the inaccurate frame selection position will also prolong the running time of the algorithm [5].

Therefore, virtual simulation technology can accurately extract different types of sewage microbial images in image segmentation technology, which is still the research focus in the field of microbial image processing.

### 3 Application of Virtual Simulation Technology in Microbial Image Segmentation in Sewage

Virtual simulation software uses dynamic mathematical model to simulate real experimental phenomena and processes in real time, and generates experimental phenomena and results consistent with real experiments through interactive operation of 3D simulation experimental device. Each student can do the experiment in person, observe the experimental phenomenon, record the experimental data, and achieve the purpose of verifying the formula and principle. It can reflect the basic experimental processes such as experimental steps and data sorting, meet the requirements of process operation and process operation training, and can operate safely and for a long period of time. In view of the uneven gray distribution of some microbial images, the existence of multiple identical microorganisms with clear edge contour and high contrast between foreground and background, and the difficulty of grabcut algorithm in multi-target image segmentation, a multi threshold segmentation method based on improved firefly algorithm is proposed in this chapter. Firstly, the optimal number of multi thresholds  $m$  is automatically obtained by counting the peak value of image gray histogram [6]; Secondly, based on the principle of two-dimensional entropy threshold segmentation, the two-dimensional entropy single threshold is extended to multi threshold, and the multi threshold objective cost function based on logarithmic entropy is designed; Finally, aiming at the problem that the traditional firefly intelligent optimization algorithm is easy to fall into the local optimal solution prematurely and the algorithm efficiency is low due to the lack of cooperation between fireflies, the optimization of firefly initialization process and the adjustment of algorithm variable parameters (step size quantization factor) are proposed  $\alpha$  And relative attraction parameters  $\beta$  j) The improved firefly algorithm can quickly and accurately find multiple optimal thresholds [7]. The algorithm is compared with one-dimensional, two-dimensional, Otsu, particle swarm optimization multi threshold segmentation and original firefly multi threshold segmentation. The particle segmentation of microbial image is shown in Fig. 1 below.

Considering the complexity of microbial image, because there may be noise in the microscopic imaging process of sewage treatment system, bilateral nonlinear filtering is used to eliminate the noise first; Then, the principle of SLIC algorithm and its application to the problem of low fit of super-pixel blocks at some fuzzy edges of microbial images that need preprocessing are carefully analyzed; Finally, the color information is introduced into the post-processing process of SLIC algorithm to improve the merging process of isolated pixels and small super-pixel blocks at the edge to enhance the connectivity [8]. At the same time, the gradient information is added to the similarity distance measurement to enhance the edge fit of super-pixel blocks.



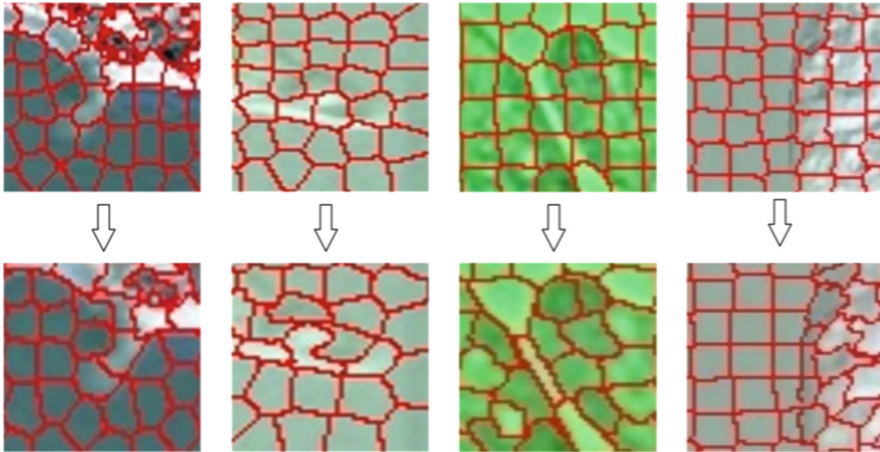


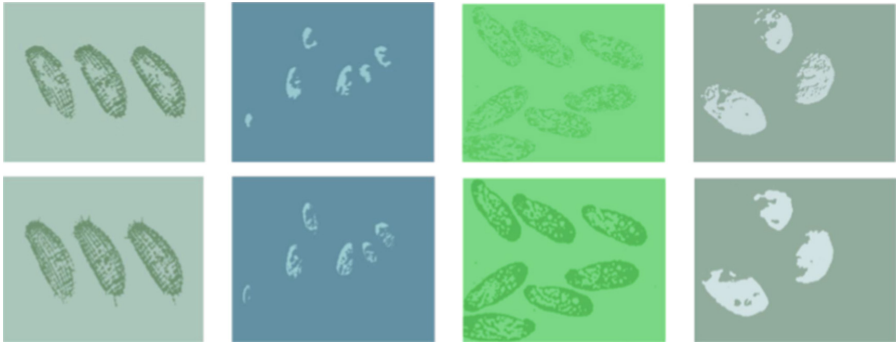
Fig. 1. Particle segmentation of microbial image

### 4 Simulation Analysis

The multi-target microbial image in the sewage microbial map mainly includes two types: multiple microorganisms of the same kind and multiple microorganisms of different kinds in a single image. The focus of this chapter and Sect. 5 is the image segmentation algorithm containing multiple microorganisms of the same kind. Multi threshold algorithm can achieve good segmentation effect, but the traditional maximum entropy multi threshold segmentation algorithm is often solved by exhaustive method, and the time complexity of the algorithm is high. In this context, the iterative optimization method combined with swarm intelligence algorithm is widely used in multi threshold segmentation [9].

The model is applied to real bacterial images with different contours and shapes, and the numerical experimental results are obtained At the same time, the original image is given. The active contour evolution and the approximate value of its piecewise constant (i.e.  $C$ , and  $C$ , average). In the experiment, the exact value of  $u$  and the initial level set equation will be given every time. As shown in Fig. 2 below, real bacterial images with different contours and shapes are shown.

However, the model still has many limitations. For the image with high noise, the segmentation effect is not ideal, and there are many wrong segmentation results, and some bacterial image segmentation is imperfect and insufficient [10].



**Fig. 2.** Real bacterial images with different contours and shapes

## 5 Conclusion

The application of virtual simulation technology in microbial image segmentation has been widely used to improve the accuracy of classification and distinguish different types of bacteria. This paper presents a new method for detecting and classifying microorganisms in sewage by using virtual simulation technology. The results show that the proposed method can accurately classify the target microorganisms and achieve high-quality classification results with low computational cost.

## References

1. The Application of Virtual Simulation Technology in Experimental Psychology Teaching. *Creative Education Studies* **09**(3), 613–616 (2021)
2. Di, X., Lian, H.: The Application of Computer “Virtual Simulation” Experimental Teaching in Basic Football Tactics (2021)
3. Robin, F.: Beyond the classroom: insights into the use of virtual simulation in veterinary education. *The Veterinary record*, 2020年186卷17期:559–561页 (2020)
4. Lu, J., Peng, W., Lv, Y., et al.: Application of cell immobilization technology in microbial cocultivation systems for biochemicals production[J]. *Indust. Eng. Chem. Res.* (2020)
5. Hanel, E., Bilic, M., Hassall, K., et al.: Virtual application of in situ simulation during a pandemic. *Canadian J. Emerge. Med.* **22**(5), 1–6 (2020)
6. Wen-Xiu, J.I., Wang, M.J., Dong, W.W.: Exploring the application of virtual simulation technology in engineering experiments. *DEStech Transactions on Social Science Education and Human Science* (2020)(icesd)
7. Jiang, H., Wei-Xing, H.U.: Development and application of virtual simulation and real arm combination system based on teleoperation Technology. In: 2021 2nd International Conference on Big Data and Informatization Education (ICBDIE) (2021)
8. Xiao, Q., Wang, L., Fang, J., et al.: Research on the application of virtual instrument technology in simulation training. *MATEC Web Conf.* **309**, 03022 (2020)
9. Wu, Y.: Analysis of the application countermeasure of virtual reality technology in the vision simulation system. *J. Phys: Conf. Ser.* **1550**, 032099 (2020)
10. Liu, W., Liang, R.: The application of virtual simulation software in university physics teaching. In: 3rd International Conference on Advances in Management Science and Engineering (IC-AMSE 2020) (2020)



# Application of Virtual Simulation Technology in Traditional Handicraft Protection

Xiong Huang<sup>(✉)</sup> and Yuanjing Zhang

Hubei University of Technology, Wuhan 430000, Hubei, China  
moonzz2022@163.com

**Abstract.** The protection of traditional handicraft industry is an important part of national industry. Handicrafts are handmade and require a lot of labor and time. In addition to the high cost, it also has some problems, such as: 1) non-standard handicrafts; 2) Handicrafts cannot be mass produced; 3) Handicrafts cannot meet the requirements of international trade. The protection of traditional handicrafts is mainly based on the inspection and classification of rival handicrafts. The detection rate of fake and inferior products in this field is very low, which is difficult to effectively protect the national economy. Therefore, it is necessary to use virtual simulation technology for product quality control. It can be used as a new tool to detect counterfeit products at any time, especially when there is no qualified inspector or there is a shortage of inspectors. It also provides an effective means to prevent counterfeit products from entering the circulation and selling in retail stores without being discovered by inspectors. Therefore, this paper studies the protection of traditional handicrafts based on virtual simulation technology.

**Keywords:** Virtual simulation technology · Traditional handicraft

## 1 Introduction

China has a vast territory, abundant resources and diverse ecological environment, which is suitable for fishing, hunting and gathering, as well as animal husbandry and farming. Since ancient times, all ethnic groups have multiplied here, creating a brilliant ancient civilization and leaving a solid cultural heritage.

In the early Paleolithic age, ancestors lived in this northeast land. The human bone fossils found at Jinniushan site in Yingkou City, Liaoning Province, are 230000 to 300000 years old [1]. The human bone fossils found at the miaohoushan site in Benxi are 140000–240000 years old. At this time, the ancestors already knew how to make stone tools and bone vessels with different uses. The polished horns, bone vessels and decorations of the late Paleolithic age unearthed from the Xiaogushan site in Haicheng are more skillfully made. In the Neolithic age, the ancestors in Northeast China had mastered relatively advanced handicraft technology and produced exquisite handicrafts such as jade, bone ornaments, wood reliefs, coal products and pottery sculptures [2].

It can be seen that traditional arts and crafts and handicrafts have a glorious source and a long history, and the innovation, dissemination and exchange of culture and skills are

inevitable. In the long history, the traditional handicrafts in Northeast China have been constantly innovated; However, no matter how the handicrafts in history have changed, the art forms and handicrafts we inherit and own today are all derived from tradition.

Because of the rapid economic development and the increasing popularity of monetization, the trend of money worship in contemporary society is becoming more and more serious. Only those with high economic strength will focus on traditional handicrafts, and because there are many art works in the works of traditional handicrafts, it greatly meets their needs. Looking at the young people now, they have too many ideas, are eager to pursue fame and wealth, and are unwilling to learn the inheritance technology accumulated over time. However, the process of making handicrafts is not difficult, but it takes time to accumulate and become more and more skilled [3]. Although the process is troublesome, you have learned real kung fu. However, because it takes too long, young people are unwilling to work hard, resulting in the slow development of traditional handicrafts. Up to now, the current situation of traditional handicrafts in China is not optimistic. With the continuous progress of science and technology, it has changed from pure handcraft to assembly line processing, and from unique to unified style, which has reduced the price of handicrafts too much [4]. In addition, only a small part of the so-called traditional skills have been completely inherited, and a large part is facing the loss of inheritance. But in my opinion, no matter how to change, the basic form cannot change. Handicraft is handicraft, just like some high-end industries, only do high-end, because its positioning is like this.

## 2 Related Work

### 2.1 Traditional Painting Skills

Traditional folk painting skills refer to the traditional handicrafts that have been created in the folk in history and passed down to this day in a living form, and that use traditional methods to draw or print folk paintings. It mainly includes: traditional folk hand painting and drawing skills, folk printmaking and folk freehand painting skills (such as pyrography, finger painting, gourd painting, new year painting, birch bark painting, etc.).

As for New Year pictures, we have to say that China's famous "Cloisonne crystal paintings".

Cloisonne originated in the Yuan Dynasty and has a development history of more than 600 years. It is one of the famous traditional arts and crafts varieties in China. It is called Cloisonne because it flourished in the Jingtai period of the Ming Dynasty and its glaze color is mainly blue.

During the reign of Jingtai in the Ming Dynasty, the art became mature and gradually became famous in the world, so it was named "Cloisonne". Cloisonne reached its peak in the Qing Dynasty. Cloisonne is unique in the field of Arts and crafts because of its special process materials, complicated process procedures, exquisite artistic modeling, elegant artistic taste, and various attributes such as shape, color, quality, and use [5]. There are many kinds of traditional cloisonne products, mainly including utensils, stationery, furniture, antique, lamps, jewelry, etc. Cloisonne products are highly decorative and practical. However, with the change of lifestyle and aesthetic habits, Cloisonne is

facing the subject of change and keeping up with the trend of appreciation. The workers of cloisonne craft constantly explored and discussed the innovation of cloisonne, and created a new form of cloisonne arts and crafts - cloisonne pojing painting.

## 2.2 Traditional Knitting Skills

Traditional weaving skills are hand-made weaving skills created and inherited by our ancestors. Broadly speaking, it can be divided into traditional silk weaving skills, traditional cotton weaving skills, traditional brocade weaving skills, etc. What we are talking about here is the traditional weaving technique in a narrow sense, which uses natural plant materials such as bamboo, wood, wool, grass and rattan to compile folk handicrafts and daily necessities.

Jilin Tongzhi: in the Qing Dynasty, the North zhuchengzi area of Changchun Prefecture was rich in sedge, trigonous grass, cattail, camphor, willow and shrubs, which provided rich raw materials for the prevalence of straw weaving, willow weaving and other weaving skills [6]. The weaving skills in this area have a long history. With the change of traditional lifestyle, most of the traditional weaving skills have disappeared. Now, corn husks are the main weaving materials, and the main products are daily necessities.

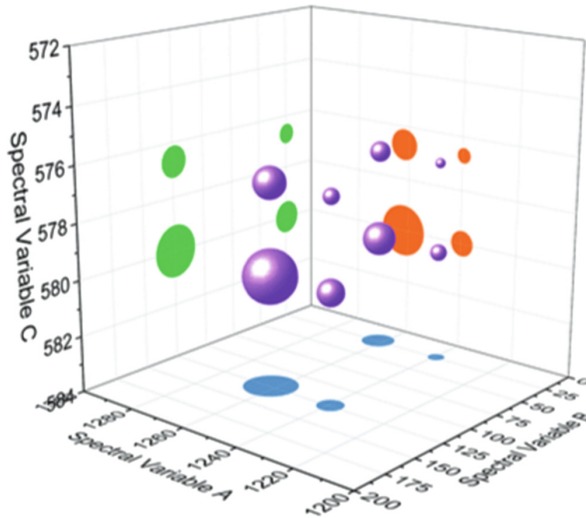
Among the traditional weaving skills, the more common ones are “bamboo weaving” and “straw weaving”.

According to the existing archaeological data, after humans gradually began to live a stable life, they began to study the production of agriculture and animal husbandry, and some rice harvested and food obtained from hunting also became abundant. In order to deal with emergencies, they will store water and food. Because they need to store food, so-called utensils exist. They use the most primitive stone axes, knives, etc. to collect plant branches, and make these branches into baskets, baskets and other storage supplies [7]. Bamboo, in particular, was the main material for making utensils at that time.

## 3 Digital Production: From “Virtual” to “Reality”

The media shift has updated the realization form of process production. The introduction of digital design and manufacturing technology into the field of handicrafts has led to great changes in the technological process of works. It can not only use computers to easily and quickly carry out creative transformation and artistic communication in digital (bit) form, but also show the final effect of work molding and generate creative prototypes, realize the “creative substitution and replacement of material media”, and change the design thinking and production logic of traditional handicrafts, A new production mode, digital production, has been derived, that is, digital design and digital manufacturing [8]. But it is precisely the changes in these two stages that have led to the revolutionary transformation of handicraft art design and manufacturing in contemporary times. Scholar zhaojianghong proposed; “Design in the 21st century will change from tangible design to intangible design, from material design to non-material design, and from physical product design to virtual product design”. “The process of digital design is also the process of creative digitization, and the existence form of its constructed artistic image is a kind of “tangible virtual”. On the one hand, CAD (Computer

Aided Design), rhinocero (rhinoceros) As well as the digital design software represented by ZBrush, contemporary process design has changed from traditional manual drawing to computer two-dimensional drawing, and even transformed its conceptual prototype into three-dimensional digital model. Nowadays, artists have realized the transformation from material two-dimensional drawing design to virtual three-dimensional visual image [9]. Through the real-time modification of the conceptual model and the real-time feedback of its artistic image, creators can more easily perceive and understand their creative prototype to improve their own design ideas, as shown in Fig. 1. On the other hand, relying on computer parametric design program, designers can freely create a variety of combinations of complex geometric forms, or more accurately imitate nature and absorb the beauty and geometric structure of nature.



**Fig. 1.** Two dimensional graph turns to virtual three-dimensional

The process innovation of digital handicrafts lies not only in the non-material turn of digital media in the process design stage, but also in the non-material expression in the manufacturing stage. “Digital technology can realize the possibility of coding, compression, transmission and conversion between all physical and material phenomena and codes”. Digitization is not only the digitization of design process, but also the digitization of material, manufacturing and processing process, that is, the digitization of manufacturing process. Digital manufacturing equipment represented by 3D printing, laser cutting, CNC milling machines, etc. can directly or indirectly act on materials such as paper, wood, cloth, resin or metal, control and adjust the relationship between raw materials and processing through computers and data information, connect abstract data and specific material materials and convert them into exquisite process products, which is conducive to simplifying the difficulty of process realization, It also helps to produce new technological forms that are different from those made by hand.

