Smart Innovation, Systems and Technologies 369

Roumen Kountchev
Srikanta Patnaik
Kazumi Nakamatsu
Roumiana Kountcheva *Editors*

Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)

Network Technologies: Mathematical Approaches and Applications, Volume 2

KES International

Springer

Smart Innovation, Systems and Technologies

Volume 369

The Smart Innovation, Systems and Technologies book series encompasses the topics of knowledge, intelligence, innovation and sustainability. The aim of the series is to make available a platform for the publication of books on all aspects of single and multi-disciplinary research on these themes in order to make the latest results available in a readily-accessible form. Volumes on interdisciplinary research combining two or more of these areas is particularly sought.

The series covers systems and paradigms that employ knowledge and intelligence in a broad sense. Its scope is systems having embedded knowledge and intelligence, which may be applied to the solution of world problems in industry, the environment and the community. It also focusses on the knowledge-transfer methodologies and innovation strategies employed to make this happen effectively. The combination of intelligent systems tools and a broad range of applications introduces a need for a synergy of disciplines from science, technology, business and the humanities. The series will include conference proceedings, edited collections, monographs, handbooks, reference books, and other relevant types of book in areas of science and technology where smart systems and technologies can offer innovative solutions.

High quality content is an essential feature for all book proposals accepted for the series. It is expected that editors of all accepted volumes will ensure that contributions are subjected to an appropriate level of reviewing process and adhere to KES quality principles.

Indexed by SCOPUS, EI Compendex, INSPEC, WTI Frankfurt eG, zbMATH, Japanese Science and Technology Agency (JST), SCImago, DBLP.

All books published in the series are submitted for consideration in Web of Science.

Roumen Kountchev · Srikanta Patnaik ·
Kazumi Nakamatsu · Roumiana Kountcheva
Editors

Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)

Network Technologies: Mathematical Approaches and Applications, Volume 2

Springer

Editors
Roumen Kountchev
Technical University of Sofia
Sofia, Bulgaria

Srikanta Patnaik
SOA University
Bhubaneswar, India

Kazumi Nakamatsu
University of Hyogo
Kobe, Japan

Roumiana Kountcheva
TK Engineering
Sofia, Bulgaria

ISSN 2190-3018 ISSN 2190-3026 (electronic)
Smart Innovation, Systems and Technologies
ISBN 978-981-99-6955-5 ISBN 978-981-99-6956-2 (eBook)
https://doi.org/10.1007/978-981-99-6956-2

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

Paper in this product is recyclable.

Preface

This is Volume 2 of the *Proceedings of the International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023): Network Technologies: Mathematical Approaches and Applications.*

The conference was held during June 9–11, 2023 in Shenzhen, China.

The aim of ICAICT 2023 was to provide a forum for researchers, academicians and practitioners to exchange ideas and discuss the latest achieved scientific results on artificial intelligence, communication technology and other related fields. The conference was focused on the establishment of an effective platform for institutions and industries to introduce the work of scientists, engineers, educators and students from all over the world.

After reviewing, 65 papers were accepted for presentation and publication, of which 32 are in this volume.

In memory of Prof. Dr. Roumen Kountchev, Dr. Sc., *Best Paper Award* was announced by the Organizing Committee. The *winner's paper* is published as Chap. 1 in Volume 1 of these proceedings. The title of the paper is "Multi-Objective Dynamic Optimization Scheme for Unmanned Aerial Vehicles", with authors: Bao-Qiang Zhuang and Ya-Jun Wang.

The chapters, collected in this volume, present research works based on AI and DL, Information Security, Multimedia and Big Data Analysis, Mathematical Approaches and Applications, etc. The chapters are arranged in four groups, defined in accordance with the used research approach:

Part I: DL and AI-based approaches (9 chapters)
Part II: Information Security and Power Management Control (7 chapters)
Part III: Approaches Based on Multimedia and Big Data Analysis (6 chapters)
Part IV: Mathematical Approaches and Applications (10 chapters)

The book editors express their special thanks to IRNet International Academic Communication Center who organized this conference in correspondence with their dedication to building platforms for scholars and researchers for better knowledge sharing, together with providing full-scale conference services that meet the standards from renowned publishing organizations.

We also thank Prof. Lakhmi Jain (Honorary chair), Prof. Dr. Srikanta Patnaik (General chair), Prof. Hang Li and Dr. Shoulin Yin (Organizing chairs), Dr. Muhammad Ibrar, Dr. Asif Ali Laghari and Dr. Anas Bilal (Programme chairs), and Dr. S. R. Roumiana Kountcheva (International advisory chair).

The editors express their warmest thanks to the excellent Springer team for making this book possible.

Sofia, Bulgaria	Prof. Dr. Roumen Kountchev
Bhubaneswar, India	Prof. Dr. Srikanta Patnaik
Kobe, Japan	Prof. Dr. Kazumi Nakamatsu
Sofia, Bulgaria	Dr. Roumiana Kountcheva
July 2023	

Contents

Editors and Contributors

About the Editors

Prof. Dr. Roumen Kountchev, D. Sc. is a professor at the Technical University of Sofia. He got his Ph.D. degree at the Institute of Telecommunications at Sankt-Petersburg, Russia and defended his dissertation for Doctor of Sciences at TU-Sofia. His scientific areas of interest are: digital signal and image processing, image compression, multimedia watermarking, video communications, pattern recognition and neural networks. Prof. R. Kountchev has 342 papers published in magazines and conference proceedings, and also: 30 books (20 published by Springer-Verlag, Germany); 50 book chapters and 20 patents (3 US). He had been principle investigator of 38 research projects. At present he is a member of Euro Mediterranean Academy of Arts and Sciences (EMAAS) and President of Bulgarian Association for Pattern Recognition (member of IAPR). Editorial board member of: Intern. Journal of Reasoning-based Intelligent Systems; Intern. Journal Broad Research in Artificial Intelligence and Neuroscience; KES Focus Group on Intelligent Decision Technologies; Egyptian Computer Science Journal; Intern. Journal of Bio-Medical Informatics and e-Health, and Intern. Journal Intelligent Decision Technologies. He has been a plenary speaker in many international scientific conferences and workshops.

Prof. Srikanta Patnaik is presently working as director of Interscience Institute of Management & Technology, Bhubaneswar, India. He has received his Ph.D. (Engineering) on Computational Intelligence from Jadavpur University, India in 1999. He has supervised 32 Ph.D. Theses and more than 60 Master theses in the area of Computational Intelligence, Soft Computing Applications and Re-Engineering. Dr. Patnaik has published more than 100 research papers in international journals and conference proceedings. He is author of 3 text books and more than 100 edited volumes and few invited book chapters, published by leading international publisher like IEEE, ACM, Elsevier, Springer-Verlag, Kluwer Academic, SPIE, IOS Press and others. Dr. Patnaik is the Editor-in-Chief of International Journal of Information and Communication Technology and International Journal of Computational Vision and Robotics published by Inderscience Publishing House, England and also Editor-in-chief of Book Series on "Modeling and Optimization in Science and Technology" published by Springer, Germany. He is the Editor of Journal of Information and Communication Convergence Engineering, published by Korean Institute of Information and Communication Convergence Engineering. Prof. Patnaik is a Guest Professor to Hunan University of Finance and Economics, Changsha and Kunming University of Science and Technology, Kunming, China and visiting professors to some of the B-Schools of Europe and South East Asia. He is a member of Institute of Electrical and Electronics Engineering (IEEE) and Association for Computing Machinery (ACM). He is also Fellow of IETE, Life Member of ISTE, and CSI. Dr. Patnaik has visited various countries such as Japan, China, Hong Kong, Singapore, Indonesia, Iran, Malaysia, Philippines, South Korea, United Arab Emirates, Morocco, Algeria, Thailand and Vietnam for delivering Key note addresses at various conferences and symposiums.