Digital media provides a bridge for craft creation from conceptual prototype to physical creation, and extends the creative space. Associate Professor Zeng Li of Shenzhen Vocational and technical college believes that digital technology may bring more possibilities for artistic creation. In the past, there was a continuous trial process from concept to work. In the digital age, the final three-dimensional effect of the work can be displayed through digital technologies such as 3D printing or virtual reality; From the perspective of artistic creation, digital technology extends the space of artistic creation and brings more possibilities; From the industrial level, digital technology reduces the process production costs, including labor costs, time costs and capital costs. "For example, natte metal art is the first decoration brand in China with metal as the base material. Its business mainly focuses on metal sculptures, metal murals and other handicrafts. In the production process of some of its craft works, the designer's creative design is transformed into three-dimensional manuscripts through 3D printing technology as the physical reference for creation, and then it is modified and improved based on three-dimensional manuscripts before it is put into production, This process greatly shortens the production cycle of handicrafts from concept to prototype, and saves the material cost of drawing, mold opening, molding and other links from design to manufacturing.

#### **4 Application of Virtual Simulation Technology in Traditional Handicraft Protection**

Virtual reality (VR) is a comprehensive information technology developed from computer graphics technology, simulation technology, multimedia technology, artificial intelligence technology, computer network technology, parallel processing technology, multi-sensor technology and other technologies. It can fully simulate the functions of human visual, auditory, tactile and other sensory organs, provide a real-time, three-dimensional virtual environment, enable people to immerse in the computer-generated virtual environment with the help of necessary equipment, and can interact with the objects in the virtual environment in real time through language, gestures, etc., to create a personalized multi-dimensional information space. Users can not only feel the realism of "immersive" through the virtual reality system, but also break through the space, time and other objective constraints, and feel the experience that cannot be experienced in the real world.

First, data acquisition[10]. The construction of virtual reality first requires the data base related to traditional handicrafts, including spatial data: such as the location, shape, outline, size, proportion and so on of the modeling object; Texture data: digital photos of different angles and elevations of modeling objects and texture data of different materials; Attribute data: text, picture, audio and video file data related to the attribute introduction of the modeling object.

The second is 3D modeling of virtual scene. After obtaining a large amount of basic data, 3D modeling is one of the key steps, and it is also the basis for the digital display of traditional handicrafts. In the virtual scene, the virtual object is the main body, and its virtual reproduction is realized by modeling. Virtual object is the focus of 3D modeling of virtual scene and the primary condition for users to experience 3D interactive immersion. At present, the commonly used methods include geometric modeling, image modeling,

and the modeling method of combining image and geometry. It mainly uses the existing mature modeling software to model the target interactively, and the modeling accuracy can meet the requirements of practical applications,

Thirdly, real-time 3D graphics generation technology. The 3D model needs further processing, because a realistic and smooth 3D virtual scene needs not only a realistic model, but also a smooth running effect. Real time generation of 3D graphics is to display the established 3D scene and target model in front of all kinds of users. The key is to achieve “real-time” generation of 3D images or virtual environment.

Finally, stereoscopic display and sensor technology. 3D virtual scene realizes interactive roaming and other functions through the application of various stereo display technologies and teleporter technologies.

With the continuous maturity of virtual reality technology and the continuous reduction of cost, its application in the field of traditional handicraft digital protection has attracted more and more attention. Undoubtedly, virtual reality technology is obviously an indispensable part of digital technology in terms of its own nature and function. It can also play a solid and effective role in the digital development of traditional handicrafts.

## 5 Conclusion

Today, with the rapid development of information, the participation of digital technology in the protection of traditional handicrafts is a new idea of sustainable development. Applying virtual reality technology to the protection and inheritance of traditional handicrafts has not only the theoretical basis of digital art, but also the practical feasibility. The digital protection scheme of traditional handicrafts based on virtual reality technology can maximize the sharing and utilization of resources, open up new living space for traditional handicrafts, let more people participate in the protection and inheritance of traditional handicrafts, and promote the prosperity of traditional handicrafts culture.

## References

1. Gu, Y., Huang, W.: Analysis of the application of virtual simulation software in food testing technology course with atomic absorption spectrum1– taking the example of determining calcium in drinking water by atomic absorption spectrometry. *IOP Conf. Ser. Earth Environ. Sci.* **680**(1), 012085 (6pp) (2021)
2. Xiao, X., Liu, X., Xiao, Z.: Construction and application of computer virtual simulation teaching platform for medical testing. *J. Phys. Conf. Ser.* **1915**(4), 042074 (2021)
3. Zhang, Z., Du, J., Zhang, W.: Application of virtual simulation experiment system in instrumental analysis teaching in colleges (2022)
4. Badowski, D., Rossler, K.L., Reiland, N.: Exploring student perceptions of virtual simulation versus traditional clinical and manikin-based simulation. *J. Profession. Nurs.* **37**(9) (2021)
5. Li, X., Chen, Y.: Virtual Simulation Technology for Innovative Design of Intelligent Products (2021)
6. Zhao, C.: Application of virtual reality and artificial intelligence technology in fitness clubs. *Math. Probl. Eng.* **2021**(20), 1–11 (2021)
7. Wang, X., Wang, J., Weng, Z., et al.: Reform exploration of proteus virtual simulation practice teaching in electronic. *Adv. Appl. Sociol.* **12**(4), 9 (2022)



8. Hyc, A., Chc, B., Cwl, C.: The effect of a virtual simulation-based educational application on nursing students' belief and self-efficacy in communicating with patients about complementary and alternative medicine. *Nurse Educ. Today* (2022)
9. Marinus, A.B.: Application of virtual anthropological techniques in the reconstruction and analysis of late Middle and Late Pleistocene hominin crania (2021)
10. Hai, N.T., Duong, N.T., Huy, D., et al.: Sustainable business solutions for traditional handicraft product in the northwestern provinces of Vietnam. *Management* **2021**(1) (2021)



# Big Data Visual Accounting Model for Ecotourism Carrying Capacity

Yang Ye<sup>(✉)</sup>

Department of Tourism and Convention, Hefei University, Hefei 230001, China  
yeyang1718@126.com

**Abstract.** With the rise of eco-tourism and the unprecedented prosperity of tourism, the increasing environmental awareness and the wide implementation of sustainable development strategy, eco-tourism has developed rapidly; But at the same time, various negative effects also appear, and gradually show their huge potential threats: over exploitation and even predatory development of tourism resources, extensive management of tourist attractions, pathological expansion of tourism facilities construction, sharp deterioration of environmental quality, etc., all of which threaten the sustainable development of tourism and ecology to varying degrees. The research on the big data visual accounting model of ecotourism carrying capacity is a research paper, focusing on the role of visualization in understanding and interpreting big data. The purpose of this study is to develop a web-based interactive tool that enables users to explore, visualize and analyze large amounts of data from different sources to improve the decision-making process. The main purpose of this study is to create a tool that allows users to interact with a large amount of information using simple tools such as charts, charts, maps and tables. These tools are designed to be used by both experts and non experts.

**Keywords:** ecotourism · Big data visualization · Bearing capacity accounting

## 1 Introduction

With the rapid development of the global economy and the increase of leisure time, tourism, especially eco-tourism, is growing rapidly all over the world. Since the first package tour organized by thomascook in 1841, tourism has achieved great success in various regions of the world. Especially since the Second World War, the world tourism has developed rapidly and reached an unprecedented stage of prosperity. People unilaterally believe that “tourism is a smoke-free industry”, so they blindly adopt an extensive development model that does not conform to the environmental moral code, and simplify the development of tourism into quantitative growth or extended reproduction [1]. The development of tourism resources often lacks in-depth investigation and research and comprehensive scientific demonstration, evaluation and planning, and pays more attention to development than protection, or even only development without protection, resulting in serious environmental pollution in the tourism area, The ecological system is out of balance, and the contradiction between tourism and ecological environment is

becoming increasingly prominent. Since the 1960s, people have gradually realized that the disorderly development of tourism has brought huge, even catastrophic damage to the environment [2]. If this problem is not fundamentally solved, the natural ecological environment, the basis of tourism development, will be weakened.

With the idea of realizing sustainable tourism development becoming popular, the research on tourism environmental carrying capacity, as an important foundation and carrier of realizing sustainable tourism development, has become one of the hot spots in the field of environmental research at home and abroad in recent years. Tourism environmental carrying capacity, also known as tourism environmental capacity, is an important measure to measure whether the tourism environment and tourism development are coordinated [3]. It is the core theoretical issue of the contradiction between the development of tourism and the protection of the environment. In order to avoid the tourism industry repeating the mistake of “pollution before treatment” in industrial development, the tourism environmental carrying capacity, as the basis for judging whether tourism activities have a negative impact on the environment, has been put forward with the development of tourism and has become the focus of tourism research. The study of tourism environmental carrying capacity is an important part of the implementation of the national sustainable development strategy [4]. It is of great significance for the development of tourism resources, the adjustment of tourism structure and the protection of tourism environment. However, in the actual research work, this problem involves resources, ecology, regional economy, social environment, tourists and other aspects, so it is difficult to study.

Many limit values of tourism environmental carrying capacity are the threshold that can not be exceeded for the development of scenic resources and environmental protection in tourist areas, the guarantee for maintaining the balance of the ecosystem in tourist areas, and the important scientific basis for people to correctly deal with the relationship between tourism activities in scenic areas and ecological environmental protection; The most appropriate value for the calculation of tourism environmental carrying capacity is an indispensable basis for correctly determining the tourism development objectives, the tourism development scale, correctly planning the facility area, capacity and designing the length of tourism lines [5]. It is also an important basis for reducing blind investment and construction and overcoming the blindness of tourism planning.

## 2 Related Work

### 2.1 Bearing Capacity Theory

In 1921, Parker and Burgess put forward the concept of carrying capacity for the first time: the maximum limit of the number of individual organisms under a specific environmental condition (mainly the combination of living space, nutrients, sunlight and other ecological factors). It usually refers to the maximum number of an individual organism that can survive under certain environmental conditions. IRMI seidll and Clem a.tisdell explored the origin of the concept of carrying capacity, and believed that the concept of carrying capacity could be traced back to Malthus' population theory in the field of human ecology and biological ecology. In North America, South America and Asian grassland areas, due to grassland reclamation, overgrazing and other reasons, the land

began to deteriorate. In order to effectively manage the grassland and obtain the maximum economic benefits, some scholars introduced the carrying capacity theory into grassland management, so the concepts of grassland carrying capacity and maximum livestock carrying capacity were put forward. With the increasing global population and decreasing cultivated land, some scholars put forward the concept of land carrying capacity. In 1949, Allan of the United States defined land carrying capacity as “a region can permanently support the number of people and the level of human activities on the premise of maintaining a certain level without causing land degradation”. In the 1950s and 1970s, many foreign scholars discussed the calculation basis of land carrying capacity and believed that: land carrying capacity is the maximum population that the land production potential can accommodate on the premise of ensuring that it will not cause irreversible negative impact on land resources. Millington et al. [6]. Applied the multi-objective decision analysis method to calculate the land resource carrying capacity of Australia based on the restrictions of various resources on the population.

Later, many scholars conducted comprehensive research and Discussion on the global carrying capacity from different aspects such as economy, society, environment and development. Therefore, understanding the theory of carrying capacity and its calculation method is of great significance to the study of eco-tourism environmental carrying capacity.

The difference between the eco-tourism environmental carrying capacity and the traditional tourism environmental carrying capacity lies in: first, the premise for the two to carry out tourism activities is different. The premise for carrying out eco-tourism activities is to protect the ecological environment and not interfere with natural regions, while the premise for carrying out traditional tourism activities is to meet the tourism needs of tourism subjects. Second, the eco-tourism carrying capacity focuses on the lowest impact on the ecological environment, and even protect and improve the ecological protection of the tourism area, while the traditional tourism environmental carrying capacity focuses on not destroying the ecological balance, and the ecological protection goal is relatively low [7]. Third, the eco-tourism environmental carrying capacity aims to make the ecological environment and people in a given natural region be used sustainably, emphasizing the protection and improvement of the ecological environment so that it can get better development in the future. However, the traditional environmental carrying capacity only emphasizes that the tourism activities will not damage the balance of the ecosystem, and does not emphasize whether the ecological environment can be used continuously with the same utilization rate. Its protection requirements are low.

## **2.2 Tourism Environmental Capacity**

### **(1) Main concepts of domestic tourism environmental capacity**

The concept of tourism environmental capacity has been widely used in tourism, which provides a scientific method to guide the healthy development of tourism. However, the tourism environmental capacity is still limited. One is to take the number of tourists accepted as the only indicator, which is bound to violate the essential connotation of tourism environment as a carrier of spatial and non spatial, physical and non physical elements. The second is that environmental capacity is widely used in Environmental Science, and the scientific system is very perfect. Its foothold is the amount of

pollutants absorbed by the environment, and it is not very accurate to measure the number of tourists[8]. Therefore, in 1995, cui Fengjun and others tended to use the term “tourism environmental carrying capacity”. Subsequently, many domestic scholars began to discuss the concept of tourism environmental carrying capacity. There are two basic definitions:

- ① The intensity of tourism activities (including tourist density, land use intensity and economic development intensity) that a tourist destination can withstand in a certain period of time, provided that the existing state and structural combination of the environment of a tourist destination do not have harmful changes to contemporary and future people.
- ② In a certain period, a certain state or condition, the environment of the tourist area can bear the closed value of the economic volume of tourism activities.

### **3 Characteristics of Eco-Tourism Environmental Carrying Capacity**

#### **(1) Spatiotemporal differentiation**

In terms of time distribution, tourism activities are seasonal. As the natural ecological environment factors, socio-economic factors, local residents' psychological factors and management suitability of the tourism area change regularly with time, the spatial bearing capacity, economic bearing capacity and social psychological bearing capacity of resources determined by these impact factors also change with time. This finally determines the “sum” of the above bearing capacity components, that is, the eco-tourism environmental carrying capacity has the characteristics of time distribution [9]. From the perspective of spatial distribution, the natural ecological environment, as the basis of tourism environmental carrying capacity, has spatial differences in its ecological resilience, sensitivity and other characteristics, as well as regional differences in socio-economic environment and social psychology. Therefore, the ecological tourism environmental carrying capacity has the characteristics of spatial differentiation.

#### **(2) Static and dynamic characteristics**

The static characteristics are mainly determined by the necessary stay time for tourists to carry out a certain tourist activity.

In the process, the environmental carrying capacity will not change, that is, it has an instantaneous environmental carrying capacity. The dynamic characteristic is determined by the relationship between the finite residence time and the infinite opening time of the tourist area. This characteristic determines that the eco-tourism environmental carrying capacity can reach a considerable amount in a certain period of time.

#### **(3) Controllable factors and uncontrollable factors coexist**

According to the control ability of human activities on the factors affecting the environmental carrying capacity of ecotourism, the factors affecting the environmental carrying capacity of ecotourism can be divided into controllable factors and uncontrollable factors. Generally speaking, among the environmental carrying capacity of ecotourism, the spatial carrying capacity of resources and the social psychological carrying capacity are

rigid and elastic; However, the economic bearing capacity and environmental management bearing capacity are relatively large and easy to regulate. Therefore, the bottleneck of the growth of eco-tourism environmental carrying capacity is usually the resource space environmental carrying capacity and social psychological carrying capacity.

#### (4) Objectivity and measurability

In a certain period and a certain natural area, the structure, function and information of eco-tourism environmental system are relatively stable and objective. Ecotourism environmental carrying capacity is the reflection of the functional structure of ecotourism environmental system. It depends on this structural function. Under the condition that the eco-tourism environmental system does not change, under a certain evaluation index system of eco-tourism environmental carrying capacity of natural regions, the amount of regional eco-tourism environmental carrying capacity exists and can be calculated. In order to ensure that human tourism activities do not interfere with the local ecological environment system and protect it, we must grasp and calculate this objective quantity through certain means.

#### (5) Comprehensiveness and complexity

Tourism is a comprehensive industry. The eco-tourism environmental system formed by its development is a composite system composed of social, natural, economic and other elements. Whether qualitative description or quantitative analysis, it involves many factors. In particular, when quantifying the eco-tourism environmental carrying capacity, it requires interdisciplinary and multi field cross cutting and comprehensive research. This is the complexity of the study of eco-tourism environmental carrying capacity.

#### (6) Variability

At different times, the ecological environment and social environment of the ecotourism environmental system in a certain natural region are different, and they are always changing. The ecological environment changes with the change of seasons, and the social environment changes with the change of human activities. The change of social environment changes people's requirements for the ecological environment. This series of changes reflected in the eco-tourism environmental carrying capacity is its "quality" and "quantity" changes. The change of "quality" is the change of eco-tourism environmental carrying capacity index system, while the change of "quantity" is the change of eco-environmental carrying capacity evaluation index value.

## **4 Big Data Visual Eco-Tourism Carrying Capacity Accounting Model**

### **4.1 Idea of Index Selection**

(1) The index system can be divided into general and specific

For the tourism environment, the natural ecosystem, population, social and economic factors have distinct regional characteristics. For different regions, not only the types and importance of the determinants of the tourism environmental carrying capacity may be different, but also the specific evaluation objectives and management requirements will be different. Therefore, the tourism environmental carrying capacity is very local, and there is no "universal" index system. However, for most tourism environmental

systems, their constituent elements have certain similarities. It is both necessary and possible to establish some form of general index database. In other words, there are two index systems of tourism environmental carrying capacity: general index system and specific index system. The former is similar to the index library and has universality, which can provide a reference index set for specific applications in different regions; The latter is a specific index system applicable to specific regions, which can be deleted, added and modified according to specific needs on the basis of the former. In addition, the evaluation method, especially the weight, is closely related to the number and structure of indicators. For different regions, different tourism environment systems and the importance of various limiting factors will lead to different weights, which also requires that a fixed index system should not always be used. This also shows that the weight is conditionally dependent [10]. Once the application background is lost, the weight becomes meaningless. This idea is quite different from the traditional approach that always focuses on giving a fixed index system.

(2) Consider the difference of indicators in time and space scale

In fact, comprehensive evaluation, especially principal component analysis, is based on certain statistical theories [80]. Indicators and their calculation methods can be divided into long-term and short-term, macro and micro indicators in time and geographical space. Some long-term statistical indicators cannot be mixed with short-term indicators, and macro indicators should not be used in micro. Attention should be paid to this difference in the selection of indicators.

(3) Focus on the general index selection methodology

At present, most of the researches on the index system of tourism environmental carrying capacity generally give a fixed index system, which is not conducive to local technicians to apply and develop their own applicable index system and evaluation methods. Therefore, we should try our best to give a general method of index selection and index system establishment, that is, the index system construction methodology, so that it can establish its own index system of tourism environmental carrying capacity according to the characteristics of local tourism environmental system and its own needs. In other words, this study attempts to provide a method of index screening and index system establishment.

(4) The construction of index system is a process of continuous improvement

Tourism natural environment system, social economic system and their interaction are constantly developing and changing; Therefore, in essence, the construction of tourism environmental carrying capacity index system should not be a static process, but a process of continuous development, improvement and enrichment. The construction process of the index system is also a management process. We must pay full attention to this dynamic and unity.

## 4.2 Structure of Eco-Tourism Environmental Carrying Capacity

The eco-tourism environmental carrying capacity reflected by the eco-tourism environmental system with complex structure is the link and intermediary between eco-tourism activities and eco-tourism environment, and reflects the coordination degree of eco-tourism activities and eco-environmental structure and function. It involves social, economic, cultural and other elements. It is a concept affected by many factors. From different perspectives and classification methods, we can get different conclusions.

## (1) Taking the main body of eco-tourism activities as the starting point

From the perspective of the main body of ecotourism activities, that is, from the classification of ecotourists, tourism economic activities include six aspects: food, accommodation, transportation, tourism, entertainment and shopping. Accordingly, ecotourism environmental carrying capacity can be divided into living environment carrying capacity, transportation environment carrying capacity, tourism environment carrying capacity, tourism land carrying capacity, natural environment pollution carrying capacity and social economic carrying capacity.

## (2) Starting from the ecotourism system

From the perspective of system theory, it includes ecotourist carrying capacity (ecotourist psychological carrying capacity), ecotourism resource carrying capacity, "ecotourism environmental carrying capacity" (natural environmental carrying capacity), and ecotourism industry carrying capacity (social and economic carrying capacity).

## (3) Starting from eco-tourism environment

From the environmental perspective of ecotourism, ecotourism environment is divided into natural environment, economic environment and social environment. Ecotourism environmental carrying capacity can include three components: natural environmental carrying capacity, economic environmental carrying capacity and social environmental carrying capacity.

This paper holds that the study of eco-tourism environmental carrying capacity is ultimately to serve the regional eco-tourism environmental planning and management, and it is more effective in practical application from the perspective of eco-tourism environment. To sum up, the structural system of eco-tourism environmental carrying capacity can be shown in Fig. 1.

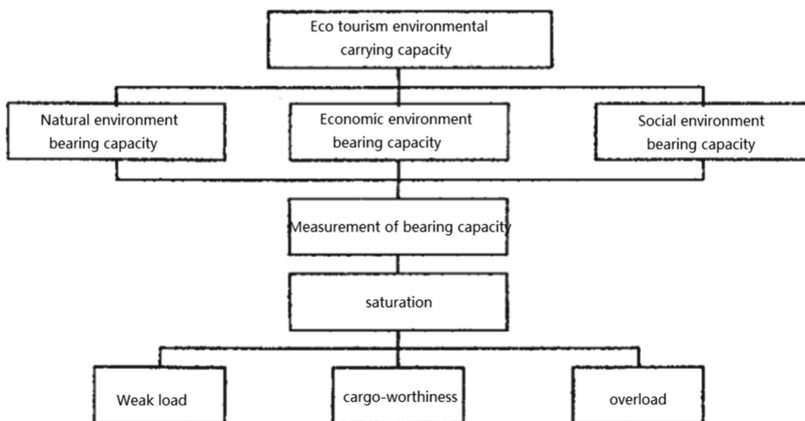


Fig. 1. The structural system of eco-tourism environmental carrying capacity

## 5 Conclusion

In this paper, based on the model of scenic spot carrying capacity of big data technology, the ecological footprint method is used to calculate the environmental ecological



footprint of soil environment, water resource environment, biological environment and pollutants in the carrying capacity analysis of scenic spots, so as to obtain the ecological environment carrying capacity of scenic spots. According to the results of the carrying capacity of the scenic spots, the ecological carrying capacity of the scenic spots should be warned to improve the application effect of the carrying capacity of the scenic spots and enhance its practical application capacity.

## References

1. Wang, X.: Research on the application of cloud accounting in small and medium-sized enterprises under the background of big data (2020)
2. Chen, Y.: Research on challenges and countermeasures of tax accounting work based on big data analysis. *J. Phys. Conf. Ser.* **1648**(3), 032162 (4pp) (2020)
3. Wang, H.: Research on object-oriented management accounting dynamic budget management based on big data. *J. Phys: Conf. Ser.* **1574**, 012021 (2020)
4. Kang, M., Ampornstira, F.: Research on data analysis of chinese public accounting firms in the big data era. *Open J. Account.* **10**(1), 1–8 (2021)
5. Xu, X.: Application research of accounting archives informatization based on big data (2020)
6. Hua, Y., Shang, D.: Research on the innovation of the new model for the cultivation of government accounting professionals under the “accrual basis of big data analysis of government comprehensive financial report. *J. Phys. Conf. Ser.* **1744**(4), 042098 (2021)
7. Aboagye-Otchere, F., Agyenim-Boateng, C., Enusah, A., et al.: A review of big data research in accounting. *Intell. Syst. Account. Finan. Manage.* **28** (2021)
8. Ibrahim, A., Elamer, A.A., Ezat, A.N.: The convergence of big data and accounting: innovative research opportunities. *Technol. Forecast. Soc. Change* (2021)
9. McBride, K., Philippou, C.: “Big results require big ambitions”: big data, data analytics and accounting in masters courses. *Account. Res. J.* ahead-of-print(ahead-of-print) (2021)
10. Balios, D.: The impact of big data on accounting and auditing. *Int. J. Corporate Finan. Account.* **8**(1), 1–14 (2021)



# Biomethanation System for Carbon Dioxide Capture in Renewable Energy Hydrogen Production Coupling Power Plant

Changchun Xu<sup>(✉)</sup>

Yangzhou University, Yangzhou 225127, Jiangsu, China  
renj1985@126.com

**Abstract.** Biomethanation system is a biocatalysis process that converts carbon dioxide into methane and hydrogen. The process of reacting biogas with CO<sub>2</sub> to produce methane and water. The main principle of the system is to use biocatalysts, which are microorganisms that can decompose organic compounds into carbon dioxide and water. The most common bacteria used in this system are methanogens. Methanogens naturally exist on rock surfaces, in soils and in deep-sea sediments, where they consume organic matter such as dead plants or animals. Syngas is also used as a fuel for power generation, which can reduce the cost of fossil fuel power generation. The biomethanation system for carbon dioxide capture in coupling power plants for hydrogen production from renewable energy has been widely used in coal-fired power plants, biomass power plants and waste incinerators. The main advantage of this process is that it does not require any external energy to operate.

**Keywords:** Renewable energy · Methane · Hydrogen production technology · Carbon dioxide capture

## 1 Introduction

In the past few decades, renewable energy, especially solar energy and wind energy, has made great progress. For example, according to the 2018 renewable energy market report of the International Energy Agency, it is expected that the market share of renewable energy in global energy consumption will rise to a historical high of 12.4% in 2023. More specifically, the share of renewable energy in the electricity market will increase from 24% in 2017 to 30% in 2023. Although these figures may vary according to the assumptions and configurations of different models, it is worth noting that the development of renewable energy (Fig. 1), especially solar and wind energy, plays an increasingly important role in energy supply and reducing greenhouse gas emissions.

Hydrogen is also an important clean energy, which is technically feasible, economically acceptable, environmentally friendly and widely used. With the continuous and rapid growth of the renewable energy market, the energy storage system is crucial to integrate intermittent renewable energy into the existing energy network. For various

reasons, hydrogen is an ideal medium for storing renewable energy of various scales. Electricity to hydrogen refers to reducing the peak load of the power grid by using water electrolysis to convert the excess electric energy into hydrogen. Electricity to hydrogen will become an important link between seasonal excess power and hydrogen production [1]. By summarizing the research status of hydrogen production by electricity, analyzing its technical bottlenecks and solutions, the future research prospects and industrial demonstration directions are put forward. The article points out that adding about 10% hydrogen to the natural gas pipeline can appropriately prolong the asset value of the existing natural gas pipeline; The separation of hydrogen rich natural gas by membrane, adsorption or other effective technologies at the downstream of the industrial chain is an effective method to obtain high-purity hydrogen. The key lies in the research and development of high selective separation technology and new materials with high hydrogen capacity [2]. In addition, the article points out that the membrane adsorption coupling process is an important development direction for the production of high-purity hydrogen from hydrogen rich natural gas. The scale, feasibility and energy consumption analysis of electrochemical hydrogen separation (hydrogen pump) is an important development direction, while the cryogenic refrigeration separation technology has certain feasibility when liquefied natural gas is the main product.

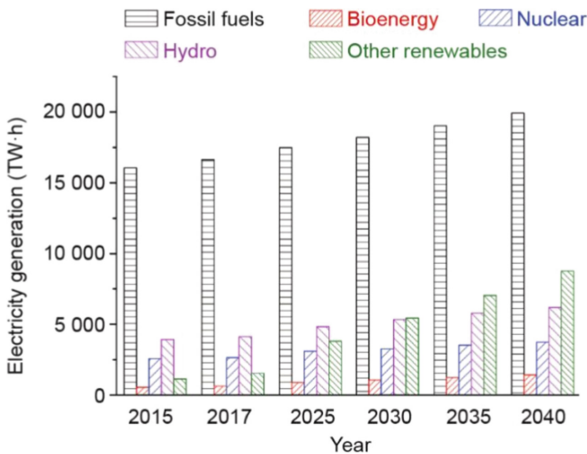


Fig. 1. Development of renewable energy

## 2 Related Work

### 2.1 Development of Hydrogen Energy

Traditional hydrogen production methods include hydrocarbon reforming, gasification, hydrocarbon pyrolysis, biomass and water decomposition (electrolysis, solar pyrolysis or photocatalysis). The current annual hydrogen production is  $1 \times 10^8$  T, mainly for production site consumption, used for oil refining and metal treatment. Nowadays, the

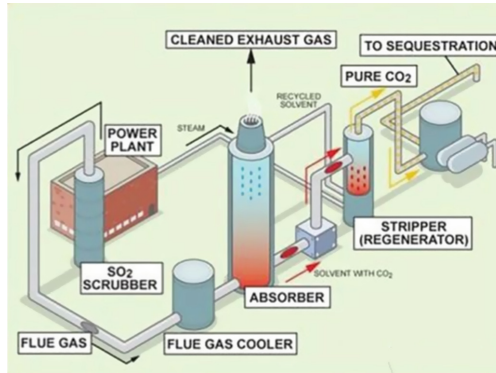
rise of fuel cells has increased the demand for hydrogen and significantly increased the demand for hydrogen production [3].

Hydrogen is an important chemical raw material for ammonia synthesis, oil refining (i.e. converting crude oil into diesel and aircraft fuel), gas purification (i.e. removing sulfur and nitrogen compounds from fuel), and other industries. It is also a substitute for carbon monoxide in the steel smelting industry. In addition, hydrogen is also an important clean energy. Because it is technically feasible, economically acceptable and environmentally friendly, it can also be used for many purposes, such as heat source, power generation, fuel cell and transportation (trucks and trains). Taking California as an example, it is reported that as of September 2017, more than 3000 fuel cell electric vehicles have been sold or leased. As of May 22, 2019, there were 40 hydrogen stations in operation. In 2017, more than 17 million people took hydrogen fuel cell buses. At the same time, there is evidence that adding hydrogen to traditional fuels may improve energy efficiency, reduce fuel consumption, and reduce carbon emissions. With the extension of the service life of fuel cells (60 000–90 000 h, the data to be proved, while the target set by the U.S [4]. Department of energy is 60 000–80 000 h), a series of automobile manufacturers such as Hyundai, Toyota, Honda and Mercedes Benz are involved in the field of fuel cells. All these advantages will increase people's confidence in the wide application of fuel cells. Due to the diversity of hydrogen sources and utilization ways and its low-carbon/low pollutant emission characteristics, hydrogen will play an important role in the clean and safe energy future. As a short-term (hourly), medium-term (daily) and long-term (seasonal) energy storage carrier, people also fully realize that hydrogen is also very important in maintaining the stable supply of renewable energy (especially solar energy and wind energy) with fluctuations and seasonal changes. Therefore, Germany, other European countries, the United States, Australia and many other countries have put forward strategic plans for the hydrogen industry.

## 2.2 Carbon Dioxide Capture Technology

The CCS (carbon dioxide capture and storage) technology can be divided into three steps: capture, transportation and storage. The commercial carbon dioxide capture has been in operation for a period of time and the technology has been relatively mature, while the carbon dioxide storage technology is still undergoing large-scale experiments in various countries. The process flow of carbon dioxide capture is shown in Fig. 2 below.

There are three kinds of carbon dioxide capture methods: pre combustion, oxy fuel combustion and post combustion.



**Fig. 2.** Carbon dioxide capture process flow

Pre combustion capture is mainly used in IGCC (integrated coal gasification combined cycle) system, which converts coal into gas by high-pressure oxygen enriched gasification, and then generates CO<sub>2</sub> and hydrogen (H<sub>2</sub>) after water gas conversion. The gas pressure and CO<sub>2</sub> concentration are very high, so it is easy to capture CO<sub>2</sub> [5]. The remaining H<sub>2</sub> can be used as fuel. The capture system of this technology is small, low energy consumption, and has great potential in efficiency and pollutant control, so it has been widely concerned. However, IGCC power generation technology still faces problems such as high investment cost and high reliability.

The oxygen enriched combustion adopts the technical process of the traditional coal-fired power station, but through the oxygen production technology, a large proportion of nitrogen (N<sub>2</sub>) in the air is removed, and the mixed gas of high concentration oxygen (O<sub>2</sub>) and part of the extracted flue gas (flue gas) is directly used to replace the air. In this way, the flue gas obtained contains high concentration CO<sub>2</sub> gas, which can be directly processed and stored.

### 3 Bio Methanation System for Carbon Dioxide Capture in Renewable Energy Hydrogen Production Coupling Power Plant

Biomethane system has the characteristics of many unit technologies and many model variables. Its topological network design and key unit modeling are the basis of system optimization and evaluation, and are of great significance. The electrolysis device is a device that uses electricity to decompose water into hydrogen and oxygen. When electricity generated by renewable energy is used, hydrogen becomes the carrier of renewable energy and complements electricity. The electrolysis device helps to integrate VRE into the power system because its power consumption can be adjusted according to wind and solar power generation, in which hydrogen becomes the storage source of renewable power [6]. Therefore, it can provide flexible load and grid balancing services (up and down frequency regulation), and can operate at the optimal production capacity to meet the hydrogen demand of the industrial and transportation departments, or the

injection demand of the natural gas pipeline network. As shown in Fig. 3 below, the framework of hydrogen production biomethanation system is shown.

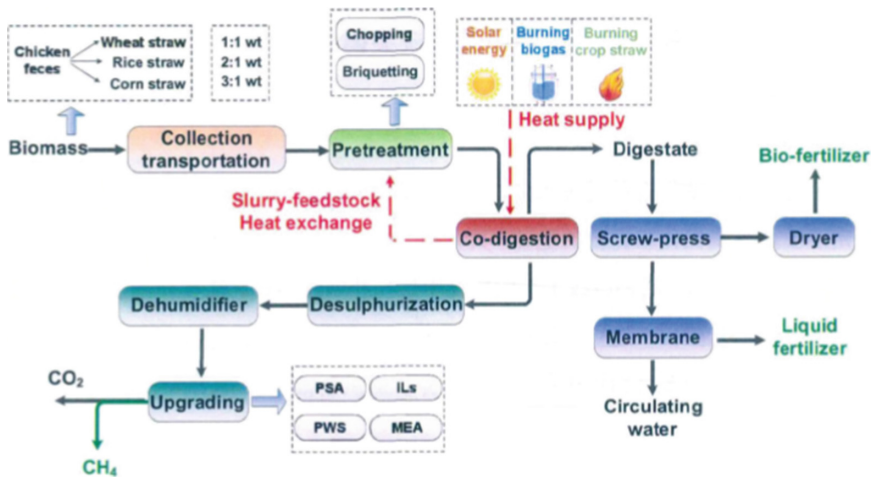


Fig. 3. Framework of biomethanation system for hydrogen production

Hydrogen produced by renewable energy power can create a new downstream market for renewable power. If part or all of the renewable energy generation capacity is sold to the electrolysis unit operator through long-term contracts, the risk of price fluctuation faced by renewable power producers can be reduced [7].

The system uses the CO<sub>2</sub> captured by the coal-fired power plant as the compressed energy storage medium of the renewable energy power plant. When the power generation capacity of the renewable energy power plant is excessive, the CO<sub>2</sub> captured and transported by the nearby coal-fired power plant is compressed into the CO<sub>2</sub> storage tank to store the excess power of the power plant in the compressed CO<sub>2</sub>. When the power generation capacity of the renewable energy power plant is at a low point and cannot meet the power grid demand, the compressed CO<sub>2</sub> in the storage tank is used to drive the gas turbine impeller to run and drive the generator rotor to complete power generation, so as to supplement the power generation capacity and meet the power grid demand. It is worth noting that before the operation of the CO<sub>2</sub> compression energy storage system, a large amount of CO<sub>2</sub> needs to be captured by the CO<sub>2</sub> capture system of the coal-fired power plant as the circulating working fluid. When the energy storage system starts to operate, the demand for CO<sub>2</sub> will be reduced. However, due to leakage and other factors, the system itself will generate CO<sub>2</sub> loss, so it is still necessary to supplement CO<sub>2</sub> from the CO<sub>2</sub> capture system regularly [8]. After the operation of the CO<sub>2</sub> compression energy storage system, due to the reduced demand for CO<sub>2</sub>, the CO<sub>2</sub> captured by the coal-fired power plant can be used for other purposes, such as CO<sub>2</sub> synthetic chemicals.