Prof. Kazumi Nakamatsu received the Ms. Eng. and Dr. Sci. from Shizuoka University and Kyushu University, Japan, respectively. His research interests encompass various kinds of logic and their applications to Computer Science, especially paraconsistent annotated logic programs and their applications. He has developed some paraconsistent annotated logic programs called ALPSN (Annotated Logic Program with Strong Negation), VALPSN (Vector ALPSN), EVALPSN (Extended VALPSN) and bf-EVALPSN (before-after EVALPSN) recently, and applied them to various intelligent systems such as a safety verification based railway interlocking control system and process order control. He is an author of over 150 papers and 20 book chapters, and 10 edited books published by prominent publishers. Kazumi Nakamatsu has chaired various international conferences, workshops and invited sessions, and he has been a member of numerous international program committees of workshops and conferences in the area of Computer Science. He has served as the Editor-in-Chief of the International Journal of Reasoning-based Intelligent Systems (IJRIS), he is now the founding editor of IJRIS, and an editorial board member of many international journals. He has contributed numerous invited lectures at international workshops, conferences, and academic organizations. He also is a recipient of numerous research paper awards. He is a member of ACM. https://www.researchgate.net/profile/Kazumi-Nakamatsu

Dr. Roumiana Kountcheva, S. R. is the Vice president of TK Engineering, Sofia and participates in R&D scientific projects of the Technical University of Sofia, Bulgaria. She got her Ph.D. at TU-Sofia, and became senior researcher in 1993. Her scientific interests are in image processing, image compression, digital watermarking, image tensor representation, CNCs, programmable controllers. She has more than 200 scientific publications and 5 patents and participated in 49 scientific research projects.

R. Kountcheva is a member of Bulgarian Association for Pattern Recognition (IAPR); International Research Institute for Economics and Management (IRIEM); Institute of Data Science and Artificial Intelligence

(IDSAI); Honorary Member of the Honorable Editorial Board of the nonprofit peer reviewed open access IJBST Journal Group, etc. Reviewer of WSEAS, NAUN, etc. She presented her research achievements in more 24 plenary speeches at international conferences and other scientific events and edited several books published in Springer SIST series.

Contributors

Ramir Santos Austria University of the Cordilleras, Baguio, Philippines

Shaobo Bai Research Institute of Highway Ministry of Transport, Beijing, China

Sohail Basheer Department of Computer Science, Sindh Madressatul Islam University, Karachi, Pakistan

Ting Cao Financial Information Management Center of Jiangsu Provincial Department of Finance Nanjing, Jiangsu, China

Jingde Chen State Grid Shanghai Electric Power Company, Shanghai, China; Qingpu Power Supply Company, State Grid Shanghai Municipal Electric Power Company, Shanghai, China

Jun Chen Guangxi Power Grid Corporation Metering Center, Nanning Guangxi, China

Qiuhua Chen Guangxi Power Grid Corporation Wuzhou Power Supply Bureau, Wuzhou Guangxi, China

Yongliang Chen China Southern Airlines Technology Branch Haikou Maintenance Baseline, China Southern Airlines Group Co.Ltd.Line, Haikou, China

Jiangyan Cheng Fujian Institute of Information Technology, Fuzhou, Fujian, China

Guoliang Dong Research Institute of Highway Ministry of Transport, Beijing, China

Bo Du Engineering University of PAP, Urumqi, China

Siyu Du Shaanxi Tobacco Company Xi'an Company, Xi'an, China

Xiaodong Du State Grid Hebei Electric Power Co., LTD, Shijiazhuang, Hebei, China

Yucheng Du Research Institute of Highway Ministry of Transport, Beijing, China

Jicheng Duan Hunan Railway Professional Technology College, ZhuZhou, China

Hua Gu State Grid Shanghai Electric Power Company, Shanghai, China;
Qingpu Power Supply Company, State Grid Shanghai Municipal Electric Power Company, Shanghai, China

Yuanfeng He Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen, China

Zhimin He State Grid Laboratory of Power Cyber-Security Protection & Monitoring Technology, State Grid Smart Grid Research Institute Co., LTD, Nanjing, Jiangsu, China

Chenhong Huang State Grid Shanghai Electric Power Company, Shanghai, China;
Qingpu Power Supply Company, State Grid Shanghai Municipal Electric Power Company, Shanghai, China

Muhammad Ibrar Software College, Shenyang Normal University, Shenyang, China

Tao Jiang Shaanxi Tobacco Company Xi'an Company, Xi'an, China

Mansoor Ahmed Khuhro Department of Artificial Intelligence, Sindh Madressatul Islam University, Karachi, Pakistan

Asif Ali Laghari Department of Computer Science, Sindh Madressatul Islam University, Karachi, Pakistan

Bo Li Electric Power Dispatching and Control Center of Guangdong Power Grid Co., Ltd., Guangzhou, China

Hao Li Research Institute of Highway Ministry of Transport, Beijing, China

Huimin Li Research Institute of Highway Ministry of Transport, Beijing, China

Jinjin Li Guangxi Power Grid Corporation Metering Center, Nanning Guangxi, China

Junlei Li State Grid Electric Space Technology Company Limited, Beijing, China

Nige Li State Grid Laboratory of Power Cyber-Security Protection and Monitoring Technology, Nanjing, China

Qing Li Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen, China

Wenba Li Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen, China

Xiaoxu Li Qilu Institute of Technology, Ji Nan City, Shandong, China

Xingnan Li Electric Power Dispatching and Control Center of Guangdong Power Grid Co., Ltd., Guangzhou, China

Zhibin Li Chang'an University, Xi'an, China

Zhangzhao Liang Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen, China

Xiuqing Lin Guangxi Power Grid Corporation Metering Center, Nanning Guangxi, China

Dan Liu China Electronic Port Data Center Branch Guangzhou, Guangzhou, China

Lisha Liu Shanghai Institute of Technology, Shanghai, China

Yanxia Liu Nanchang Institute of Technology, Nanchang, China

Zhaohua Long School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing, China

Huqin Luo Wuhan Technology and Business University, Wuhan, China

Chuntian Ma State Grid Electric Space Technology Company Limited, Beijing, China

Yanfang Mao Nantong Power Supply Company, State Grid Jiangsu Electric Power Co., Ltd., Nantong, Jiangsu, China

Ma Mingyu University of the Cordilleras, Baguio, Philippines

Fanghua Mo Guangxi Power Grid Corporation, Nanning Power Supply Bureau, Nanning Guangxi, China

Zhujian Ou Nantong Power Supply Company, State Grid Jiangsu Electric Power Co., Ltd., Nantong, Jiangsu, China

Sen Pan State Grid Laboratory of Power Cyber-Security Protection and Monitoring Technology, State Grid Smart Grid Research Institute Co., LTD., Nanjing, Jiangsu, China

Lin Peng State Grid Laboratory of Power Cyber-Security Protection and Monitoring Technology, State Grid Smart Grid Research Institute Co., LTD., Nanjing, Jiangsu, China

Wenbo Peng Krirk University, Bangkok, Thailand

Zuhaib Phul Department of Computer Science, Sindh Madressatul Islam University, Karachi, Pakistan

Junfeng Qiao State Grid Laboratory of Power Cyber-Security Protection and Monitoring Technology, State Grid Smart Grid Research Institute Co., LTD., Nanjing, Jiangsu, China

Yupeng Qin Guangxi Power Grid Corporation Metering Center, Nanning Guangxi, China

LianTeng Shen China Electric Power Research Institute Co.Ltd, Beijing, China

Xiaofeng Shen State Grid Shanghai Electric Power Company, Shanghai, China; Qingpu Power Supply Company, State Grid Shanghai Municipal Electric Power Company, Shanghai, China

Zhanyu Shen Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen, China

Zhan Shi Electric Power Dispatching and Control Center of Guangdong Power Grid Co., Ltd., Guangzhou, China

Yang Su Guangzhou Customs Technology Center, Guangzhou, China

Min Sun Wuhan Institute of Design and Sciences, Wuhan, China

Xiaoting Sun Krirk University, Bangkok, Thailand

Zijun Tan Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen, China

Taiqing Tang School of Electrical and Information Engineering, Heilongjiang University of Technology, JiXi City, Heilongjiang, China

Maojie Tian State Grid Electric Space Technology Company Limited, Beijing, China

Gang Wang State Grid Laboratory of Power Cyber-Security Protection and Monitoring Technology, State Grid Smart Grid Research Institute CO., LTD., Nanjing, Jiangsu, China

He Wang State Grid Laboratory of Power Cyber-Security Protection and Monitoring Technology, State Grid Smart Grid Research Institute Co., LTD, Nanjing, Jiangsu, China

Lijun Wang Hainan Health Vocational College, Haikou, China

Lingyun Wang Guangzhou Customs, Guangzhou, China

Ning Wang State Grid Jiangsu Electric Power Co., LTD, Nanjing, Jiangsu, China

Zhenwei Wang Guangdong University of Science & Technology, Dongguan, China

Xiwen Wei School of Electrical and Information Engineering, Heilongjiang University of Technology, JiXi City, Heilongjiang, China

Guofang Wu Research Institute of Highway Ministry of Transport, Beijing, China

Zhaotao Wu School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing, China

Lingyue Xie South China University of Technology, Guangzhou, China

Min Xu State Grid Laboratory of Power Cyber-Security Protection and Monitoring Technology, State Grid Smart Grid Research Institute Co., LTD, Nanjing, Jiangsu, China