In addition, the excess power generation of wind power and solar power plants under favorable power generation conditions can provide energy sources for photocatalytic conversion or electrochemical conversion of CO<sub>2</sub>, so that CO<sub>2</sub> can be converted into

high-energy compounds such as formic acid, so as to realize the storage of excess power generation.

### 4 Simulation Analysis

The calculation of the solar cell model shows that the open circuit voltage fluctuation range of the solar cell is 18.6–23.5 v, the fluctuation range of the maximum power point voltage is 14.9–18.8 v, and the protection voltage of the electrolytic cell is 15 V, which will work before the maximum power point of the solar cell; Since the fluctuation range of the short-circuit current of the solar cell is 0–14.5 a, according to the ideal Faraday efficiency of 100%, 1 A generates 6.97 ml per minute, the maximum flow rate of the electrolytic cell with five membranes in series is  $14.5 \times \text{six point nine seven} \times 5 = 505 \text{ ml/min}$ , meeting the requirement of the maximum flow of the electrolytic cell of 500 ml/min, because the electrolytic cell can allow the current to exceed the maximum value in a short time, and the Faraday efficiency of the electrolytic cell can not reach 100% [9]. Therefore, the direct coupling of the solar cell and the electrolytic cell is feasible. As shown in Fig. 4 below, the change of external light intensity of the current module.

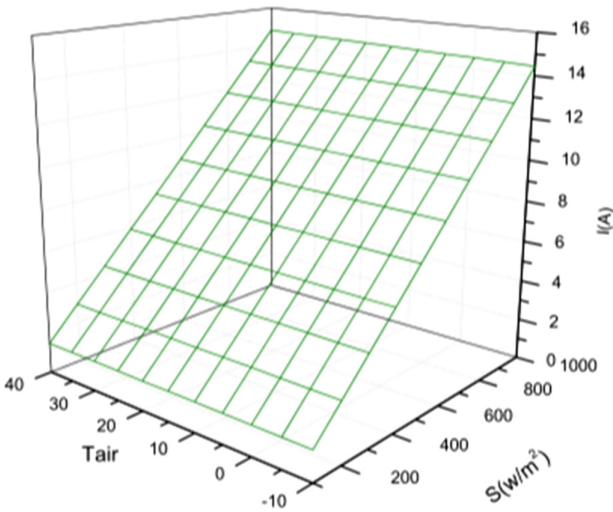


Fig. 4. Changes of external light intensity in current module

According to the calculation that the current of theoretical hydrogen production 1A can produce 6.97 ml of hydrogen, it can be obtained that the hydrogen production is in direct proportion to the light intensity [10]. See Fig. 4. The relative accuracy of temperature on the hydrogen production can be 5%. With the increase of light intensity, the hydrogen production also gradually increases.

## 5 Conclusion

The biomethanation system for carbon dioxide capture in renewable energy hydrogen generation combined shaft power plants is a new type of bioreactor that can be used to convert CO<sub>2</sub> into useful products, such as methanol, ethanol and other organic compounds. The main advantage of this process is that it can be combined with any existing power generation technology, including solar photovoltaic and wind turbines. The most important feature of biomethanation system is that it can use renewable energy (wind energy or solar energy) to generate electricity and heat at the same time. However, with the recent technological progress, the application of this method in power plants has become more feasible. The main advantage of this technology is that it produces pure hydrogen instead of other methods, in which impurities such as water or carbon monoxide may exist in the final product.

## References

1. Liu, Y., Zhu, Q., Zhang, T., et al.: Analysis of chemical-looping hydrogen production and power generation system driven by solar energy. *Renew. Energy* **154** (2020)
2. Kassem, N., Hockey, J., Lopez, C., et al.: Integrating anaerobic digestion, hydrothermal liquefaction, and biomethanation within a power-to-gas framework for dairy waste management and grid decarbonization: a techno-economic assessment. *Sustain. Energy Fuels* **4** (2020)
3. Herdem, M.S., Mazzeo, D., Matera, N., et al.: Simulation and modeling of a combined biomass gasification-solar photovoltaic hydrogen production system for methanol synthesis via carbon dioxide hydrogenation. *Energy Convers. Manage.* **219C**, 113045 (2020)
4. Yuksel, Y.E., Ozturk, et al.: Energetic and exergetic performance evaluations of a geothermal power plant based integrated system for hydrogen production. *Int. J. Hydrogen Energ.* (2018)
5. Daglioglu, S.T., Ogut, T.C., Ozdemir, G., et al.: Comparative evaluation of two packing materials (Glass Pipe and Ceramic Ball) for hydrogenotrophic biomethanation (BHM) of CO<sub>2</sub>. *Waste Biomass Valorization* **1**, 1–10 (2020)
6. Dalolu, T., Tuba Ceren üt, Ozdemir, G., et al.: Comparative analysis of the effect of cell immobilization on the hydrogenotrophic biomethanation of CO<sub>2</sub>. *Greenhouse Gases Sci. Technol.* (2021)
7. Bellotti, D.: Thermo-economic analysis of innovative systems for renewable energy storage by chemicals production (2018)
8. Corumlu, V., Ozturk, M.: Performance analysis of a novel solar energy-based combined plant for alternative fuels production. *Int. J. Energy Res.* (2021)
9. Fy, A., Mo, A., Rs, B.: Design and thermodynamic modeling of a renewable energy based plant for hydrogen production and compression. *Int. J. Hydrogen Energy* **45**(49), 26126–26137 (2020)
10. Moioli, S., Pellegrini, L.A.: Fixed and capture level reduction operating modes for carbon dioxide removal in a natural gas combined cycle power plant. *J. Clean. Product.* **254**,120016 (2020)





# Children's Safety in Smart Home Based on Intelligent Optimization Algorithm

Wenxin Yang(✉)

Guangzhou Vocational University of Science and Technology, Guangzhou 510550, Guangdong, China

1656907686@qq.com

**Abstract.** The research on children's safety in smart home based on intelligent optimization algorithm is a new research work. We are developing a system based on intelligent optimization algorithm to analyze children's safety in smart home. The system will be used to monitor and control intelligent devices in the house to ensure that they will not cause any harm to children. The purpose of this study is to design an optimal control strategy for the smart home system, which can prevent children from being injured by minimizing the risk of accidents and injuries, while maximizing the benefits obtained from the system. This study was conducted using a simulation software called matlab r2010a (the MathWorks Inc., Natick, MA). In order to achieve this goal, we consider various aspects, such as: 1) defining the model of security level; 2) Model for calculating accident probability; 3) Algorithm. The system adopts intelligent algorithm design and can analyze all data related to child safety from different sources (such as sensors, cameras or other information systems). It then uses this data to determine whether there is any danger or risk of injury to specific equipment in your home.

**Keywords:** Intelligent optimization algorithm · Smart home · Child safety

## 1 Introduction

For us adults, home is just a tool or a background board that has no sense of existence in most scenes. But you know, in the world of children, home is not only a paradise for fun, but also may be full of fatal traps. In this children's festival, I will briefly discuss various hidden dangers that may exist in children's homes and are often ignored according to relevant national standards [1]. If you feel useful, I hope you can share them with more parents to reduce unnecessary child injuries, so that every baby can avoid these avoidable risks and grow up safely and happily.

Parents know to buy car seats for their baby, but it is easy to ignore that the baby stays at home for the longest time, and all kinds of things at home may cause potential harm to the baby. It is very necessary to do a good job in children's home safety in advance to prevent trouble! When the baby starts to learn to turn over and climb, it will soon move around the home by itself, but the parents' eyes cannot keep an eye on the child 24 h a day, so it is very important to provide a safe home environment for the child.

Because children's home is more used in children's life scenes, they pay more attention to safety than ordinary home. For children's home furnishings, China has a special standard GB 28007-2011 "general technical conditions for children's home furnishings". Everyone should note that this standard starts with GB rather than gb/t, that is, it is a compulsory standard, which is the basic requirement that any children's home furnishings must meet [2]. If it does not meet this standard, it will be regarded as unqualified products and prohibited from domestic circulation. Based on this, this paper studies the research of children's safety in smart home based on intelligent optimization algorithm.

## 2 Related Work

### 2.1 Smart Home Concept

With the help of the developed telephone network and wireless network, the intelligent home control system designed in this paper can monitor and control the situation at home far away from home. It has become a tool to make the room more comfortable, safer, efficient and economical. Smart home is one of the hottest topics in modern society. Its goal is to realize the control of smart devices and communicate through the Internet and information technology, so that they can work together according to everyone [3]. At present, more and more institutions and individuals have started the research of smart home.

With the development of network technology, especially the development of wireless network, smart home has great room for development. Combined with remote monitoring, home appliance control, lighting control, indoor and outdoor remote control curtain automation, alarm system, telephone remote control, multiple functions and control devices, it is used to control programming and computer synchronization, making life more comfortable, convenient and safe.

Smart home is a residential platform, using generic cabling technology, network communication technology, security technology, automatic control technology, audio and video technology to integrate the facilities related to home life, build an efficient residential facilities and family schedule management system, improve the safety, convenience, comfort, artistry of home, and realize an environment-friendly and energy-saving living environment [4]. It is the integration of Internet technology and Internet of things technology.

Internet of things technology provides more services and better user experience for smart home. The traditional smart home connects products through wired mode, and can only be controlled locally in a short distance through the wall mounted touch control screen and remote control board. The system debugging is also relatively complex, and the functions that can be realized are relatively simple. With the landing and acceleration of new technologies such as NB-IoT, the attention of the smart home market has increased significantly, which shows that the overall development of the industry is in a rapid development trend. Smart home based on Internet of things technology can not only provide control functions, but also provide data analysis and other functions. It can remotely control smart home products through mobile terminal app to provide real intelligent life conveniently and safely. Figure 1 below shows the smart home security device.



Fig. 1. Smart home security device

## 2.2 Threats to Children's Home Safety

1. Prevent the baby from swallowing small things and suffocating  
 Many babies like to put things into their mouths, so mothers should put away all kinds of small things at home, such as coins, electronics, beads and other small things should be cleaned up to prevent babies from accidentally swallowing them. All small things that can be stuffed into the baby's mouth should be put away, and whether they can pass through the roll paper core can be used as the standard.
2. Avoid poisoning by eating various chemicals  
 Cosmetics, detergents, drugs, health products and other things containing chemical ingredients at home should be placed out of the reach of the baby, or in the locked cabinet door. If there are potted plants at home, try to choose non-toxic varieties and put them out of the reach of your baby [5].
3. Avoid electric shock to baby  
 All power sockets are covered with covers, but the baby may suffocate by picking up the cover and stuffing it in his mouth. Therefore, it is recommended to directly replace the power socket board with the elastic one.  
 If you use an extended wire board, remember to seal the useless plugs with power tape, and the unused electrical plugs should be pulled out and put away in time.
4. Avoid furniture overturning or things falling and hitting the baby  
 Every year, there are many cases of baby injuries caused by furniture overturning or things falling from shelves. Heavy furniture such as TVs, wardrobes and bookcases can be fixed to wall piles with special anti tip anchors angle iron sheets or belts, so as to ensure that they will not overturn.

All cabinets with unstable center of gravity, longer height than width, or more drawers are necessary - once fixed, the center of gravity of many drawers and cabinets is unstable after they are pulled apart. Just a few nails on the wall can reduce the potential safety hazards.

5. the sharp corners of furniture should be wrapped to avoid hurting the baby  
The sharp corners of the furniture and the hard edges of the steps should be covered with bumpers to prevent the baby from being hurt and reduce the baby's injury of accidentally falling down from the stairs.

### **3 Research on Children's Safety in Smart Home Based on Intelligent Optimization Algorithm**

The security system in smart home requires the integration of anti-theft, robbery prevention, fire protection, gas leakage alarm and other functions, and uses various advanced and intelligent electronic security equipment to conduct real-time monitoring. In case of alarm events such as children's safety hazards, it will automatically alarm and generate linkage through mutual communication and cooperation [6].

The processing of a series of alarm events in the security system is basically that the security controller uses the alarm detection module to detect all kinds of detectors in real time. Once an abnormal event such as out of range or state change occurs, it sends an event notification service, and then requires the corresponding security equipment to take measures to make a series of chain reactions. The implementation of this alarm linkage requires a set of preset linkage rules. When the security controller detects an alarm event, it calls this set of linkage rules to query and match the trigger conditions of the rules. If the trigger conditions are met, the linkage action of the rules will be executed.

It can be seen that the formulation of linkage rules and scheduling matching, that is, the design of security strategy, are the core and key of alarm linkage in the security system [7]. Because BACnet protocol provides a unified communication language for various devices, it is easy to find a unified expression for the design of security policy based on BACnet to realize the linkage between various security devices.

In general, the operation mode of security policy is divided into global policy and local policy. The local policy runs on the local security controller, while the global policy runs on the HIC. Dividing the strategy into global strategy and local strategy design can greatly enhance the flexibility and self-organization of the system. The strategies that can be implemented in the local security controller are solved locally without the transfer of HIC, which improves the efficiency, makes the security controller independent and self-contained, and reduces the burden of HIC at the same time; On the other hand, simple strategies that are limited to local execution are executed by the local security controller, while complex strategies that need to deploy different controllers are executed on the HIC, which reduces the load of the security controller and improves its flexibility. The security controller and HIC are interconnected through BACnet network for information exchange, and HIC obtains the status information of field equipment through the security controller [8]; The security controller notifies the HIC of alarm events; The

decision command of global strategy issued by HIC to the security controller; The security controller informs the implementation of its local strategy. The structure of the two operation modes in the system is shown in Fig. 2.

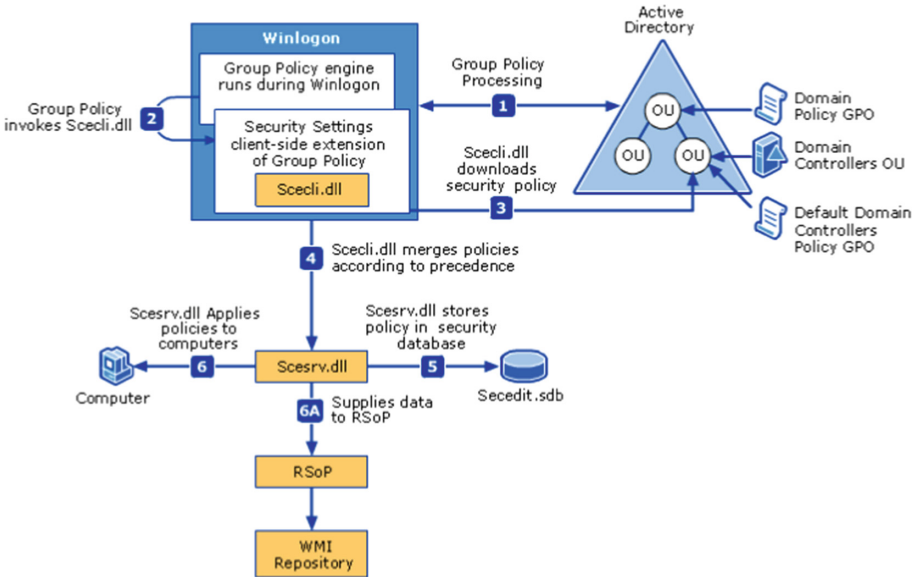


Fig. 2. Security policy operation system framework

## 4 Security Data Structure Design

The alarm linkage of the security system is that once an alarm occurs in a certain device, it will immediately link with other devices or the system will automatically cooperate to deal with the alarm. From the previous research on the representation of policy rules, we can get a series of specific linkage rules. Executing each policy action is actually an operation on the attribute value of one or some BACnet objects of a specific device [9]. We need to abstract a data structure from these policy rules to completely describe the policy. By studying these specific linkage rules, we can see that each rule is composed of two parts: the motivator (trigger condition) and the responder (linkage action).

In addition, the security strategy is divided into three modes according to the actual security needs: master at home mode, master out mode and master rest mode. Therefore, the security mode should be selected for each security strategy. Moreover, considering the conflict of policy implementation, we also need to assign priority to each policy, and the priority will be reduced from small to large. In case of policy execution conflict, the policy with high priority shall be executed first [10].

Among them, the node definitions of alarm equipment linked list, auxiliary equipment status linked list and response equipment operation linked list are shown in Fig. 3 below respectively:

```
typedef struct s_activeObj
{
    BACnetObjectIdentifier activeObjID;
}BACnetactiveObj;
typedef struct s_subactiveObj
{
    BACnetObjectIdentifier activeObjID;
    BACnetrealorbool activeObjval;
}BACnetsubactiveObj;
typedef struct s_jointObj
{
    BACnetObjectIdentifier objID;
    BACnetObjectIdentifier deviceObjID;
    BACnetrealorbool jointObjval;
}BACnetjointObj;
```

Fig. 3. Project code

## 5 Conclusion

The research on children's safety in smart home based on intelligent optimization algorithm is the research on children's safety in smart home. The main focus of this study is to find out how to use smart devices such as smartphones, tablets, laptops and other devices to reduce the risk of children and their parents. The project aims to identify the risks that can be reduced by using these devices with children. It also tries to find out whether there are any risks that cannot be reduced or even increased by using these technologies. The purpose of this study is not only to improve the understanding of children's safety, but also to improve children's safety awareness.