Guozhu Yang State Grid Electric Space Technology Company Limited, Beijing, China

Pei Yang State Grid Laboratory of Power Cyber-Security Protection and Monitoring Technology, State Grid Smart Grid Research Institute Co., LTD., Nanjing, Jiangsu, China

Jianguang Yao Nantong Power Supply Company, State Grid Jiangsu Electric Power Co., Ltd., Nantong, Jiangsu, China

Zilu Ying Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen, China

Xiao-Rong You ChongQing Industry and Information Vocational College, Chongqing, China

Hai Yu State Grid Laboratory of Power Cyber-Security Protection & Monitoring Technology, State Grid Smart Grid Research Institute Co., LTD, Nanjing, Jiangsu, China

Zhao-Jun Yu Engineering University of PAP, Urumqi, China

Chen Yuan Engineering University of PAP, Urumqi, China

Wenbin Yuan Shaanxi Tobacco Company Xi'an Company, Xi'an, China

Wei Zeng Yibin Vocational and Technical College, Yibin, Sichuan, China

Ying Zeng Electric Power Dispatching and Control Center of Guangdong Power Grid Co., Ltd., Guangzhou, China

Yikui Zhai Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen, China

Jian Zhang Electric Power Dispatching and Control Center of Guangdong Power Grid Co., Ltd., Guangzhou, China

Kechang Zhang Hunan Railway Professional Technology College, ZhuZhou, China

Xike Zhang Shenzhen National Climate Observatory, Shenzhen, China

Xinli Zhang Guangdong Medical University, Dongguan, China;
Philippine Christian University, Manila, The Philippines

Yanbin Zhang Guangzhou Customs Technology Center, Guangzhou, China

Yifeng Zhang Chang'an University, Xi'an, China

Bangbo Zhao State Grid Electric Space Technology Company Limited, Beijing, China

Aihua Zhou State Grid Laboratory of Power Cyber-Security Protection and Monitoring Technology, State Grid Smart Grid Research Institute CO., LTD., Nanjing, Jiangsu, China

Jianhong Zhou Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen, China

Yafeng Zhu Guangzhou Customs Technology Center, Guangzhou, China

Part I
DL and AI-based Approaches

Chapter 1
Command and Control Modeling Method for Federated Intelligence

Bo Du, Chen Yuan, and Zhao-Jun Yu

Abstract To explore the problems related to joint intelligent command and control modeling, analyze the scale correlation characteristics of the C2 model from the perspective of operational process, and compare and analyze the C2 operation mechanism under different scales; propose two types of C2 framework models based on edge intelligence and cloud-brain intelligence; propose a collaborative C2 framework based on joint cloud-edge-end intelligence, and lay the foundation for joint intelligent command and control modeling.

1.1 Introduction

With the development of artificial intelligence, edge computing, mobile cloud computing and other new generation information technology, modern warfare intelligence process continues to accelerate, artificial intelligence-driven joint combat command and control has become the supremacy of the military forces of various countries. In modern warfare, the importance of any one activity cannot be compared with command and control, and integrated joint warfare also needs a joint intelligent C2 (Command and Control, C2) model to adapt with it. Exploring the construction method of joint C2 model is of great significance for developing and improving the command and control theory system.

From the development of command and control models, the earliest can be traced back to the early 1950s when John Boyd proposed the OODA (Observe-Orient-Decide-Act, OODA) ring model [1], emphasizing that the key to gaining the battlefield initiative lies in accelerating the command and control "tempo" and disrupting the enemy's C2 "tempo", which became the classic model of micro-scale C2. Since then, based on its own military strategy and operational concept development, the U.S. Army has successively proposed the HEAT (Headquarters Effectiveness Assessment Tool, HEAT) ring [2], The OODA ring model has been developed

B. Du (✉) · C. Yuan · Z.-J. Yu
Engineering University of PAP, Urumqi 830049, China
e-mail: 726682174@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_1

by the U.S. military, including the SHOR (Stimulus-Hypothesis-Option-Response, SHOR) model, Lawson model, Rasmussen model, SA model, PDCA (Plan-Do-Check-Action, PDCA) ring, RPDM (Recognition-Primed-Decision-Making, RPDM) model, etc. Although the model design has its own focus, it is essentially a revision and development of the OODA ring model. The OODA loop model fails to address the deficiencies of the OODA loop model in terms of vague concept definition, lack of operational planning process, lack of cognitive and knowledge reasoning limited to regular reasoning, and lack of action monitoring and feedback [3].

Yang Dongsheng, Zhang Weiming et al., proposed the scale correlation of command and control model [4], established a general process description model PREA ring for command and control of the combat system from the macroscopic scale [5–8], which provides a systematic approach for C2 modeling at the macroscopic scale. From the current research, the following problems still exist in the study of intelligent C2 models: first, what kind of correlation characteristics exist between command and control models at different scales, and what impact these characteristics have on the construction of models is not clear; second, what kind of synergistic transformation relationship exists between C2 models at the macroscopic and microscopic scales, and how can C2 models at different scopes be "jointly". The second is what kind of synergistic transformation relationship exists between C2 models at the macro and micro scales, and how to establish a unified framework for joint intelligent perception of combat situation, joint optimal scheduling of combat resources, and joint intelligent decision making for combat planning through the "joint" approach.

Based on this, this paper firstly analyzes the scale correlation characteristics of C2 model from the perspective of the combat process; accordingly, it proposes a C2 framework model based on F-OODA (Feedback Observe-Orient-Decide-Act, F-OODA) for edge intelligence corresponding to tactical edge intelligence model for the microscopic scale, and a C2 framework model based on AI-PREA (Artificial Intelligence on Planning-Readiness-Execution-Assessment, AI-PREA) corresponding to cloud-brain intelligence for the macroscopic scale; and then unifies the federated intelligence command and control process in the tactical edge cloud service. Under, a collaborative C2 framework based on joint cloud-edge-end intelligence is proposed to provide a model reference for joint intelligent command and control.

1.2 C2-Scale Correlation Characteristics for Operational Processes

Research shows that the C2 model has scale correlation characteristics, and such operational scale correlation characteristics not only include time scale, space scale and force size, but also show different correlation characteristics and characteristic mechanisms in different stages. Only an accurate insight can better guide the construction of the C2 model. To this end, this section splits joint warfare into five stages: situational awareness, combat operations, command and decision, operational control, and evaluation and feedback from the spatial scale, and analyzes the operation mechanism of C2 under two operational scales: micro and macro.

1.2.1 Mechanisms Characterizing the Situational Awareness Process

The mechanism analysis of the scale-related characteristics of the situational awareness process is shown in Table 1.1. Situational awareness is the initial stage of joint combat command and control. Whether it is a single combat platform under the microscopic scale or a joint combat system under the macroscopic scale, situational awareness is the prerequisite and basis for operational decision making. The difference between situational awareness at different operational scales is that situational awareness at the microscopic scale usually focuses on local battlefield environmental target awareness, establishing a channel from target situational information to the weapon control system, providing the weapon system (platform) with decision data based on target parameters, and realizing the rapid response from “observation” to “action”. The aim is to establish the advantage of countermeasures at the microscopic scale. Macro-scale situational awareness can be understood as, through the cooperative sensing of multi-source sensing devices, multi-source information fusion and situational integration, to achieve the fusion of the full-dimensional battlefield space, as well as the global situation affecting the battlefield, such as national defense and military, political economy, history and culture, so as to provide the basis for situational decision-making, form trend analysis and prediction, and achieve macro-scale situational suppression, such as typical aircraft carrier battle groups The most typical ones are carrier battle groups, amphibious assault groups, etc.

1.2.2 Combat Course of Action Characteristic Mechanism

The mechanism analysis of the scale correlation characteristics of the combat operation process is shown in Table 1.2. The combat operation process is the key link to establish battlefield superiority and achieve the goal of victory over the enemy. The domain of combat operations at the microscopic scale (domain of influence) is mainly limited to the physical domain, and the combat effectiveness or operational effects of weapons platforms are mainly determined by physical attributes such as platform shape and layout, mobility performance, and strike progress, so they can be achieved by optimizing structural layout, improving platform design, and increasing strike accuracy, such as adding terminal guidance and autonomous maneuvering devices to tactical missiles to reduce the probability of interception. The system is a very important part of the war. The effectiveness of system operations at the macroscopic scale, which extends to the physical, information, cognitive, and social domains, is

Table 1.1 Mechanism for characterizing the situational awareness process

Operational phase	Scale of operations	Sensing range	Perception mode	Perceptual content	Objectives
Situational awareness	Microscopic	Local posture (line-of-sight reconnaissance)	Local sensing (reconnaissance, detection, surveillance)	Enemy situation, our situation, battlefield environment	Target-awareness-based, providing target parameters for weapon systems to gain microscale countermeasure advantage
	Macro	Global posture (over-the-horizon)	Cooperative sensing (multi-source sensing device collection, multi-source information fusion, multi-source situational awareness judgment, situational intelligence integration, etc.)	Land, sea, air, sky, electricity 5-dimensional battlefield space, history and culture, political economy, national defense and military, etc.	Multi-dimensional battlefield space situational awareness and situational fusion, both early warning intelligence, trend analysis, situational integration capabilities, to achieve macro-scale situational suppression

Table 1.2 Mechanisms for characterizing combat operations processes

Operational phase	Scale of operations	Domain of influence	Influencing factors	Advantage building
Combat operations	Microscopic	Physical domain	Weapon platform shape layout, power performance, firing accuracy, gun power, etc., combat environment, physical properties of combat opponents	Optimize layout, improve design, enhance accuracy, etc.
	Macro	Physical domain, information domain, cognitive domain, social domain	Situational awareness, situational understanding and situational judgment of commanders and command structures, information and social attributes such as force deployment, execution logic, and platform synergy efficiency	Intelligent cognition, intelligent judgment, intelligent decision-making, intelligent collaboration

greatly limited by the commanders' and command agencies' understanding, judgment, and cognition of the situation, and by the use, deployment, coordination, and layout of forces, and is also affected by the execution logic of combat units and the efficiency of platform coordination, requiring C2 systems to provide decision making, judgment, and coordination capabilities based on intelligent situational awareness, judgment, and establish operational advantage.

1.2.3 Mechanisms for Characterizing the Command Decision Process

The scaled correlation characteristics of the command decision process are mechanistically analyzed as shown in Table 1.3. The command decision-making process is the soul of C2, and all operational results are attributed to the merits of command decisions. The command decision at the microscopic scale is usually a tactical-level combat unit due to its limited force size, narrow decision opening, and fast conversion of the combat process, requiring the commander to make a quick combat decision judgment based on the basic combat model and combat methods through a quick judgment of the current situation, and to issue combat orders to guide the tactical detachment to complete tactical operations. Command decisions at the macroscopic scale with a relatively long operational conversion cycle make for a large decision window, enabling the commander and command structure to perfect the operational concept through collaborative decision making and operational preplanning, form a determined program, and rely on the C2 system to issue instructions to each operational platform for coordination, strike, reinforcement, and security.

Table 1.3 Mechanism of command decision process characteristics

Operational phase	Scale of operations	Decision-making approach	Timing of decision	Decision content	Decision-making organizations
Command decision	Microscopic	Rule-based decision making	Narrow, fast-paced decision-making window	Execution instructions	Individual decision making
	Macro	Prognosis-based decision making	Large decision window and slow pace	Operational concepts, determination programs, operational guidance, and various directives	Collaborative decision making

1.2.4 Mechanisms of Action Control Process Characteristics

The mechanism analysis of the scale correlation characteristics of the operational control process is shown in Table 1.4. The operation control process is an important link to control the operation process, master the rhythm of the operation, and monitor the effect of the operation. Operation control at the microscopic scale usually divides the control interface by combat mission or weapon system, and realizes the operation control of weapon system or tactical unit through direct control and command control. Macro-scale operational control is usually divided into control interfaces by military service or theater, through indirect control, planning control, or cooperative control to achieve the operational process, operational tempo, operational spectrum, and situational monitoring of the operational domain.

Table 1.4 Mechanism of action control process characteristics

Operational phase	Scale of operations	Control objects	Control method	Control method
Action control	Microscopic	Operational control (weapon systems, tactical units)	Direct control and command control	Distinguished by mission, weapon system
	Macro	Situational control (process control, tempo control, spectrum control, airspace control, sea area control), the	Indirect control and planning control	Distinguished by military service, combat operations

1.2.5 *Assessing the Mechanism of Feedback Process Characteristics*

The scale correlation mechanism analysis of the assessment feedback process is shown in Table 1.5. The main role of the assessment feedback process is to provide a basis for mission transformation, adjustment of action plans, and optimization of force deployment, and to determine whether the charge control process is accurate and reasonable. The assessment and feedback process at the microscopic scale is mainly deterministic due to its narrow operational window and restricted operational territory, and is usually based on the commander's rapid decision to switch the function or adjust the status of the weapon system through instructions or operations. The assessment feedback under macroscopic scale is not the case, the complex composition of combat nodes under system confrontation and the obvious nonlinear characteristics of the combat process make the system effectiveness assessment process possess huge disturbance and uncertainty, so the assessment under macroscopic view is usually based on uncertainty of fuzzy comprehensive evaluation, and the combat system completes the transformation of the combat process, adjustment of combatant, and improvement of the combat plan based on the assessment results.

1.3 Federated Intelligence-Based C2 Modeling

The scale-related characteristics of the C2 model indicate that the scope of application, structural features, functional characteristics, organizational forms, and control methods of the C2 model at different operational scales differ significantly and cannot be generalized. For this reason, this paper proposes a C2 framework model for federated intelligence at different operational scales.

Table 1.5 Characterization mechanism of the feedback process

Operational phase	Scale of operations	Evaluation model	Conversion mode	Expression form
Evaluation feedback	Microscopic	Deterministic assessment	Implicit conversion (does not exist independently)	Function switching or status change of weapon system
	Macro	Uncertainty assessment	Explicit conversion (independent existence)	Operational process transition, operational planning adjustments

1.3.1 F-OODA-Based Edge Intelligence C2 Framework Model

In view of the command and control adaptability under the microscopic scale of OODA ring model, combined with the problems of the model in concept definition, cognitive judgment, reasoning and decision making, and supervisory feedback [3], this paper proposes a self-feedback OODA loop (F-OODA) based edge intelligence C2 framework model, shown in Fig. 1.1. This model is mainly used for tactical operations or single weapon platform command and control. In addition to the improvement of the intelligence of each link, it is noteworthy that the model adds the operational feedback link from the structure, and its main functions include, first, based on battlefield situational monitoring, assessing the operational effectiveness through effectiveness assessment, providing a basis for adjusting the action plan and optimizing the operational deployment, making up for the lack of monitoring feedback in the tactical edge C2 process; second, resource optimization scheduling, given the limited bandwidth, storage, computing and other resources at the tactical edge Third, model optimization, with limited resources at the tactical edge, higher requirements are placed on the model volume and model parameter scale, and firstly, federal learning technology is used for model initialization training, and later, based on the model self-feedback mechanism, model initialization training is conducted. Based on the knowledge distillation technique, distributed reinforcement learning techniques are used to optimize and prune the model parameters, and finally, through model segmentation and model compression, adaptive deployment, adaptive offloading, and joint computation are performed according to the task detachment requirements.

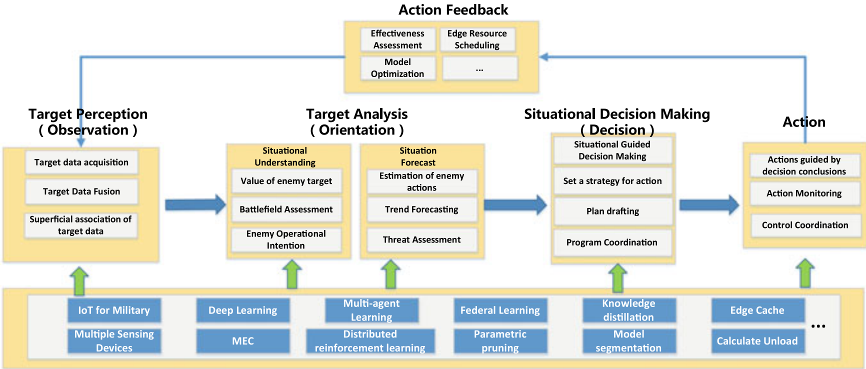


Fig. 1.1 F-OODA-based edge intelligence C2 model

1.3.2 Cloud Brain Intelligence C2 Framework Model Based on AI-PREA

PREA model, the basic idea is to examine the mechanism of command and control and decision-making activities from the system perspective and macroscopic scale, and establish a process model of command and control applicable to system operations. In this paper, we propose an AI-PREA-based cloud brain intelligence C2 framework model based on the PREA model infrastructure, as shown in Fig. 1.2, to adapt the macro-scale C2 process of joint warfare.