## References

1. Cheng, J., Ying, L.I.: Application of resistance psychology in packaging design based on children's safety assurance. *Packag. Eng.* (2018)
2. Guo, Q., Zhang, W., Zhang, F., et al.: Research and analysis on the identification algorithm based on the random near parameter estimation of children's running. *J. Phys: Conf. Ser.* **1802**(3), 032102 (2021)
3. Liu, Q., Shen, L.: Design of intelligent jackets based on security concept for children. *Wool Textile J.* (2018)

4. Liu, X., Duan, S., Pei, F., et al.: Safety risk assessment for children's products based on reinforcement learning. In: ICAAI 2020: 2020 The 4th International Conference on Advances in Artificial Intelligence (2020)
5. Chi, H., Kang, S.U.: Multiobjective metaheuristic load control algorithm for interaction between smart home and humans. *Math. Probl. Eng.* **2022** (2022)
6. Shankpal, S.V., Brahmananda, S.H.: Design and development of trust management scheme for the internet of things based on the optimization algorithm. In: 2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT). IEEE (2020)
7. Vanus, J., Fiedorova, K., Kubicek, J., et al. Wavelet-based filtration procedure for denoising the predicted CO<sub>2</sub> waveforms in smart home within the Internet of Things. *Sensors* **20**(3), 620 (2020)
8. Melhem, F.Y.: Méthodes d'optimisation et de gestion de l'énergie dans les réseaux intelligents "Smart Grids" (2018)
9. Mc, A., Kes, B.: Long-term efficacy and safety of repeated botulinum toxin a applications based on function and anesthesia type in children with cerebral palsy - ScienceDirect. *J. Orthopaed.* (2022)
10. Hassan, A.S.: Analysis on the safety aspects in malaysian children playgrounds. *Adv. Sci. Let.* (2018)



# Clustering Analysis of English Scores Based on K-means Clustering Algorithm

Yueying Shen<sup>(✉)</sup>

School of Foreign Languages, Guangzhou Institute of Science and Technology,  
Guangzhou 510000, Guangdong, China  
syy\_winnie@163.com

**Abstract.** The integration of Internet information technology and education has brought tremendous changes to the whole education system. At the same time, English curriculum standards require teachers to make full and reasonable use of modern educational technology in daily teaching according to the actual situation of students and teaching content. This paper introduces a k-means algorithm. The k-mean algorithm based on LAGSA and the traditional K-mean algorithm were simulated and compared with the measured data. Finally, the clustering results show that the Load-clustering algorithm based on LAGSA has less iteration times, fast convergence speed, good clustering accuracy, strong anti-noise performance and good robustness. Experimental results show that the improved algorithm can not only improve the efficiency of curve clustering, but also refine the mining of English score curves of different students, reflecting the rules and characteristics.

**Keywords:** K-means Clustering · English Achievement · Cluster Analysis · Intelligent Algorithm

## 1 Introduction

Student achievement level is not only an objective and important comprehensive index reflecting the quality of the comprehensive teaching practice of the school, but also an important symbol to evaluate the overall effect of students' learning practice and the accuracy of key knowledge content. At the same time, the school also regards improving student achievement as an important teaching goal. With the continuous and rapid expansion of the total scale of China's university system, the number of students in schools is increasing, which brings great pressure to the effective management of schools [1, 2]. Along with the education the further rapid development of information technology, further the reform of college management information system construction and deepen constantly, the relevant management concept to use information technology to the colleges and universities have all within the school of education and teaching information management and working system, in which the management level has greatly improved. Colleges and universities in undergraduate education and teaching administration practice in a large number of the use of a variety of systems, such as student information management system, teaching evaluation system and so on, transport some systems



accumulated a lot of data. However, administrators only carry out simple operations such as statistics of male and female students and passing rate of examinations, so they can't get the information hidden in these data that is helpful for teachers' teaching and students' learning [3, 4]. In order to utilize these data effectively, k-means clustering algorithm is introduced.

The application of clustering analysis of English score data is based on clustering and analysis of historical English score data. The prediction day model is constructed by extracting some sample data sets of simulated training results with similarity to historical prediction day results from cluster analysis research results [5, 6]. In the whole process of students' learning and application analysis in the study, through the integrated use of intelligent English curve statistical methods such as clustering method to respectively on the data characteristics of the above student achievement has carried on the comprehensive statistical analysis, better able to further the above all kinds of English level is not below a certain level of the same type share of students learning degree requirements is open, in order to establish the student needs The research of the lateral feedback response analysis system has laid and laid some more extensive and solid effective and reliable theoretical foundation. In recent years due to as smart quickly learn English software, information and other areas widely involved with the scope and content of expanding constantly, all kinds of English education has been in the stage of leading-edge technology learning automation intelligent platform technology and English information and management automation application system are also gradually obtained the very good and effective practice research and promotion of innovation and development and utilization. The scientific and efficient clustering method is not only a mathematical foundation that can directly provide enrollment majors of colleges and universities at all levels of society, but also a service that can easily and accurately provide relevant information for students of colleges and universities all over the country [7, 8]. The basic structure, functions and forms of information are different, and it is difficult to compare unified processing information and accurately distinguish its authenticity [9, 10]. However, the existing large number of unsolved classification algorithm items in our real life system have been difficult to fully meet the above classification requirements, especially when processing massive data, the difficulty is gradually increasing. If only the pre-processing class method is used to make the mass data completely meet the classification requirements of the items of our classification algorithm, the cost may be very high, so clustering algorithm can be completely tried [11, 12].

In this paper, k-means clustering algorithm is used to analyze students' performance. On the one hand, students can deeply understand the influencing factors of learning, more targeted education of their own learning status, put forward suggestions to improve learning methods, reduce the blindness of learning, and improve their academic performance. On the other hand, it can also make students realize their own objective and subjective factors that affect their academic performance, reduce confusion and annoyance, and improve their academic performance.

## 2 Proposed Method

### 2.1 K-means Clustering Algorithm

In the k-means algorithm model, the actual distance between each object and the center point in each cluster may need to be recalculated in general. The following numerical methods are often used to directly measure the actual distance: for example, the actual distance between Minkowski, the actual distance between hamanton and euclidean. These formulas can be selected for different types of discrete distances. Euclidean distance operation has always been the most effective and preferred algorithm for solving the distance of K-means algorithm program under the normal computer working and running environment. In the process of actually solving the k-means algorithm program in the actual working and running environment, the maximum similarity of the data distance between each calculated object point and the central point of its data cluster also lies in the need to calculate the distance continuously or multiple times. It can really play a great role and improve the operation and iterative calculation speed of the algorithm itself to a certain extent. The best problem-solving method is to directly arrange several data points in any lower one-dimensional Euclidean space cluster set, which can temporarily avoid considering and calculate the system through the minimum similarity of data points between any number of other data point sets. Clustering algorithm analysis is a completely unsupervised machine learning method In a general technical sense, the clustering algorithm can cluster all the target data sets directly according to the spatial distance between the data of multiple target record sets (objects), so that the spatial distance of the data of each set of multiple target records in the same data set is as small as possible, The distance between multiple target record sets (objects) with different class numbers is as large as possible Therefore, the recording system (object) in the same type of system has the greatest similarity, while the same recording subsystem (object) formed in different subclasses has the greatest difference. Because of its simplicity and fast convergence speed, K-means algorithm has been widely used in various research fields For example, K - means is used to delete noisy data However, the disadvantage of K-means is that it is easy to fall into local optimization as a remedy. At present, the most popular trend is to combine genetic algorithm with K-means to obtain genetic k-means algorithm. The task of cluster analysis is to cluster a group of N data elements ( $I = 1, 2, 3, 4 \dots N$ ) into class K Each gene can be expressed as a d-dimensional vector, for  $I = 1, 2, 3 N$  and  $K = 1, 2, 3 N$  is defined as follows:

$$w_{ik} = \begin{cases} 1, \\ 0, \end{cases} \quad (1)$$

In addition, the definition of class mark matrix  $w = WIK$  is as follows: we require that each record belongs to exactly one class, and each class contains at least one record. Therefore:

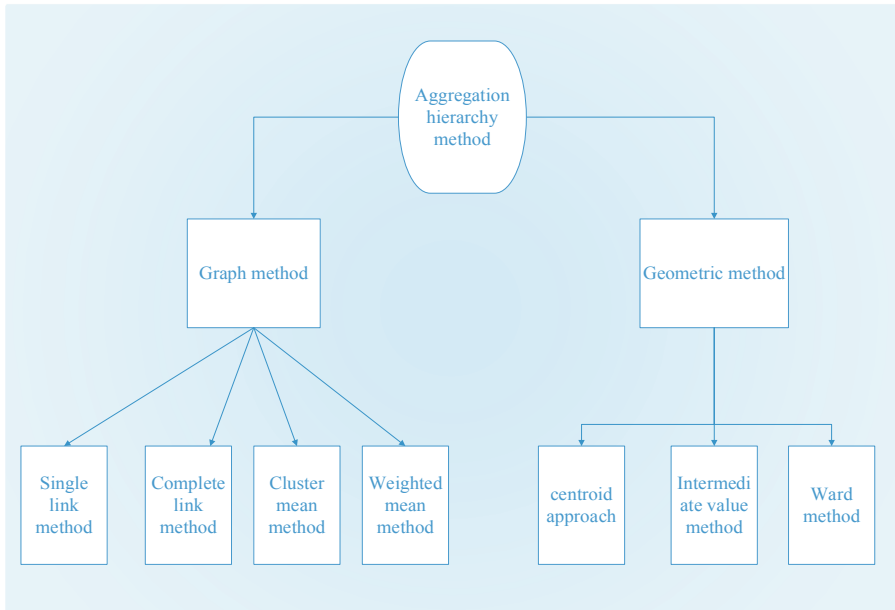
$$\sum_{i=1}^k w_{ik} = 1, i = 1, 2, 3, \dots, n$$

$$1 \leq \sum_{i=1}^n w_{ik}, k = 1, 2, 3, \dots, k \quad (2)$$

## 2.2 Hierarchical Clustering Algorithm

Hierarchical clustering algorithm uses a mathematical process of bottom-up aggregation. First, each aggregated data should be regarded as a separate class, and then it should be measured according to a certain distance, or two or more aggregated classes in a small range of this distance should be combined in turn, and these combined aggregation steps should be performed repeatedly until there can be only one class at last. Implementation process of aggregation hierarchical clustering algorithm:

1. Select an appropriate distance measure according to data set  $X$  and calculate distance matrix  $D$ .
2. Repeat
3. Select the two nearest clusters  $C_a$  and  $C_b$  and merge them into  $C_a \cup C_b = C_a \cup C_b$ .
4. Delete the rows and columns of  $C_a$  and  $C_b$  in the distance matrix, add  $C_a \cup C_b$ , and update the distance.
5. Finally, output the hierarchical tree (Fig. 1).



**Fig. 1.** Aggregation hierarchy method

### 3 Experiments

#### 3.1 Determine Research Objects and Objectives

The research object of the experiment is all undergraduate students in a university. Clustering algorithm is used to analyze the English performance of experimental students. Taking the score of CET-4 as an example, this paper analyzes the impact on the score of CET-4, and provides scientific basis for English teachers' teaching and help students' learning through clustering algorithm.

#### 3.2 Data Preprocessing

Since the collected and detected experimental data do not accurately meet the detection requirements, technicians need to carry out various processing operations on various original data to improve the analysis efficiency and quality of the detection experimental analysis results. This chapter mainly preprocesses the score of CET4.

##### 1. Data cleaning

In the information column of filling in CET-4 examination results, some students in some classes can only take the examination once at a time, some students in some classes can take the examination many times, and some examination class results are not filled in so completely. Forgive me, some do not fill in gender, some student numbers and some majors. The basic methods to deal with these vacancy values are as follows:

1. 1 fill in the vacancy value manually:
2. Fill in the vacancy value with the average value;
3. Ignore numerical method

When dealing with CET-4 score information, the score data is incomplete and there are vacancy values, which almost all come from students' absence from the exam. The w adopts the ignore numerical method and uses the screening function of Excel to directly delete the score records of these students; If the student number, class, gender and major are not filled in completely, the vacancy value shall be filled in manually.

2. Data selection

After data cleaning, the wrong data has been deleted and cleaned up, and the incomplete data has been supplemented by various methods. Data selection is to consider the relationship between each attribute and mining task. In order to reduce the time required for data mining and not affect the mining results, data attributes need to be selected. The work in this paper is as follows: in the students' CET-4 test score information, use the screening function of Excel to select the score records of undergraduate students in the three test scores; Delete unnecessary attributes in the dataset and retain the required attributes, such as student number, gender, total score, absence, individual score, major and so on.

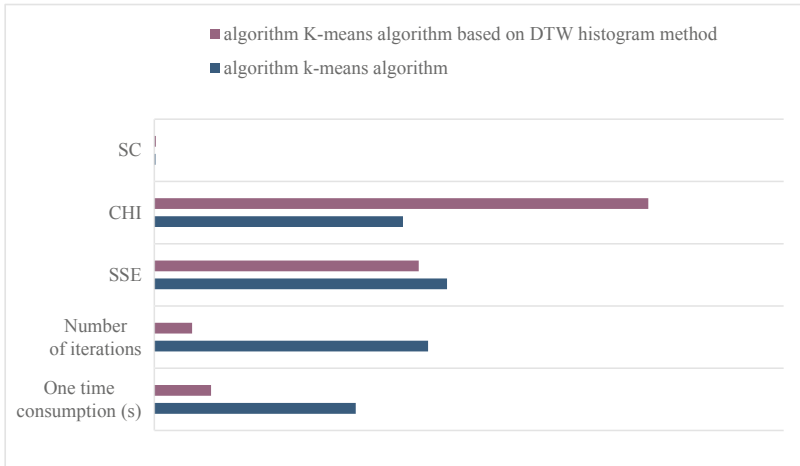
## 4 Discussion

### 4.1 Data Sources

Relevant experiments are carried out to verify the research of K-means clustering algorithm on clustering analysis of English scores. The data are shown in the table below (Fig. 2 and Table 1):

**Table 1.** Algorithm performance optimization comparison

index	algorithm	
	k-means algorithm	K-means algorithm based on DTW histogram method
One time consumption (s)	64	18
Number of iterations	87	12
SSE	93	84
CHI	79	157
SC	0.35	0.47



**Fig. 2.** Algorithm performance optimization comparison

K-means clustering are from the above chart, the algorithm has the English the clustering analysis method and the research object of many related analyses experiment, the results have proved that the k-means algorithm are based on the English data as the validity of clustering analysis, the experimental result shows that the subjects of English in good condition, During the experiment, the iterative times of the algorithm are reduced obviously, and the iterative convergence speed of the algorithm results is accelerated, and the number of clustering can achieve the good effect of obtaining the global optimal solution. It effectively improves the efficiency of analyzing students' English scores.

## 5 Conclusions

In this paper, K-means clustering algorithm is used to analyze the English scores of all undergraduates in a university; The data obtained by relevant software are analyzed. Finally, through experimental comparison, it shows that K-means clustering algorithm is scientific and feasible to cluster students' English scores, and the efficiency of K-means clustering algorithm is higher than other algorithms, and the reliability of the results is also higher. Therefore, K-means clustering algorithm can be used in the research and analysis of students' English scores.

## References

1. Kumar, A., Kumar, S.: Density based initialization method for k-means clustering algorithm. *Int. J. Intell. Syst. Appl.* **9**(10), 40–48 (2017)
2. Yunoh, M., Abdullah, S., Saad, M., et al.: K-means clustering analysis and artificial neural network classification of fatigue strain signals. *J. Braz. Soc. Mech. Sci. Eng.* **39**(3), 757–764 (2017)
3. Jajoo, G., Kumar, Y., Yadav, S.K., et al.: Blind signal modulation recognition through clustering analysis of constellation signature. *Expert Syst. Appl.* **90**(dec.30), 13–22 (2017)

4. Riaz, M., Habib, A., Khan, M.J., et al.: Correlation coefficients for cubic bipolar fuzzy sets with applications to pattern recognition and clustering analysis. *IEEE Access* **99**, 1 (2021)
5. Balik, P.Y., Demirc, E., Konca, M.: Comparison of European countries' health indicators and health expenditures by clustering analysis. *Ömer Halisdemir Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi* **14**(2), 365–377 (2021)
6. Bulla-Cruz, L.A., Barrera, L.L., Darghan, A.: Complete-linkage clustering analysis of surrogate measures for road safety assessment in roundabouts. *Revista Colombiana de Estadística* **44**(1), 91–121 (2021)
7. Bhowmick, K.: Clustering analysis for residential areas based on neighborhood amenities. *Int. J. Adv. Res.* **9**(1), 957–965 (2021)
8. Ansari, K., Bae, T.S.: Clustering analysis of seismicity in the space–time–depth–magnitude domain preceding the 2016 Kumamoto earthquake, Southwestern Japan. *Int. J. Earth Sci.* **110**(1), 253–261 (2021)
9. Turkoglu, B., Uymaz, S.A., Kaya, E.: Clustering analysis through artificial algae algorithm. *Int. J. Mach. Learn. Cybern.* **13**(4), 1179–1196 (2022)
10. Close, L., Kashef, R.: Combining artificial immune system and clustering analysis: a stock market anomaly detection model. *J. Intell. Learn. Syst. Appl.* **12**(4), 83–108 (2020)
11. Akramunnisa, A., Fajriani, F.: K-Means Clustering Analysis pada PersebaranTingkat Pengangguran Kabupaten/Kota di Sulawesi Selatan. *Jurnal Varian* **3**(2), 103–112 (2020). <https://doi.org/10.30812/varian.v3i2.652>
12. Mascandola, C., Barani, S., Massa, M., et al.: Clustering analysis of probabilistic seismic hazard for the selection of ground motion time histories in vast areas. *Bull. Earthq. Eng.* **18**(7), 2985–3004 (2020)



# College Students' Learning Engagement Based on Engineering Thinking

Mingyuan Ma<sup>(✉)</sup>

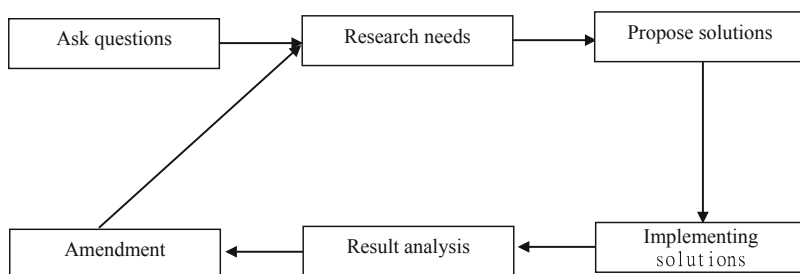
Xi'an Eurasia University, Xi'an 710065, Shaanxi, China  
mamingyuan292@126.com

**Abstract.** One of the cores of engineering thinking is system thinking, which solves problems systematically with scientific theory. Engineering thinking is problem-solving oriented. It conducts research and planning around the problem, determines the requirements and assumptions, proposes solutions and implementation schemes, tests and evaluates the results, and modifies the scheme according to the evaluation results. Using engineering thinking, this paper developed a questionnaire survey of College Students' learning engagement. Through the questionnaire, the relevant data were obtained, and spss24.0 was used to open the black box of how engineering college students learn.