Its characteristics include three aspects: First, the focus on the overall process of system operations. That is, from the beginning of the received combat mission to combat mission operations, always around the realization of the mission objectives of command and control and decision-making activities, rather than a moment of real-time command activities, immediate and efficient OODA ring may gain advantages for local or instantaneous confrontation, but the instantaneous micro-level advantages in the system at the macro-level does not necessarily have a characterization, and vice versa. Second, the focus is on continuous changes in battlefield posture. At the system level, the time scale is larger than that of the combat platform, resulting in a situation that is no longer transient, but a continuous evolutionary process. For the commander, the situation it faces is the process of gradually converging to the real situation, that is, before the war to face the historical statistics of the predicted situation, in the

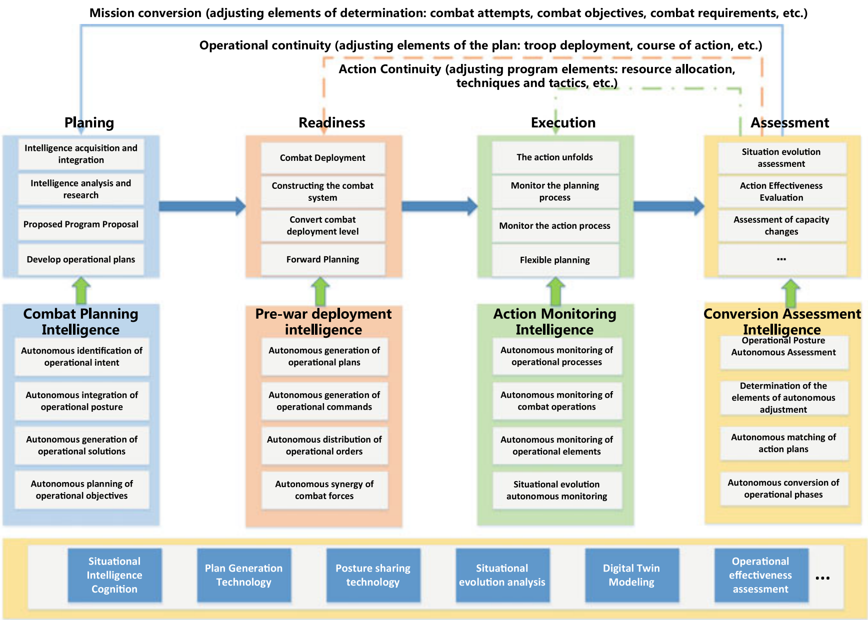


Fig. 1.2 Cloud brain intelligence C2 framework model based on AI-PREA

straight ahead of the preparation stage to face the straight ahead based on forecasts, in the implementation stage of combat operations to face the real-time situation based on the fusion of sensor-acquired data, in the operational evaluation stage to face the historical situation based on the evolutionary process. Third, focus on the dynamic process of command and decision-making. Because the combat situation faced by the commander is a gradually real and continuously evolving process, its command decision-making activities are also dynamically adjusted. Based on the predicted situation can be organized pre-planning, based on the straight forward situation can be organized straight forward planning, based on the real-time situation to organize the critical planning, pre-planning, straight forward planning. Critical planning is in the system level command decision-making problems gradually and accurately seek the help process.

1.4 C2 Collaborative Framework Based on Joint Cloud-Edge-End Intelligence

The scale correlation of C2 processes determines that the scope of application, structural features, and intelligence characteristics of C2 processes under different viewpoints (macro/micro). The above discussed the construction of two types of C2 models under different scale viewpoints, and only when they are synergistic and symbiotic and unified under a unified framework can they better serve operations and maximize their effectiveness. To this end, this paper designs a collaborative C2 framework based on joint cloud-edge-end intelligence, as shown in Fig. 1.3. The intelligent command and control process is unified under the tactical edge cloud service system.

The framework adopts the "cloud-edge-end" synergy approach, and the synergy modes mainly include edge-cloud synergy, edge-edge synergy, edge-end synergy, and cloud-edge-end synergy.

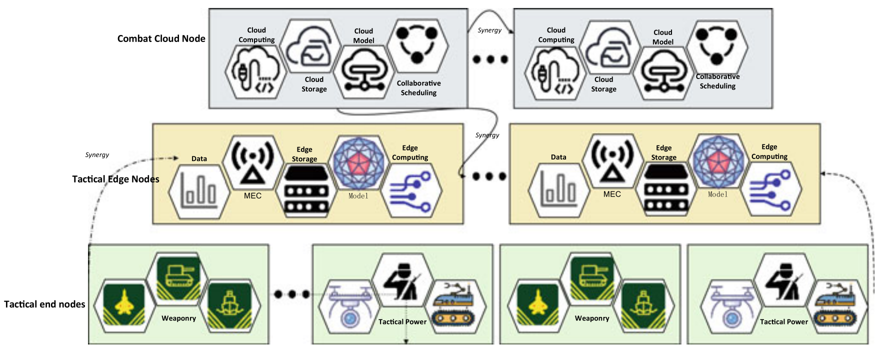


Fig. 1.3 C2 collaborative framework based on joint cloud-edge-end intelligence

1.4.1 The Edge-Cloud Collaboration

It mainly used to solve the contradiction between the limited computing resources of tactical edge nodes and the demand for combat computing resources. By reasonably optimizing the task allocation strategy, disassembling and slicing intelligent algorithms, using the powerful computing capacity of the combat cloud to undertake such public basic class computing tasks as situation aggregation, geographic information services, scheduling queries, etc., linking the tactical edge-side computing pressure to meet the intelligent algorithms and computing power requirements. Edge-cloud collaboration can be achieved by migration learning, model splitting, model compression and other technologies [9–11].

1.4.2 Edge-Side Collaboration

The main purpose is to solve the two problems of limited computing capacity of a single node and difficulties in collaborative sharing of data resources by allocating computing power, network, storage and other resources to tactical edge nodes under the same mission framework to achieve collaborative model training and collaborative data sharing in a distributed collaborative training mode. One is distributed collaborative training, in which each tactical edge node uses the local combat force, combat mission, enemy and enemy situation and battlefield environment data collected by edge sensing devices to train intelligent decision models in a distributed manner, and then changes the model parameters to the edge center node to complete global model training and application. Second, federal collaborative training [12]. The federated learning is proposed based on data security and privacy purposes. The federated learning in edge-side collaboration described by the author mainly uses the federated architecture for distributed data collaboration and model training, in which the tactical edge nodes can decide data collaboration, participation in training and parameter sharing timing according to the operational process. The key technologies involved in edge-side collaboration include distributed training [13], federation learning [12] etc.

1.4.3 Side End Collaboration

It mainly refers to the synergy between tactical edge nodes and tactical terminal nodes, which can be understood as weapon systems and tactical forces equipped with multi-source sensing devices with ubiquitous sensing capabilities. It is used to solve the contradiction between complex and diverse combat styles and single terminal node capability. Using the advantages of first-line deployment, rapid sensing, and three-dimensional deployment and control of tactical terminal nodes, they can collaborate on combat tasks and improve the effectiveness of tactical edge charges. In its collaborative process, the tactical terminal node is mainly responsible for data collection and transmission to the tactical edge node, while receiving instructions from the edge node and executing them; the tactical edge node is responsible for parallel collection and centralized calculation of multiple data, issuing instructions and providing edge services to the face-to-face tactical terminal node. The main problem faced in the process of edge-end collaboration is how to deploy through model tailoring, compression and lightweighting under resource-constrained and complex battlefield environments [14, 15]. The main problem is how to meet the operational requirements in intelligent computing scenarios through model tailoring, compression and lightweight deployment in resource-constrained and complex battlefield environments.

1.4.4 Cloud-Side End Collaboration

It mainly refers to the collaborative approach taken to make full use of the resources of the entire combat command link and to bring into play the advantages of different nodes in the command and control system. In this process, tactical terminal nodes are responsible for the collection of combat data, tactical edge nodes are responsible for the pre-processing of data resources, and combat cloud nodes are responsible for the aggregation, governance, sharing and service of data. Due to the limited arithmetic resources, under the MEC (Mobil Edge Computing, MEC) framework, various types of combat nodes at different levels carry out distributed collaboration based on different arithmetic power, networks, and models, driven by combat mission requirements, thereby enhancing the flexibility of command and control organizations, improving the resistance to destruction of the accusation system, and shortening the adjustment time of the accusation organization structure.

It is important to emphasize that the key to cloud-side end-integrated collaboration lies in the selection of different modes of collaboration according to the accusation tasks, and the core lies in the marginalized deployment of the accusation model, which enables the entire accusation architecture to dynamically adjust the deployment structure according to the complex and changing battlefield links, respond to the emergence and uncertainty of combat under complex confrontation conditions, and achieve dynamic and flexible self-adaptation of command and control according to the battlefield posture. The command and control is dynamically and flexibly self-adaptive and self-coordinated according to the battlefield posture, and the combat force is “assembleable”, “cuttable” and “transformable” to improve the overall effectiveness of the accusation.