**Keywords:** engineering thinking · student learning engagement · NSSE · CCSS

## 1 Engineering Thinking

One of the cores of engineering thinking is system thinking, using scientific theories to systematize problem-solving. Engineering thinking is oriented to problem solving, and the main process includes the following aspects: researching and planning around the problem, determining requirements and assumptions, proposing solutions, implementing solutions, testing and evaluating the results, and revising the solutions according to the evaluation results [1, 2] (Fig. 1).



**Fig. 1.** Engineering thinking model

Ask questions, how do college students learn? What are the effects of college students' learning experience? The research object is engineering students in a university;



The scale was designed based on the theory of student engagement; Implement the solution, collect data through questionnaire survey, and propose specific measures through expert discussion; Results analysis, analysis of the status quo of students' learning engagement, influencing factors; Revise the solution and evaluate whether the proposed problems have been solved [3–5].

## 2 Theoretical Dimensions and Indicators of Students' Learning Engagement

There are several different understandings of the concept of College Students' learning engagement. The behavioral perspective refers to the time and energy that students devote to learning activities, including students' learning behavior engagement and practical activities; Psychological perspective refers to the process of students' learning psychology, which changes with time, including learning motivation, learning cognition, etc.; Social cultural perspective refers to the influence of a wide range of social environment on students' learning experience, including learning environment, situation, culture and school spirit. From the perspective of school management, it refers to the use of resources and effective educational practice to guide students to do the right thing and direct students' energy to the right activities [4]. According to George Kuh, the concept of learning engagement has two key characteristics: one is the time and energy that students devote to learning and other purposeful educational activities; Second, the teaching resources, related courses and other learning opportunities provided by colleges and universities provide service support for students to participate in learning activities [7]. Based on the reference of ness and CCSS projects, this scale designs the following theoretical framework and index system [6] [7, 8] (Table 1).

**Table 1.** Theoretical dimensions and indicators of students' learning engagement

Learning behavior	Learning behavior before and after class
	Classroom learning behavior
	Learning interaction behavior
learning emotion	Learning motivation
	Resilience learning
	Interest in learning
academic challenge	Higher order learning behavior
	High impact teaching practice
educational environment	School spirit and teaching style Study style
	Learning support and service
learning harvest	Learning satisfaction
	Educational gains

### 3 An Analysis of the Characteristics of Students' Learning Engagement

#### 3.1 Sample Size

Minimum sample size:

$$S = \frac{Z^2 p(1-p)/e^2}{1 + z^2 p(1-p)/e^2 N}$$

$N$  Represents the overall size,  $e = 0.05$  Preset error range,  $z = 1.96$  Significance level,  $p = 0.5$  The proportion of the population with a measurement of 1 (Table 2).

**Table 2.** Comparison of overall scale and minimum sample size

Overall scale	3000	5000	7500	10000	25000	50000
Minimum sample size	341	357	366	370	379	382

The subjects of this study are engineering students, with an overall scale of 8000, and 400 questionnaires are effectively collected. From the above table, we can see the composite sample requirements.

#### 3.2 Reliability and Validity of the Survey Scale

Reliability is a measurement concept, which refers to the credibility of measurement, Cronbach  $\alpha$  The reliability coefficient is a reliability measure about the questionnaire.

Cronbach  $\alpha$  Reliability coefficient:

$$\alpha = \frac{k(\overline{\text{cov}}/\overline{\text{var}})}{1 + (k - 1)(\overline{\text{cov}}/\overline{\text{var}})}$$

among  $k$  The number of items in the questionnaire,  $\overline{\text{cov}}$  Mean value of item covariance,  $\overline{\text{var}}$  Mean value of item variance (Table 3).

**Table 3.** Cronbach reliability test  $\alpha$  coefficient

Dimension of learning engagement	Learning engagement indicators	Cronbach $\alpha$
Learning behavior	Classroom learning behavior before and after class	0.828
	Learning interaction behavior	0.925
Learning emotion	Learning cognition	0.878
Academic challenge	Advanced learning	0.961
	High impact teaching activities	0.866
educational environment	School style, study style and class style	0.931
	Learning support and service	0.951
	Teaching by teachers	0.968
Learning outcomes	Learning gains	0.968
Students' learning engagement scale as a whole		0.973

A Cronbach  $\alpha$  coefficient below 0.7 indicates that the scale needs to be redesigned. Above 0.8 is acceptable and above 0.9 indicates good reliability and high reliability. As can be seen from the above table, the scale passed the reliability test and is of high reliability, indicating that the evaluation questionnaire scale has high internal consistency (Table 4).

**Table 4.** Validity test coefficient of the scale

Kmo and Bartlett's test		
Kmo sampling suitability quantity		0.964
Bartlett sphericity test	Approximate chi square	21149.75
	Degree of freedom	1176
	Significance	0.000

It can be seen from the above figure that the value of kmo is 0.964, indicating that the scale in this questionnaire is suitable for factor analysis. The result of the Bartlett sphericity test at the bottom: the chi square value is 21149.75, which proves that the corresponding p value (0.000) < 0.05, so Bartlett sphericity test is significant.

### 3.3 Current Situation Analysis

Scoring calculation formula:

$$x_i = \frac{\sum d_{ij}}{n_j}$$

According to the statistical content of the questionnaire, the best degree was recorded as 4, good degree was recorded as 3, the moderate degree was recorded as 2, and the worst degree was recorded as 1. According to the formula, calculate the average score of each option as the score of each index.  $x_i$  Represents the average score of the index,  $d_{ij}$  Indicates the score of the option,  $n_j$  Indicates the number of options [9–11]. (Table 5).

**Table 5.** Basic data of each dimension of learning engagement

		average value	standard deviation
1. Learning behavior	1.1 learning behavior before and after class	2.94	0.61
	2.1 interactive learning behavior	2.84	0.81
2. Learning emotional cognition	2.1 learning cognition	3.12	0.63
3. Academic challenges	3.1 higher level learning	3.05	0.72
	3.2 high impact teaching practice	2.96	0.58
4. Educational environment	4.1 school spirit, class style and study style	3.08	0.69
	4.2 learning support and service	3.13	0.68
	4.3 teacher teaching	3.28	0.64
5. Learning effectiveness	5 learning gains	3.00	0.68

Full score is 4, The highest score of teacher teaching is 3.28, which shows that students are very satisfied with teacher teaching. The scores of learning cognition, advanced learning, school spirit, class atmosphere, learning support and service, learning harvest were all higher than 3, and the other indicators were close to 3, indicating that the evaluation results of the above indicators were in an excellent state.

### 3.4 Correlation Analysis

Pearson correlation coefficient:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

among  $\bar{x}$ ,  $\bar{y}$  They are variables  $x$ ,  $y$  Mean value,  $x_i$ ,  $y_i$  They are variables  $x$ ,  $y$  Observation value,  $n$  It's the number of respondents. Spss24.0 was used to analyze the correlation of questionnaire data (Table 6).

**Table 6.** Correlation coefficient of each index

	1.1	1.2	2.1	3.1	3.2	4.1	4.2	4.3	5
1.1 learning behavior before and after class	1								
1.2 interactive learning behavior	.654**	1							
2.1 learning cognition	.683**	.717**	1						
3.1 higher level learning	.671**	.735**	.714**	1					
3.2 high impact teaching practice	.569**	.621**	.637**	.680**	1				
4.1 school spirit, class style and study style	.578**	.632**	.595**	.695**	.629**	1			
4.2 learning support and service	.557**	.626**	.619**	.733**	.652**	.757**	1		
4.3 teacher teaching	.548**	.528**	.552**	.603**	.524**	.688**	.757**	1	
5 learning gains	.609**	.702**	.690**	.793**	.743**	.690**	.751**	.643**	1

\*\*Indicates a significant correlation at the 0.01 level, \*Indicates a significant correlation at the 0.05 level,

Note: the contents in the first row are the same as those in the first column. For example, line 1.1 indicates the classroom learning behavior before and after class. The correlation coefficient range is generally [-1, +1] The greater the absolute value, the higher the correlation. [0.2, 0.4] Weak correlation, [0.4, 0.6] Moderate correlation, [0.6, 0.8] Strong correlation. From the correlation coefficient, we can see that there is a significant positive correlation between the indicators. Among them, the highest correlation between higher-order learning and learning harvest is 0.793, among which learning interaction behavior, high-level learning and high impact teaching practice have the most prominent impact on learning harvest.

### 3.5 Establishment of Regression Equation

#### Multiple Linear Regression Model

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n$$

Y: dependent variable,  $X_n$ : No  $n$  Independent variables,  $a$ : constant term,  $b_n$ : No  $n$  Regression coefficients,

Taking learning gains as dependent variables and learning behaviors, learning emotions, academic challenges, teachers’ teaching and campus environment as independent variables, the regression equation is established. Spss24.0 was used to analyze the influencing factors and calculate the regression coefficient (Table 7).

**Table 7.** Modified linear regression coefficients

	Non standardized coefficient		Standardization coefficient	t	Significance	Collinear statistics	
	B	Standard error	Beta			tolerance	VIF
(constant)	-0.141	0.106		-1.328	0.185		
2.1 interactive learning behavior	0.091	0.034	0.108	2.632	0.009	0.374	2.677
2.1 learning cognition	0.075	0.044	0.069	1.700	0.030	0.382	2.617
3.1 higher level learning	0.281	0.044	0.297	6.417	0.000	0.297	3.367
3.2 high impact teaching practice	0.308	0.044	0.264	7.034	0.000	0.450	2.223
4.2 learning support and service	0.178	0.047	0.178	3.794	0.000	0.287	3.486
4.3 teacher teaching	0.100	0.041	0.095	2.422	0.016	0.415	2.410

Significant (P value) = <0.05, The fitting effect of regression equation has passed the test and has significant statistical significance.

### Multiple Linear Regression Equation:

Learning gains =  $-0.141 + 0.091 * \text{learning interaction behavior} + 0.075 * \text{learning cognition} + 0.281 * \text{higher level learning} + 0.308 * \text{high impact teaching practice} + 0.178 * \text{learning support and service} + 0.1 * \text{teacher teaching}$ .

The larger the regression coefficient is, the higher the influence degree of independent variable is. From high to low, the top three indicators are: high impact teaching practice, high-level learning, learning support and service.

## 4 Conclusion

### 4.1 Open the Black Box of Learning

This study clearly shows how students learn, their learning experience and learning outcome, and the influencing factors of learning gains. Learning behavior includes preview

before class, taking notes in class, interaction in class, review and discussion after class, interaction between students and students, and interaction between teachers and students; Learning emotion includes learning cognition, learning goal and learning sense; High impact teaching practice includes participation in subject competition, teaching practice, project teaching, etc.; The learning experience includes the quality of interaction between students and their classmates, roommates and teachers, as well as the support they get in the learning process; Learning gains include the growth of students' knowledge, ability and accomplishment.

## 4.2 Existing Problems

This survey found that there are still some students who are deficient in learning goals, sense of learning meaning, learning resilience, the frequency and quality of learning interaction. Therefore, teachers and teaching administrators should guide students to realize their main responsibility in learning and realize their irreplaceable position and role in self-learning and development. Institutions and teachers should study students and learning, create a positive learning environment while helping students solve problems that arise at all stages of the learning and development process [12–15].

## 4.3 Recommendations

The survey found that the three most important factors affecting students' learning achievement are: high impact teaching practice, high-level learning, learning support and service. High impact teaching practice includes subject competition, project-based learning, practice project, research with teachers, etc. Higher order learning is not to know and memorize, but to apply and evaluate, to solve practical problems with the knowledge acquired. Learning support and service refer to the support provided by schools, teachers and administrators when students encounter difficulties in learning.

## 4.4 Inadequate Research

Learning gains have hidden characteristics, and some effects should take a long time to show up. Therefore, the data collected by the questionnaire has certain limitations, but it does not affect the conclusions of this paper.

**Acknowledgments.** Project Name: Study on Countermeasures to Improve College Students' Learning Engagement from the Perspective of NSEE—Taking Xi'an Eurasian University as an Example (Project No. 2021YB022).

Project name, research on the demand for physical fitness services for the elderly population (project number: 2022125).

## References

1. Yurtaeva, L.V., Marchenko, R.A., Kaplyov, E.V., et al.: Development of engineering thinking with the use of research activities in the field of plant raw materials deep processing. *J. Phys. Conf. Ser.* **1691**, 012182 (2020)

2. Kuh, G.D.: What we're learning about student engagement from NSSE: benchmarks for effective educational practices. *Change Magaz. High. Learn.* **35**(2), 24–32 (2003)
3. Can, L.I., Lang, H.U., Lei, Z.X.: A study of influence factors about college students' learning engagement in local colleges. *High. Educ. Sci.* (2015)
4. Drewery, D.W., Stewart, N.W., Wilson, A.W.: Engagement in mindfulness exercises during large lectures and students' writing self-efficacy. *Can. J. Scholarship Teach. Learn.* **13** (2022)
5. Du, C., Yu, C., Wang T., et al.: Impact of Virtual Imaging Technology on Film and Television Production Education of College Students Based on Deep Learning and Internet of Things (2022)
6. Blakeslee, J., Miller, R., Uretsky, M., et al.: Efficacy of the Project Futures self-determination coaching model for college students with foster care backgrounds and mental health challenges. *Children Youth Serv. Rev.* (2022)
7. Callender, C., Melis, G.: The privilege of choice: how prospective college students' financial concerns influence their choice of higher education institution and subject of study in England. *J. High. Educ.* **93** (2022)
8. Kadir, Z.A., Mohamad, F., Rathi, N., et al.: The perceived effectiveness of student engagement strategies in open and distance learning. *Int. J. Asian Soc. Sci.* **12** (2022)
9. Pearson, W.S.: Student engagement with teacher written feedback on rehearsal essays undertaken in preparation for IELTS. *SAGE Open* **12**(1), 1–20 (2022)
10. Sun, Y., Tu, S., Zhou, L.: The effect, deficiency and thinking of the implementation of the excellent engineer program – Based on the empirical investigation of 7 majors in East China University of science and technology. *High. Eng. Educ. Res.* **2021**(1), 7 (2021)
11. Xu, B.: Center for teaching and learning: an institution for improving the quality of American University Education -- a case study of Stanford University. *Comparat. Educ. Res.* **2021**(2013–5), 74–78 (2021)
12. Jin, M., Hu, S.: Gender differences in learning engagement of engineering undergraduates. *Fudan. Educ. Forum* **16**(5), 9 (2018)
13. Zhao, J., Gao, X.: On the implementation of “student-centered” Undergraduate Teaching Reform. *China High. Educ. Res.* **2017**(8), 5 (2017)
14. Zhang, X.: Countermeasures for optimizing teaching evaluation in Colleges and universities with learning as the center. *China Adult Educ.* **2021**(20), 4 (2021)
15. Chang, T.: A comparative study on learning expectation and students' learning engagement in undergraduate courses between China and the United States. *China High. Educ. Res.* **2019**(4), 10 (2019)





# Comparison and Research of Roof Waterproof Materials Based on Computer Algorithm

Xinjun Wang<sup>(✉)</sup>

School of Urban Construction Chengdu Polytechnic, Chengdu 610000, Sichuan, China  
maolv505@126.com

**Abstract.** Comparison of roof waterproof materials based on computer algorithm: use computer algorithm to compare various roof waterproof materials. This is the most common way to compare different roof materials. The results obtained from this method are very accurate and reliable because it uses advanced technology for these comparisons. How to compare and study roof waterproof materials based on computer algorithm? The comparison process of different types of roofs involves a series of steps, including: before online purchase, you can compare the prices, functions, specifications and other details of different products. This comparison can help you find out which product is better than others. It also provides information about how much you need to spend to get the product you need.

**Keywords:** Computer algorithm · Waterproof materials · Data analysis

## 1 Introduction

Roof waterproof and leak proof is the most concerned and headache problem in the society at present. It is related to the vital interests of the people and the durability of buildings. It is one of the most difficult and troublesome problems in construction engineering. Generally, it takes a period of time for a new house to leak, and sometimes early leakage occurs due to the influence of various reasons. At this time, it needs continuous maintenance and renovation. With a series of maintenance work, people have a sense of disgust. Some people living on the top floor are disgusted when it rains. People's disgust has aroused great attention in the engineering and technical circles, and is working hard to innovate and improve the durability of roof waterproof materials.

Roof structure is the most basic building structure in industrial and civil building structures. At present, according to the design reference period, the roof structure is in the process of repeated maintenance and transformation. Due to the reasons of design, construction and use, their safety, applicability and durability are reduced. In order to prolong the service life of the structure and realize the good maintenance of the building, it is necessary to formulate the existing roof renovation and maintenance procedures and explore the advanced materials and methods of structural maintenance and renovation [1]. With the arrival of the new century, there are many changes in the use of materials and methods in the fields of waterproof, leak proof maintenance and repair of roof layer. New materials are increasing and gradually replacing traditional materials. In order to

make the roof waterproof material durable, we should first analyze its causes and take measures for better application.

This paper studies the comparison of roof waterproof materials based on computer algorithm. According to the nature, importance, service life, functional requirements and service life of the building, the roof waterproof is divided into two levels and fortified according to different levels [2]. The common types of waterproof roofs are coiled waterproof roof, coated waterproof roof and rigid waterproof roof.