1.5 Conclusion

In this paper, we propose an intelligent C2 framework model with scale adaptability based on the scale correlation characteristics of the C2 model, distinguish between macroscopic and microscopic scales, design a C2 collaborative framework based on joint cloud-side end intelligence, unify C2 models of different scales under the cloud service system, and analyze their collaborative mechanism to realize the integrated intelligence of situational awareness, command and control, and collaborative decision making, and provide model reference for the application of distributed domain-wide command and control.

Acknowledgements Supported by Defense Science and Technology Innovation Program Projects NO. ZZKY20222306 and National Social Science Foundation of China No. 2019-SKjj-C-092.

References

1. Boyd, J.: A Discourse on Winning and Losing. Maxwell Air Force Base, AL: Air University Library Document No. M-U 43947 (Briefing slides) (1987)
2. Hayes, R.E., Wheathey, G.: The evolution of the headquarters effectiveness assessment tool (HEAT) and its application to joint experimentation. In: The 6th International Command and Control Research and Technology Symposium (2001)
3. Ant, T., Kooter, B.: Comparing OODA & other models as operational view C2 architecture. In: Proceedings of the 10th International Command and Control Research and Technology Symposium. CCRP, McLean, Virginia (2005)
4. Yang, D.S., Zhang, W.M.: Relationship between mechanism and scale of command and control. J. Command Control **5**(3), 191–208 (2019)
5. Yang, D.S., Jiang, J., Wang, F.Y.: From platform to systems of systems: on mechanism evolution of command confrontation and its PREA loop. J. Command Control **4**(2), 101–110 (2018)
6. Yang, D.S., Yan, J.J.: Command and control behaviors and mechanism of macro operation systems of system. J. Command Control **6**(4), 393–401 (2020)
7. Yang, D.S., Zhu, C., Xiao, W.D., Zhang, W.M., Li, Q., Xiu, B.X., Wang, F.Y.: Process mechanism for macro command and control: coordinate conflicts and resolution model for multi-PREA loop in multi-domains. J. Command Control **7**(1), 11–27 (2021)
8. Yang, D.S., Li, Q., Liu, Y.C., Zhu, C., Xiu, B.X., Zhang, W.M., Wang, F.Y.: Macro-scale C2 process mechanism: revision and application of theory model of PREAloop. J. Command Control **8**(4), 389–440 (2022)
9. Pan, S.J., Yang, Q.: A survey on transfer learning. IEEE Trans. Knowl. Data Eng. **22**(10), 1345–1359 (2009)
10. Kang, Y., Hauswald, J., Gao, C., Rovinski, A., Mudge, T., Mars, J., Tang, L.: Neurosurgeon: collaborative intelligence between the cloud and mobile edge. ACM SIGARCH Comput. Archit. News ACM **45**(1), 615–629 (2017)
11. Han, S., Mao, H., Dally, W.J.: Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding (2015). arXiv:1510.00149
12. McMahan, B., Ramage, D.: Federated Learning: Collaborative Machine Learning Without Centralized Training Data. Google Research Blog (2017)
13. Dean, J., Corrado, G., Monga, R., Chen, K., Devin, M., Mao, M., Ranzato, M., Senior, A., Tucker, P., Yang, K., Le, Q., Ng, A.: Large scale distributed deep networks. Adv. Neural Inf. Process. Syst. 1223–1231 (2012)

14. Iandola, F.N., Han, S., Moskewicz, M.W., Ashraf, K., Dally, W.J., Keutzer, K.: SqueezeNet: AlexNet-Level Accuracy with 50x Fewer Parameters and <0.5 MB Model Size (2016). arXiv:1602.07360
15. Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H.: Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications (2017). arXiv:1704.04861

Chapter 2
Research on the Influence Mechanism of External Environment Impact on University Students' Psychological and Emotional Fluctuations Based on Deep Learning

Huqin Luo

Abstract With the continuous advancement and development of information technology, the psychological problems of university students have attracted more and more social attention. Teachers should pay attention to the psychological problems of university students, actively strengthen psychological crisis pre-warning and intervention measures, and constantly eliminate psychological problems of students. In this paper, a prediction model of university students' psychological mood fluctuation based on fuzzy neural network (FNN) is proposed, and the pre-alarm of university students' psychological health status is completed through the cooperation among neurons such as psychological data collection and psychological well-being evaluation. Compared with the traditional psychological prediction algorithm, the accuracy of this algorithm has obvious advantages. Therefore, it is feasible to analyze the influence mechanism of external environmental impact on university students' psychological and emotional fluctuations through this model. The impact of external environmental shocks on the psychological and emotional fluctuations of college students includes: an increase in uncertainty in life and a lack of psychological security; Increasing difficulty in employment and further education stimulates the generation of negative emotions; Some students have intensified family conflicts and amplified the negative impact of family relationships.

H. Luo (✉)
Wuhan Technology and Business University, Wuhan, China
e-mail: 598073347@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_2

2.1 Introduction

In recent years, with the increasingly fierce social competition, the quantity of students who drop out of school or drop out of school due to psychological problems is increasing year by year. As a special group, university students are going through the transition period of life and the reconstruction period of personality from teenagers to adults, both physically and psychologically. Psychological immaturity makes university students easily suffer psychological troubles under the background of drastic social changes [1]. As an environmental product, university students' psychological problems are closely related to social background, family and school environment and even individual interpersonal environment. To do a good job of psychological counseling for university students, we must start with optimizing the psychological counseling environment [2]. As the front position of ideological and political education, universities will be in the upstream position of the whole social ideological and political education chain because of their school-running purpose and training objects [3]. To do a good job of psychological counseling for university students, we should closely follow the ideological and political work from the source, adhere to the principle of people-orientation and pay attention to the trend and requirements of the times. From the perspective of developmental psychology, university students are at the critical stage of physical and mental development, and they are easily influenced by adverse factors. Coupled with individual's wrong cognitive style and bad coping style, psychological distress will gradually turn into a psychological crisis that affects individual's normal development over time [4]. Some students make some excessive behaviors because they can't dissolve their inner emotions or face the pressure, which may have a life-long impact [5]. This will not only affect the reputation of the university, but also be a heavy blow to the family. In order to find the psychological problems of university students in time and provide psychological help, this paper studies the intelligent psychological state evaluation method. Traditional psychological well-being education, pays more attention to the prevention and treatment of university students, and amplifies their negative psychological emotions, such as depression, anxiety and fear, hoping to find a solution by amplifying these negative emotions, but without any effect, and it does not eliminate their psychological problem [6]. For the current psychological education of university students, it is extremely important to implement effective psychological crisis prevention strategies for university students, which can timely prevent students from psychological crisis [7]. As an auxiliary means, scientific and reasonable frustration education is conducive to improving the psychological quality of university students and promoting their correct and comprehensive understanding of society and self [8]. This paper puts forward a prediction model of university students' psychological mood fluctuation based on deep learning, and completes the pre-alarming of university students' psychological well-being through the cooperation among agents such as psychological data collection and psychological well-being evaluation. This method is based on the text of internal student forums in universities and introduces the FNN model in deep learning.

2.2 Methodology

2.2.1 The Causes of University Students' Psychological Crisis and the Application of Positive Psychology

Sudden and irresistible major events usually include major natural disasters and social injuries. These uncontrollable and sudden major events often make people unable to fight back and there exist stress-related obstacles because of their suddenness. According to incomplete statistics, many university students have produced some extreme behaviors because of psychological problems, such as running away from home and hurting people physically. The occurrence of these events not only caused some losses to the families concerned, but also had a great impact on the study and daily life of university students [9]. Positive psychology believes that people with serious mental illness in chaotic psychology need more satisfactory attitudes and a happy environment. Positive psychology pays more attention to the positive qualities of human individuals, and emphasizes the potential of discovering positive aspects in human psychology. It is necessary to change the traditional psychological research direction and avoid magnifying the negative emotions in human psychology.

After entering the university, the academic study is not as easy as expected, and the frustration of continuous academic efforts and the failure of grades can easily lead to psychological crisis. Nowadays, university students' study and life are inseparable from new media, affecting university students' psychological growth to a certain extent [10]. We can solve human psychological problems through positive psychology, and confront the problems in life through positive psychology, so as to promote human physical and psychological well-being, experience happiness in life, and bring positive psychology to everyone around us to create a harmonious social environment.