## 2 Related Work

### 2.1 Computer Algorithm – Python

For computer science, the concept of algorithm is very important. Algorithm is a series of clear instructions to solve the problem, that is, it can obtain the required output in a limited time for a certain standard input. If an algorithm is defective or not suitable for a problem, executing the algorithm will not solve the problem. Different algorithms may be used to complete the same task in the same space or time. The advantages and disadvantages of an algorithm can be measured by space complexity and time complexity.

The algorithm can be understood as a complete problem-solving step composed of basic operation and specified operation sequence. Or as a finite exact calculation sequence designed according to the requirements, and such steps and sequences can solve a class of problems. The algorithm can be described by many different methods, such as natural language, pseudo code, flow chart and so on. The complexity of the algorithm is a measure of the efficiency of the algorithm and an important basis for evaluating the advantages and disadvantages of the algorithm. The complexity of an algorithm is reflected in the amount of computer resources required to run the algorithm. The more resources required, the higher the complexity of the algorithm; On the contrary, the lower the required resources, the lower the complexity of the algorithm.

$$e_k(t) = y_d(t) - y_k(t) \quad (1)$$

$$\|e_{k+1}(t)\| = \|C(t)\| \|\Delta x_{k+1}(t)\| \quad (2)$$

Computer resources, the most important is time and space (i.e. memory) resources. Therefore, the complexity of the algorithm can be divided into time complexity and space complexity.

It goes without saying that for any given problem, designing an algorithm with as much complexity as possible is an important goal we pursue in designing an algorithm; On the other hand, when there are many algorithms for a given problem, choosing the one with the lowest complexity is an important criterion we follow in selecting algorithm adaptation [3]. Therefore, the complexity analysis of the algorithm has important guiding significance and practical value for the design or selection of the algorithm.

As for the complexity of algorithm, there are two problems to be clarified: what kind of quantity is used to express the complexity of an algorithm; For a given algorithm, how to calculate its complexity.

## 2.2 Roof Waterproof Materials

At present, the mainstream roof waterproof engineering generally adopts waterproof coiled material construction. There are three series of waterproof coiled materials used: traditional asphalt waterproof coiled materials, polymer modified asphalt waterproof coiled materials and synthetic polymer waterproof coiled materials. The traditional asphalt waterproof coiled material is rarely used now, unless it is used for the parts with low waterproof requirements. At present, the coiled materials used in waterproof construction are mainly two typical high polymer modified asphalt waterproof coiled materials app and SBS.

These two kinds of waterproof rolls are high polymer modified asphalt waterproof rolls, but their modifiers are different: the modifier of SBS waterproof roll is SBS (styrene butadiene styrene); APP waterproof coiled material modifier is app (random polypropylene).

Due to the difference of modifiers, the construction scope of the two coiled materials is also different. SBS waterproof coiled materials can be used in the temperature range of  $-25$  to  $+100$  °C, with high elasticity and fatigue resistance. They are especially suitable for building waterproof projects in low-temperature and cold areas and frequent structural deformation [4].

APP waterproof coiled material has higher aging resistance, high temperature resistance and other properties, especially suitable for building waterproof construction in hot areas and waterproof construction with special requirements for heat resistance. Because SBS elastomer coil has good flexibility, waterproof and anti-aging properties, and has high temperature not flowing, low temperature is not brittle crack, simple construction, no pollution, long service life and so on. Compared with app, plastic coiled material is more widely used, and it is also the most common coiled material on the construction site. As shown in Fig. 1 below, it is asphalt waterproof coiled material.



**Fig. 1.** Asphalt waterproof coiled material

There are two common cases of back water surface waterproof, one is chemical grouting. Generally, foamed polyurethane is used. This material is simple, easy to operate, fast water stop and inexpensive (do not find one priced according to the weight of the material). Note that the construction must be carried out under the condition of water, and the effect is obvious [5]. However, this material has water expansion and anhydrous shrinkage, and cracks and re leakage will occur under repeated. Epoxy material is ideal, but the requirements for construction machines and materials are particularly high, which is rare in household waterproof; One is a backwater technology, which uses inorganic and organic materials to repair the self waterproof function of the structure, with reliable effect and long working time. However, there are strict requirements for construction and materials, and few people can master this technology.

### **3 Comparison and Research of Roof Waterproof Materials Based on Computer Algorithm**

The comparison of roof waterproof materials based on computer algorithm is a very important aspect in the field of building construction. It helps us choose the best materials for our buildings and also helps us compare different types of roof materials. The basic idea behind this comparison is that all these roofs are made of similar materials, but there are some differences in design and characteristics. In order to correctly determine which type of roof should be used in our buildings, we need to know how these roofs behave under certain conditions [6].

Import the data into the computer algorithm, such as rigid waterproof roof: it refers to the roof with rigid waterproof material as waterproof layer. The roof is mainly made of waterproof sheet, waterproof sheet metal and waterproof sheet metal.

Waterproof concrete: waterproof concrete, also known as structural self waterproof, can be adjusted by adjusting the mix proportion of concrete or adding admixtures Steel fiber, synthetic fiber, etc., and cooperate with strict construction and construction management to reduce the void ratio in the concrete or change the pore shape and distribution characteristics, so as to achieve the purpose of waterproof (anti-seepage).

Roof tile: the roof tile is divided into concrete roof tile, asphalt felt tile and PVC tile. Concrete roof tiles are made of concrete; The asphalt felt tile is made of glass fiber felt as the tire base. After being impregnated with petroleum asphalt, one side is covered with colored mineral aggregate and the other - side is covered with isolation material; Polyvinyl chloride tile is made of polyvinyl chloride resin (PVC) as the main raw material and other chemical raw materials. After scientific proportion, it is made of a full set of advanced technology of three-layer co extrusion composite micro foaming [7].

Metal roof: metal roof refers to the roof form that uses metal plate as the roof material and combines the structural layer and waterproof layer into one. There are many kinds of metal plates, including zinc plate, aluminum zinc plate, aluminum alloy plate, aluminum magnesium alloy plate and titanium alloy plate Copper plate, stainless steel plate, etc.

Lead alloy waterproof roll: it is a kind of formula of lead, zinc, tin and other metal alloys as the main material, which is melted The waterproofing membrane made of alloying elements, proportioning, heat preservation and casting rolling has good extension, solderability, stable performance, X ray resistance and anti-aging ability.

Sunshine plate; It is a transparent reinforced hollow board or solid board formed by extrusion process with polycarbonate plastic (PC) as raw material. It is a new type of roof material with high strength, waterproof, light transmission and energy saving [7].

Membrane structure waterproof roof: membrane structure is a new type of building structure developed in the middle of the 20th century. It is a new type of building roof material with a sense of the times, which is made of a variety of high-strength membrane materials (PVC or Teflon) and reinforcing members (steel frame, steel column or steel cable) to generate a certain amount of pretension stress inside in a certain way to form a certain spatial shape. It is both a waterproof material and a roof structure.

## 4 Selection of Waterproof Materials

There are two kinds of waterproof materials on the market: polyurethane waterproof coating and polymer cement-based waterproof coating.

Polyurethane waterproof coatings made of polyurethane and coal tar as raw materials have been banned because the volatile tar gas is highly toxic and difficult to remove. What is still on sale in the market is polyurethane waterproof coatings made of asphalt instead of coal tar as raw materials. It is generally used for parts with low waterproof requirements [7].

Polymer cement based waterproofing coating is composed of a variety of waterborne polymers, Q, and high quality cement mixed with various additives. The flexibility of polymer (resin) is integrated with the rigidity of cement, which makes it excellent in impermeability and stability. It has the advantages of convenient construction, low comprehensive cost, short construction period, non-toxic and environmental protection. At present, polymer based coatings have been widely used.

TPO waterproof coiled material combines the performance advantages of EPDM and PVC, and has the weather resistance of EPDM, low temperature flexibility and easy welding characteristics of PVC. After adding a layer of polyester fiber fabric between the two layers of TPO materials, the physical properties, breaking strength and puncture resistance are enhanced [8]. In practical application, TPO has the characteristics of anti-aging, high tensile strength, 9 elongation, wet roof construction, no need to protect the exposed layer, convenient construction and no pollution. It is very suitable for the waterproof layer of steel roof.

## 5 Conclusion

The comparison and research of roof waterproof materials based on computer algorithm is the process of comparing different roof materials to find out which one has better performance in waterproof penetration. The comparison is done by using a computer program that calculates the amount of water that can penetrate into the material under specific conditions. This test helps you understand how much rain the roof will retain before it begins to leak. The purpose behind this method is to compare different types of roofs and their performance with rainwater infiltration.

## References

1. Inwan, M., Thuwapanichayanan, R., Sayasoonthorn, S.: Comparison of Papaya Cushioning Materials by Ellipsoid Evaluation Method (2021)
2. Sobakin, A.A , Nikolaeva, D.A., Burtsev, R.M.: Comparison of curved roofs of various configurations. *IOP Conf. Ser. Earth Environ. Sci.* **988**(5), 052017 (6pp) (2022)
3. Fritschen, A., Bell, A.K., Stark R.W., et al. Investigation and comparison of resin materials in transparent DLP-printing for application in cell culture and organs-on-a-chip (2022)
4. Kai, S.E.: Comparison of the conventional volcano analysis with a unifying approach: material screening based on a combination of experiment and theory. *J. Phys. Chem. C* **124**(1), 822–828 (2020)
5. Zhou, Y., Liu, D., Su, M., et al.: Comparison of fluorine sources on the electrochemical property of  $\text{Li}_{1.2}\text{Ni}_{0.2}\text{Mn}_{0.6}\text{O}_2$  cathode materials. *Funct. Mater. Let.* **13**(5), 2050027 (2020)
6. Mathe, C.: Protocol comparison for organic residue analyses from waterproofing materials and shards of roman archaeological amphorae. *Crystals* **11** (2021)
7. Papaioannou, C., Katsoulas, N., Kitta, E.: Reduction in blockage property of uv-blocking greenhouse covering material: In Situ and Lab Measurement Comparison (2022)
8. Humphreys, D.C., Shifferd, M.: Peel and stick waterproofing material:, US10857759B2[P] (2020)



# Comprehensive Evaluation System of Cultural Industry Development Based on Computer-Aided

Jianghong Yuan<sup>(✉)</sup>

Henan University, Henan 745001, China

11712970@qq.com

**Abstract.** The research on the comprehensive evaluation system of cultural industry development based on computer-aided is a research that provides a basis for the development and expansion of China's cultural industry. The study also provides new methods and technologies that can be used by various enterprises in the cultural industry sector. It aims to provide an efficient and reliable method for evaluating the effectiveness of the cultural industry. The main problem we encountered in this process is that there is insufficient research and literature on this subject, so it is difficult to find what should be evaluated. However, after a period of time, we found that in order to make a good assessment, we need to assess many aspects: 1) social impact (positive or negative), 2) economic impact (positive).

**Keywords:** computer-aided · Cultural industry development · Comprehensive evaluation system

## 1 Introduction

With the development of the Internet, the comprehensive contribution of tourism to GDP has not only increased year by year, accounting for more than 10% of the total GDP, but also the products and formats in the industry are gradually enriched. It can be seen that tourism plays a more prominent role in economic development. Since the reform and opening up, with the growth of China's economy and national income, tourism is no longer just the enjoyment of specific classes and a few people, and has gradually become the normal daily life of the general public.

Liushengyi, senior executive vice president of Tencent group, and Liam Fox, British Secretary of state for international trade, signed a memorandum of strategic cooperation in London, announcing that Tencent has carried out cooperation in the field of digital culture and innovation with a number of well-known enterprises and institutions, including the British Tourism Administration and the BBC. After Tencent put forward the "new cultural and creative" strategy, what does this cultural and creative cooperation in Britain mean for Tencent and the development of the whole Chinese cultural entrepreneurship field?

“The United States and the United Kingdom are countries with developed creative industries in the world, and they have advanced experience in the development of creative industries, which is worth learning and using for reference by China.” tongqingyan, Professor of the school of media and communication, Shanghai Jiaotong University, and director of the media literacy research center [1], believes that “at present, it is the best time for China to invest in the United Kingdom for mutual benefit and win-win results, and it is also a powerful opportunity for China’s cultural and creative industries to go global.” She talked about her views on new cultural and creative industries and the opportunities for China’s cultural and creative development brought by cooperation with the UK from four dimensions: the relationship between cultural and creative industries and people, the composition of cultural and creative economy, the fields that cultural and creative industries span and China’s new model of cultural and creative globalization [2].

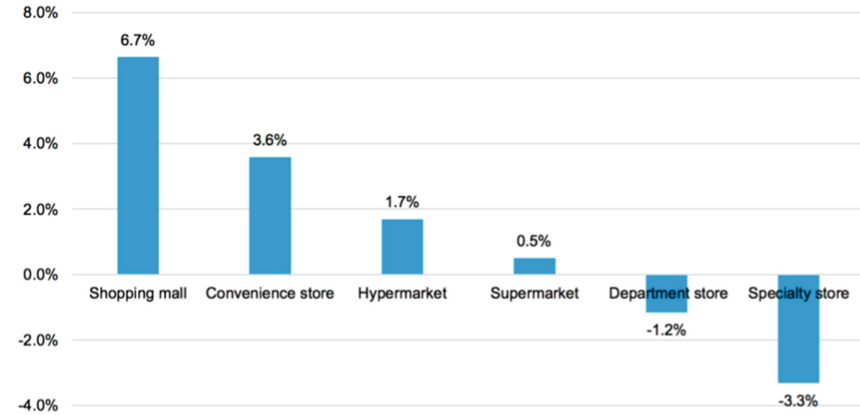
How to make Chinese cultural and creative products not only face the Chinese market, but also face the world? Make advanced and advanced products? Make it a work that integrates the global creative production concept, collides the wonderful Eastern and Western cultures and the current global trend, and comprehensively releases the influence of China’s new generation of cultural creativity? Based on this, this paper studies the comprehensive evaluation system of cultural industry development based on computer-aided.

## 2 Related Work

### 2.1 Development Status and Trend of Cultural Industry

The performance of the cultural industry in recent years can be described as “ice and fire”. In 2018, under the situation of “fighting against the four evils”, 60% of the performance market and stage production companies closed down, and the art and gallery industries continued to be depressed; At the same time, due to the homogenization, real estate and shell development of the cultural industry in recent years, 70%–80% of the film and television bases, cultural industry parks and theme parks nationwide are at a large loss; On the other hand, some of China’s cultural industries are showing a trend of rapid growth. In particular, the cultural industry has made extraordinary development by using the capital market. Cultural enterprises have made outstanding performance in the gem and become the “new favorite” of the capital market [3]. At the same time, due to the small scale of cultural enterprises on the growth enterprise market before listing, they have the motivation to actively use the capital market to seek the extension expansion of enterprises, carry out industrial integration and improve the overall level of enterprises, so as to change the competitive situation in the past. By the end of 2021, there has been a merger and acquisition in the market every six days. At the same time, with the participation of leading Internet enterprises such as Alibaba, Baidu and Tencent, under the joint efforts of policy, capital and market, the cultural industry is undergoing a qualitative and quantitative leap, that is, from blindly imitating foreign countries to using their own national culture to develop towards independent innovation. Figure 1 below shows the development trend of the cultural industry.





Data source: CCFA, Deloitte Analytics

**Fig. 1.** Development trend of cultural industry

The current situation of the cultural industry shows that under the conditions of low level, small enterprise scale and lack of market competition, China's cultural industry is facing huge market needs and major adjustment opportunities under the new situation [4]. The basic situation of these cultural industries is also the basis for the development of the cultural industry.

By industry type, the added value of cultural manufacturing industry in 2019 increased by 7.6% over the previous year, accounting for 38.6% of the added value of cultural and related industries; The added value of cultural wholesale and retail increased by 13.0%, accounting for 9.3%; Among them, the added value of cultural service industry increased by 17.5%, accounting for 52.1%, which has become the main force to promote the development of cultural industry. In addition, in 2019, the per capita consumption expenditure of national residents for culture and entertainment was 800 yuan, an increase of 38.7% over 2018; The cultural and entertainment expenditure belonging to the cultural service industry accounted for 4.7% of the total consumption expenditure, higher than the level of 4.4% in 2018. The structure of the cultural industry was further optimized and upgraded.

## 2.2 Policy Advantages of Artificial Intelligence + Cultural Industry

In made in China 2025 issued by the State Council in 2015, "intelligent manufacturing" was positioned as the main direction of manufacturing in China. On July 5, 2015, the "Internet +" action guidance issued by the State Council also proposed to vigorously develop intelligent manufacturing [5].

In April, 2016, the Ministry of industry and information technology, the national development and Reform Commission and the Ministry of Finance jointly issued the robot industry development plan (2016–2020). In May 2016, the national development

and Reform Commission, the Ministry of science and technology, the Ministry of industry and information technology and the Internet Information Office jointly issued the three-year action implementation plan for “Internet +” artificial intelligence.

In July 2016, the State Council issued the 13th five year plan for national science and technology innovation, which mentioned intelligent manufacturing and robotics in the “major project of science and technology innovation 20309”. With the overall goal of intelligent, efficient, collaborative, green and safe development, we will build a network collaborative manufacturing platform, develop intelligent robots, high-end complete sets of equipment, three-dimensional (3D) printing and other equipment, and consolidate the basic support capacity of manufacturing.

In December 2016, the State Council issued the 13th five year plan for the development of national strategic emerging industries, which requires the development of artificial intelligence, the cultivation of artificial intelligence industry ecology, the promotion and application of artificial intelligence in key economic and social fields, and the creation of an internationally leading technology system.

By March 2017, “artificial intelligence” had been written into the national government work report for the first time.

In April, the Ministry of Culture issued the guiding opinions on promoting the innovative development of the digital cultural industry. Subsequently, the Ministry of Culture issued the development plan of the cultural industry during the 13th Five Year Plan period of the Ministry of culture, proposing to vigorously cultivate new cultural industries based on new technologies such as big data, cloud computing, Internet of things, artificial intelligence, etc., and form new growth points of the cultural industry; In July, the State Council issued the new generation of artificial intelligence development plan I [6]. by 2030, the competitiveness of China’s artificial intelligence industry will reach the international leading level. The scale of the core artificial intelligence industry will exceed 1 trillion yuan, driving the scale of related industries to exceed 10 trillion yuan. The introduction of a series of policies shows that China attaches importance to the field of artificial intelligence and the relevant development prospects.

Needless to say, the continuously evolving and upgraded artificial intelligence is also bound to have a greater and greater impact on the cultural industry. This also proves the further expansion of the impact of artificial intelligence on people’s lives. It is expected that there will be more places to use artificial intelligence in people’s lives in the near future, from the current TV voice recognition, face recognition, smart home system to more technologies that will facilitate people’s lives in the future [7].

As a pioneer in the field of artificial intelligence as early as 2012, Changhong has made considerable progress in the development and application of sensor modules, speech recognition chips, artificial intelligence technology based on big data, machine vision (face recognition) technology, industrial robots and factory automation technology, and home service robots, And there are many achievements transformation and commercial application.

### 3 Research on Comprehensive Evaluation System of Cultural Industry Development Based on Computer-Aided

Generally, the difference between the predicted value of the cultural industry development model evaluation and the value of the sample itself is called error. When machine learning trains training data, the error of cultural industry development model evaluation on the training data is called training error, also known as experience error. When we use our own cultural industry development model to predict a new data set, the error on the new data set is called generalization error. In an ideal state, we hope to get an evaluation of the cultural industry development model with the generalization error  $Q$  as small as possible, or even 0, which is equivalent to a prophet. Obviously, we do not know the real output of the data to be predicted when modeling, so the generalization error is impossible. Taking a step back, we can make the training error as small as possible. With the increase of the complexity of the evaluation of the cultural industry development model and the increase of the training times, we can even do this [8]. However, such cultural industry development model evaluation is only nearly perfect in the training set, but it is often unsatisfactory when it is really used in the new data set, which is called over fitting. Now, for over fitting, the methods to reduce the impact of over fitting basically include regularization, reducing the input parameters of cultural industry development model evaluation, and so on.

$$x_k(0) = x_d(0), k = 0, 1, 2, 3, \dots \tag{1}$$

$$\Delta \dot{e}_k(t) = \dot{e}_{k-1}(t) - \dot{e}_k(t), \Delta \dot{e}_{k+1}(t) = \dot{e}_k(t) - \dot{e}_{k+1}(t) \tag{2}$$

Since we don't know the real cultural industry data, we take some cultural industry data from the training set as the test set, and the rest as the training set to train the model. For the test set, after applying the model, we can get a test error  $Q$ , and we use the test error to estimate the generalization error. The test set and training set ° here are mutually exclusive, that is, the trained cultural industry data cannot be used as a test

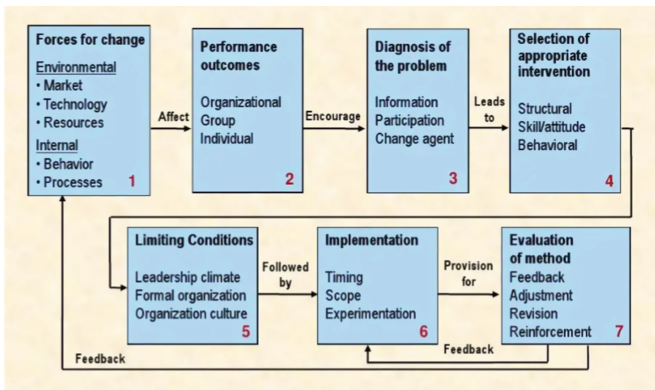


Fig. 2. Data forecast process of cultural industry

set. Obviously, such cultural industry data will be predicted accurately. Figure 2 below shows the data prediction process of the cultural industry.

Of course, by default, the division method in sklearn is random. This method is not a problem when the amount of cultural industry data is very large [9]. However, when there are not many cultural industry data sets, stratified sampling is still required to ensure that the proportion of various samples in  $S$  and  $t$  is consistent with that in the original cultural industry data set  $D$ . such a result is more representative. This can be achieved by using the stratify in sklearn.

About test\_ In theory, we should use the cultural industry data in  $D$  to model the selection of size. Therefore, the larger the proportion of the training set, the closer the model will be built. However, at this time, it will appear that there are too few cultural industry data in the test set, and the test results are not universal. Therefore, it needs to be selected according to the actual situation. Generally, about 20% of the cultural industry data will be selected as the test set [10].

In fact, the effect of dividing cultural industry data sets by setting aside method is closely related to the selection of test sets. In order to reduce its impact, we can choose to divide cultural industry data sets several times and average the final results.

## 4 Conclusion

Industrialization is the development of cultural activities, which has a great impact on social and economic development. This paper mainly introduces the computer-aided comprehensive evaluation system of cultural industry, including cultural content analysis, output value analysis and market research. The process is divided into four stages: (1) data collection; (2) Data processing; (3) Presentation and interpretation of data; (4) Policy decision-making recommendations. However, no research has been done in the field of culture and tourism development. The purpose of this study is to establish a cultural industry development evaluation system based on computer-aided technology. The purpose of this study is to find out how to use computer-aided technology to evaluate the performance of the cultural industry.

## References

1. Qin, W., Lin, Q.: Research on comprehensive evaluation method of Bayu cultural resources development based on DEMATEL. *Int. J. Electric. Eng. Educ.* **2021**,002072092199659 (2021)
2. Chen, F., Song, Q., Hua, L.I.: Sustainable development of wannian rice cultivation culture system. **13**(5), 6 (2021)
3. Gao, X., Guan, Y., Sun, D., et al.: An ecological, power lean, comprehensive marketing evaluation system based on DEMATEL–CRITIC and VIKOR: a case study of power users in northeast China. *Energies* **15** (2022)
4. Wang, W., Jia, X.: Establishing an Evaluation System for Interdisciplinary Graduate Students Based on Improved AHP (2021)
5. Ding, Y., Sun, Y., Wang, F.: A corpus-based analysis of news on cultural industry system innovation in shaanxi pilot free trade zone. *Open Access Library J.* **9**(5), 16 (2022)
6. Qin, W., Lin, Q.: Construction of cultural industry development factor model based on factor analysis, artificial intelligence and big data - ScienceDirect. *Microprocess. Microsyst.* (2021)

7. Krajnik, D., Li, M.S.: An Analysis and Evaluation Methodology as a Basis for the Sustainable Development Strategy of Small Historic Towns: The Cultural Landscape of the Settlement of Lubenice on the Island of Cres in Croatia (2022)
8. Liu, W., Sun, M., Han, K., et al.: Comprehensive evaluation of stable neuronal cell adhesion and culture on one-step modified polydimethylsiloxane using functionalized pluronic. *ACS Omega* **5**(50) (2020)
9. Rui, L., Liu, Z.: Evaluation of Coordination between Cultural Tourism and Urban Economy of Gansu 02015 (2021)
10. Zhang, X., Liu, J., Chen, Q.: Comprehensive evaluation of jianyang city's economic high-quality development level based on entropy method. In: 2020 2nd International Conference on Economic Management and Cultural Industry (ICEMCI2020) (2020)



# Correction to: Innovative Computing Vol 1 - Emerging Topics in Artificial Intelligence

Jason C. Hung, Jia-Wei Chang, and Yan Pei

## Correction to:

**J. C. Hung et al. (Eds.): *Innovative Computing Vol 1 - Emerging Topics in Artificial Intelligence*, LNEE 1044,  
<https://doi.org/10.1007/978-981-99-2092-1>**

The affiliation for the author “Hailin Tang” of Chapter 64 “Double Attention Mechanism Text Detection and Recognition Based on Neural Network Algorithm” has been updated as “Mongolian University of Science and Technology, Bayanzurkh District, Ulaanbaatar 13341, Mongolia”.

The affiliation for the author “Yongjun Qi” of Chapter 65 “A Text Detection and Recognition System Based on Dual-Attention Mechanism with Artificial Intelligence Technology” has been updated as “Mongolian University of Science and Technology, Bayanzurkh District, Ulaanbaatar 13341, Mongolia”.

The chapters and the book have been updated with the changes.

---

The updated original version of these chapters can be found at  
[https://doi.org/10.1007/978-981-99-2092-1\\_64](https://doi.org/10.1007/978-981-99-2092-1_64)  
[https://doi.org/10.1007/978-981-99-2092-1\\_65](https://doi.org/10.1007/978-981-99-2092-1_65)

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023  
J. C. Hung et al. (Eds.): IC 2023, LNEE 1044, p. C1, 2023.  
[https://doi.org/10.1007/978-981-99-2092-1\\_131](https://doi.org/10.1007/978-981-99-2092-1_131)

# Author Index

## A

Aer, Yougui 937  
An, Xin 168  
An, Younghoon 3

## B

Bai, Jing 784  
Ben, Nenjun 890  
Bian, Hengran 471  
Bo, Zonghui 448

## C

Chai, Xiaokang 363  
Chen, Jinfeng 951  
Chen, Jinhua 269, 280  
Chen, Jinxia 896  
Chen, Rui 408  
Chen, Sitong 372  
Chen, Xi 448  
Chen, Xiaoguang 800  
Cheng, Hao 396  
Chi, Jingyuan 37  
Choi, Sujeong 26  
Chu, Yongling 262  
Cong, Rijie 831  
Cong, Yuyue 315  
Cui, Lei 571  
Cui, Xinhe 168

## D

Dai, Dangdang 597  
Dai, Hongcai 408  
Dai, Jing 555  
Ding, Jie 479, 564  
Ding, Wenting 417  
Dong, Chenxi 555  
Dong, Jie 402, 571  
Du, Liming 571  
Du, Lingpan 944

## F

Fan, Jie 805  
Fan, Yangyang 448  
Fan, Yuxiao 448  
Feng, Chao 841, 937  
Fu, Lijun 37  
Fu, Yi 659

## G

Gao, Liang 120  
Gao, Qifeng 128  
Guan, Yanling 37  
Gui, Quanan 215  
Gui, Yu 448  
Guo, Jing 848  
Guo, Qingfeng 486  
Guo, Wenyan 571  
Guo, Xin 707  
Guo, Yongjie 944  
Guo, Zhishuai 389

## H

Han, Jieping 831  
Han, Yanqi 128  
Han, Zeping 138  
Hong, Shaoxian 684  
Hou, Fangbo 716  
Hou, Meng 184  
Hou, Yuanhong 841, 937  
Hu, Huanjun 530, 547, 587  
Hu, Rong 92  
Hu, Shuyu 815  
Hu, Yanjun 363  
Hu, Zhangrong 908  
Huang, Bin 381  
Huang, Chunxiang 890  
Huang, Jie 597  
Huang, Li 500, 511  
Huang, Mingfei 951

Huang, Ting 606  
 Huang, Wei 381  
 Huang, Xiaobo 168  
 Huang, Xiong 964

**J**

Jeong, Hwa-Young 10  
 Jeong, Hwayoung 3  
 Ji, Xiangchun 767  
 Jiang, Huan 424  
 Jiang, Shengchao 614  
 Jiang, Xiaolin 642  
 Jiang, Yuzhuo 659  
 Jin, Man 633

**K**

Kan, Wenting 700  
 Ke, Hongyu 555  
 Kong, Chen 159

**L**

Lao, Cuijin 579  
 Lee, Juhee 17  
 Li, Chenggao 511  
 Li, Hangkang 622  
 Li, Haoran 110  
 Li, Heng 128  
 Li, Hongtao 784, 792  
 Li, Hua 721  
 Li, Jing 597  
 Li, Lei 522, 530, 538, 547, 555  
 Li, Longfei 381  
 Li, Ren 231  
 Li, Ruichao 342  
 Li, Shanshan 951  
 Li, Shaochun 262  
 Li, Shijun 841  
 Li, Tailin 614  
 Li, Xue 831  
 Li, Yue 128  
 Li, Yujing 280  
 Liang, Jianquan 37  
 Liang, Yantao 667  
 Liang, Yijing 350  
 Liang, Yongdong 823  
 Liao, Rongtao 522, 547, 587  
 Lin, Hui 152  
 Liu, Bing 823  
 Liu, Changyu 269

Liu, Jia 951  
 Liu, Jingshuang 199, 206  
 Liu, Jun 299  
 Liu, Lu 848  
 Liu, Min 291  
 Liu, Qi 280  
 Liu, Shaowei 914  
 Liu, Shuchao 928  
 Liu, Wenting 884  
 Liu, Xinying 223  
 Liu, Xiqi 159  
 Liu, Xiran 60  
 Liu, Yefan 350  
 Liu, Yi 471  
 Liu, Zhijie 784  
 Liu, Zhiwei 884  
 Lu, Jiaqi 622  
 Lu, Xingming 784, 792  
 Luan, Zhiling 262  
 Luo, Lingxiang 402  
 Lv, Ming 417

**M**

Ma, Jun 759  
 Ma, Mingyuan 1001  
 Ma, Qianyu 902  
 Meng, Fanbo 168  
 Miao, Yuxin 448  
 Mu, Qi 767

**N**

Nam, Sinhye 26  
 Ning, Hao 530  
 Niu, Hao 775

**P**

Park, Haeyoung 3  
 Pei, Yunqing 614  
 Peng, Huiwen 775  
 Peng, Nan 848  
 Peng, Shirui 439  
 Peng, Yuefeng 606  
 Piao, Zheyong 831

**Q**

Qi, Guodong 280  
 Qi, Yongjun 500, 511, 737  
 Qin, Guohe 120  
 Qin, Shen 579



Qin, Shengbo 159  
Quan, Shuangwu 307

**R**

Ren, Haige 606  
Ren, Shunzi 448

**S**

Shan, Sicong 464  
Shao, Mingming 68  
Shen, Haijie 357  
Shen, Heng 692  
Shen, Qing 622  
Shen, Xuemei 363  
Shen, Yueying 993  
Shu, Jingjun 44  
Song, Yingying 745  
Sun, Jian 381  
Sun, LinLin 862  
Sun, Shuai 128, 138  
Sun, Shuang 464  
Sun, Xiaobo 493

**T**

Tang, Hailin 500, 737  
Tian, Jian 448  
Tong, Ning 633  
Tong, Ningning 800

**W**

Wan, Xuan 667  
Wan, Yameng 775  
Wang, Bin 614  
Wang, Bo 334  
Wang, Feifeng 614  
Wang, Fengying 571  
Wang, Gui 896  
Wang, Guojuan 767  
Wang, Han 37  
Wang, Hongchao 937  
Wang, Jingying 76  
Wang, Juyi 752  
Wang, Kunrui 633  
Wang, Lijun 767  
Wang, Lili 246  
Wang, Liying 784, 792  
Wang, Naiqing 622  
Wang, Peng 651

Wang, Qian 110  
Wang, Sanjiong 269  
Wang, Shanshan 159  
Wang, Shengwei 522, 530, 538, 597  
Wang, Xiaocheng 176  
Wang, Xinjun 1010  
Wang, Yaguang 823  
Wang, Yan 145  
Wang, Yixi 530, 538, 587  
Wen, Hongchang 269, 280, 775  
Wen, Ximeng 684  
Wen, Zhen 84  
Wu, Haiyan 328  
Wu, Wei 54  
Wu, Zhong 239

**X**

Xia, Fei 168  
Xia, Pin 102  
Xia, Weilun 138  
Xia, Yanduo 307  
Xiang, Zhongzheng 363  
Xiao, Bin 700  
Xie, Fei 606  
Xie, Jianxiang 299  
Xie, Shengli 110  
Xing, Xu 675  
Xu, Changchun 979  
Xu, Licheng 792  
Xu, Wenguang 958  
Xu, Yuanyuan 928  
Xue, Hui 841  
Xue, Shunhai 841, 937

**Y**

Yan, Guojun 890  
Yang, Guang 707  
Yang, Shaoqiang 54  
Yang, Shengjie 606  
Yang, Wenxin 986  
Yang, Yang 128  
Yang, Yanqin 922  
Yang, Yulai 841  
Yang, Zixin 855  
Ye, Dan 622  
Ye, Jingyu 381  
Ye, Weiwen 322  
Ye, Yang 971  
Ye, Yuxuan 522, 547, 587

Yi, Zhongli 159  
 Yin, Yina 350  
 Yu, Dan 878  
 Yu, He 199, 206  
 Yu, Miao 417  
 Yu, Wuzhan 823  
 Yuan, Jianghong 1016  
 Yuan, Li 606

## Z

Zeng, Fanjun 269  
 Zeng, Kangyan 84  
 Zeng, Mingquan 622  
 Zeng, Weiwei 439  
 Zhang, Fengxia 800  
 Zhang, Fucheng 389  
 Zhang, Guozhi 307  
 Zhang, Ji 299  
 Zhang, Jian 522, 538, 587  
 Zhang, Jianwei 254  
 Zhang, Liang 37  
 Zhang, Lina 191  
 Zhang, Qian 729  
 Zhang, Qiong 675  
 Zhang, Shuxin 716  
 Zhang, Xi 555

Zhang, Xilin 675  
 Zhang, Xingrong 870  
 Zhang, Xinyu 307  
 Zhang, Xuemei 431  
 Zhang, Yi 76  
 Zhang, Yuanjing 964  
 Zhang, Yujia 606  
 Zhao, Heqi 767  
 Zhao, Rongguo 456  
 Zhao, Wanting 128  
 Zheng, Chenxi 792  
 Zheng, Kai 363  
 Zheng, Yushan 831  
 Zhi, Xiaomin 120  
 Zhong, Weihua 299  
 Zhou, Pengpeng 784  
 Zhou, Xueyun 737  
 Zhou, Yuqiang 937  
 Zhou, Zheng 538, 547  
 Zhu, Guowei 597  
 Zhu, Qiankun 138  
 Zhu, Ying 775  
 Zhuo, Haoze 614  
 Zou, Chengwu 622  
 Zou, Haidong 54  
 Zuo, Zhengqin 606