At this stage, many universities archive students' psychological problems and record each one, which can effectively understand students' psychological well-being information. Often, some behaviors are the main manifestations of students' psychological activities, so students' psychological health supervision can also be realized through behavioral activities and hobbies. Positive psychology can effectively prevent mental illness and achieve good results in practical application. The main reason is that positive psychology can reshape various functions of the human body, not correcting the psychological defects of the human body, but changing it as a whole, which can really enter the inner activities of human beings and develop the positive side of human psychology so that it can use its own strength to prevent mental illness.

2.2.2 Prediction Model of Psychological and Emotional Fluctuation of University Students

Universities should establish adequate pre-plans, sum up and integrate them in time after the crisis, build a full-time psychological crisis intervention mechanism linked by home and school, and build a protective barrier for students' psychological well-being through the whole process of psychological crisis intervention system before, during and after the crisis. In order to obtain more accurate results, we should first improve the collection of information, and only by ensuring the accuracy of information can we have an accurate understanding of students' psychological well-being activities [11]. Based on the big data Internet, schools can appropriately improve the quality of questionnaires. Through the information collection based on big data, we can have a clear understanding of students' behavioral activities and psychological activities, and establish a personalized resource database for students, so as to have a basic judgment of students through their basic information, and then give more detailed and appropriate psychological well-being consultation corresponding to related health problems. When determining the number of iterations of the model, the number of manually labeled samples is very small. If too many iterations are carried out, the FNN network will be over-fitted; If the quantity of iterations is too small, the accuracy of the model cannot meet the requirements.

Based on the classical learning and classification ability of the neural network, a prediction model of university students' psychological and emotional fluctuation is established, and a large quantity of classified samples are used to train the network, so that the network system with analytical ability can objectively evaluate psychological well-being. The multilayer neural network model is shown in Fig. 2.1.

The agents in the multi-agent system cooperate to achieve the common goal. When some agents in the system fail, there is usually no problem in the whole system, mainly because each agent has certain learning ability. When conducting big data analysis, we should grasp the dimensions of investigation and conduct psychological big data analysis on students under the premise of ensuring that students are not disgusted, so as to achieve a win–win goal. At the same time, students can be judged deeply

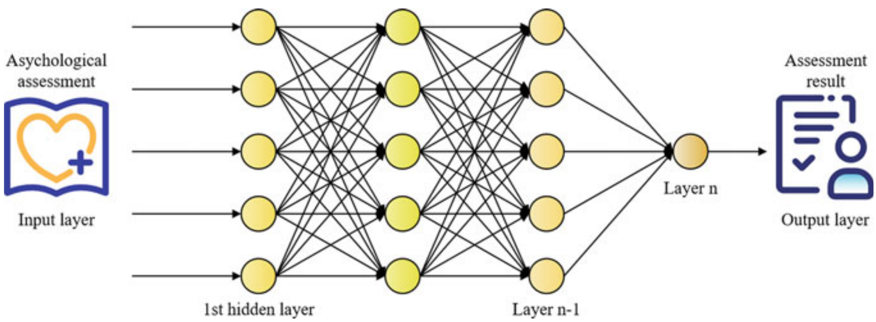


Fig. 2.1 Multi-layer neural network model

through high-quality questionnaires, so as to achieve double control of students. Only in this way can students' psychological problems be effectively controlled and monitored.

In this paper, the core indicators and auxiliary indicators are combined to evaluate the psychological and emotional fluctuations of university students more comprehensively and objectively:

$$w^* = \sum_{i=1}^{n} \alpha_i^* y_i x_i \tag{2.1}$$

x is the subjective factor of college students' psychological influence, y is the objective factor of college students' psychological influence, and n represents the number of terms in the summation equation. Select the Langrangian multiplier α_i^* corresponding to a support vector, and calculate the b^* as:

$$b^* = y_j - \sum_{i=1}^{n} \alpha_i^* y_i \left(x_i, x_j\right) \tag{2.2}$$

Substitute w^*, b^* into $wx + b = 0$ to get the optimal hyperplane equation:

$$\sum_{x_i \in SV} y_i \alpha_i^* (x_i, x) + b^* = 0 \tag{2.3}$$

Finally, the optimal classification function for this classification problem is obtained as:

$$f(x) = \operatorname{sgn}\left[\sum_{x_i \in SV} y_i \alpha_i^* (x_i, x) + b^*\right] \tag{2.4}$$

Whether it is the construction of student team or the psychological crisis identification system, it is the early stage of psychological crisis intervention. Once a crisis occurs, the emergency in the event and the aftermath are important links to determine the impact of the crisis [12]. For schools, in order to ensure the integrity of data, it is necessary to expand the indicators of psychological pre-alarm, so as to have a more in-depth and comprehensive understanding of students and the data support will be more convincing. Information processing also plays an important role in the analysis of psychological problems, so psychological information should also be effectively summarized and analyzed, so as to have a more accurate positioning of psychological well-being.

Set the factor set U and the assessment grade set V of the analysis object of psychological and emotional fluctuation:

$$U = \{u_1, u_2, \ldots, u_m\} \tag{2.5}$$

$$V = \{v_1, v_2, \ldots, v_m\} \tag{2.6}$$

Fuzzy assessment is carried out on each factor in U according to the grade index in the assessment set, and the assessment matrix is obtained:

$$R = \left(r_{ij}\right)_{n \times m} \tag{2.7}$$

In which r_{ij} indicates the degree of u_i's membership in v_i. After determining the importance index of each factor, record it as:

$$A = \{a_1, a_2, \ldots, a_m\}, \quad \sum_{i=1}^{n} a_i = 1 \tag{2.8}$$

Synthetic:

$$\overline{B} = AR = \left(\overline{b}_1, \overline{b}_2, \ldots, \overline{b}_m\right) \tag{2.9}$$

In this way, students' psychological status can be evaluated.

For data processing, we should also follow certain processing steps and methods. First, we can simulate the psychological information of students through the network information platform, and give a certain similar model after the simulation; Then, through other information about students' psychological well-being, students can be positioned to a deeper degree, which is mainly manifested in the analysis of students' daily activities and academic achievements, so as to realize the all-round information grasp of students.

2.3 Result Analysis and Discussion

Considering that psychological problems are difficult to detect by themselves and students generally have resistance to psychological consultation and investigation, this paper uses the method of text information analysis to identify psychological problems. The integration experiment is based on the unit experiment, which is a kind of overall comprehensive test that combines the functional modules according to the design requirements and then carries out step by step. When setting the model parameters, it is considered that the quantity of samples of different categories in the data set is quite different. However, the classification accuracy of the deep learning network will gradually deteriorate with the increase of the imbalance of sample number, so the weight of different types of samples is distinguished in this paper.

In order to verify the reliability and availability of the proposed student psychological prediction algorithm, experiments were carried out. The experiment is designed by Matlab. The simple semantic unit set is 200, the sample training set is 100, the number of iterations is 300, the semantic attribute set is 60, the number of instances

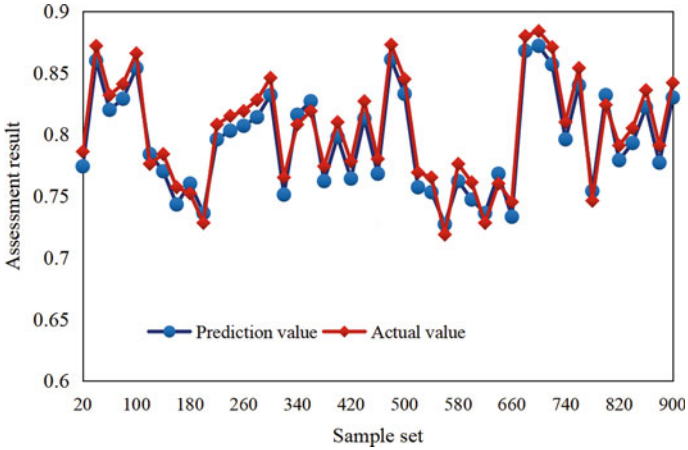


Fig. 2.2 FNN learning results

of similarity semantic feature distribution is 160, the number of deep learning is 300, and the convergence step is 20. Select appropriate neural network parameters and establish an appropriate prediction model of university students' psychological and emotional fluctuations. Use the sample set to train the network, train the network weight, and test the network performance. Finally, find the sample data to process and verify whether the expected goal can be achieved. Compare the output data of the psychological fluctuation prediction model with the real students' psychological data, as shown in Fig. 2.2.

There will be a lot of data in the network, which can truly reflect the psychological health of university students, so we can establish a psychological database of university students to collect, store, obtain and share more data resources.

By integrating the analyzed data, a more specific psychological well-being model and psychological crisis model are given to realize the pre-alarm and intervention of students' psychological crisis, so as to realize the real-time monitoring and grasp of students' psychological problems and promote the all-round development of students' mental and physical health. Pretreatment of unsatisfactory data sources can greatly improve the execution efficiency of Fuzzy C-clustering (FCM) and the process of knowledge discovery. The performance comparison results of the prediction model of university students' psychological fluctuation after psychological data processing are shown in Fig. 2.3.

Establishing psychological archives database is a long-term project, in which we should constantly update, collect and screen in time, scientifically analyze, summarize and evaluate the psychological health status of university students. Instead of the cumbersome traditional information processing methods, we should actively use Internet technology to process psychological well-being data, so as to speed up the processing of students' psychological problems, and effectively integrate students'

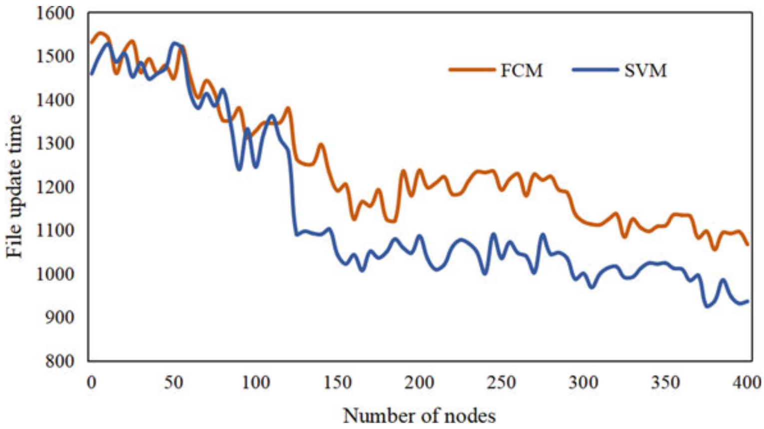


Fig. 2.3 Performance comparison of algorithms

psychological problems according to their behavior and learning characteristics, thus forming more effective psychological well-being information.

Under the background of normalization of epidemic prevention and control, the risk of students' psychological crisis increases, so universities need to investigate all students' psychological crisis to prevent unexpected situations. For the students themselves, each student's psychological status and psychological problems are different, so it is possible to file a psychological pre-alarm for each student. In this way, a relatively perfect psychological pre-alarm file can be made for every student, which can not only give corresponding answers to students' psychological problems, but also give corresponding preventive measures for students' psychological development. In the information education mode driven by big data, the accuracy of different algorithms in university students' psychological fluctuation prediction is shown in Fig. 2.4.

When the quantity of samples reaches a certain level, the accuracy of this psychological fluctuation prediction model will be improved, while when the quantity of samples is too small, the prediction accuracy will be greatly reduced. Compared with the traditional psychological prediction algorithm, this algorithm has obvious advantages. Therefore, the proposed algorithm can complete the whole process of university students' psychological well-being pre-alarm, and the overall operation is good. As far as university students are concerned, their personal subjectivity is strong, so the ways of handling things and the behaviors they encounter in different periods will be very different. Therefore, for the monitoring of university students' psychological problems, information should be dynamically detected. Through the analysis of big data, teachers can also understand the commonness and individuality of university students' psychological problems, improve the efficiency of dealing with psychological problems, and promote the development of psychological well-being education courses in universities.

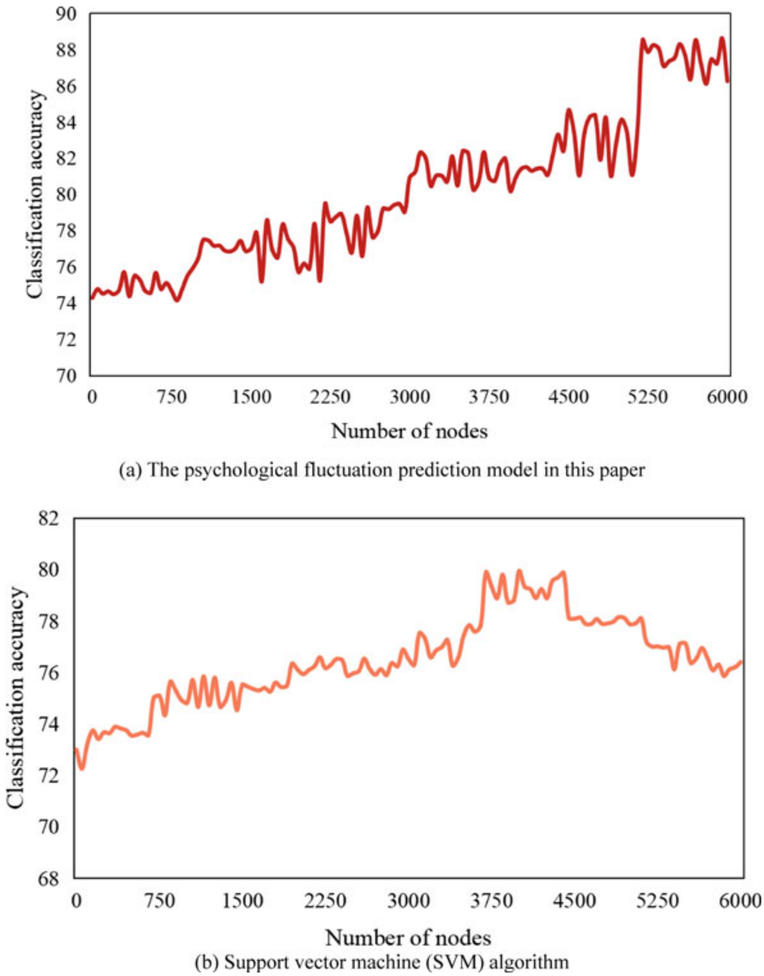


Fig. 2.4 Prediction accuracy of different algorithms in university students' psychological fluctuation prediction

2.4 Conclusion

From the perspective of developmental psychology, university students are at the critical stage of physical and mental development, and they are easily influenced by adverse factors. Coupled with the individual's wrong cognitive style and bad coping style, psychological distress will gradually turn into a psychological crisis that affects individual's normal development. Traditional psychological well-being education pays more attention to prevention and treatment of university students, amplifies their negative psychological emotions, and mainly emphasizes their psychological emotions of depression, anxiety and fear. This paper puts forward a prediction model of university students' psychological mood fluctuation based on FNN, and completes the pre-alarm of university students' psychological well-being through the cooperation among agents such as psychological data collection and psychological well-being evaluation. The results show that compared with the traditional psychological prediction algorithm, this algorithm has obvious advantages. Therefore, the proposed algorithm can complete the whole process of university students' psychological well-being pre-alarm, and the overall operation is good. Carrying out positive psychological well-being education for university students can effectively reduce the impact of external environmental impact on university students' psychological and emotional fluctuations, thereby reducing the probability of psychological crisis. Schools should vigorously support university students to set up psychological well-being education organizations, take activities as the carrier, and promote exchanges and interactions between peers through games, salons, writing and other colorful forms to achieve the purpose of peer counseling.

Acknowledgements Research on the Conduction and Influence Mechanism of External Environment Impact on University Students' Psychological and Emotional Fluctuations (A2021002).

References

1. Wong, L.W., Tan, W.H., Lee, V.H., et al.: Psychological and system-related barriers to adopting blockchain for operations management: an artificial neural network approach. IEEE Trans. Eng. Manag. **2021**(99), 1–15 (2021)
2. Christopher, R.B., Torres, W.: Aligning an undergraduate psychological medicine subject with the psychological well-being needs of the local region. BMC Med. Educ. **18**(1), 118 (2018)
3. Jondani, J.A.: Locus of control in university students with and without visual impairments, and the visual characteristics that affect it. J. Vis. Impairment Blind. **115**(1), 42–54 (2021)
4. Barr, M.L.: Encouraging university student active engagement in learning: student response methods and anonymity. J. Comput. Assist. Learn. **33**(6), 621–632 (2017)
5. Jia, J., Li, D., Li, X., et al.: Psychological security and deviant peer affiliation as mediators between teacher-student relationship and adolescent Internet addiction. Comput. Hum. Behav. **73**(8), 345–352 (2017)
6. Harris, C., Linder, C.: The racialized experiences of students of color in higher education and student affairs graduate preparation programs. J. Univ. Stud. Dev. **59**(2), 141–158 (2018)
7. Perez, R.J.: Enhancing, inhibiting, and maintaining voice: an examination of student affairs graduate students' self-authorship journeys. J. Univ. Stud. Dev. **58**(6), 833–852 (2017)
8. Kim, M., Oja, B.D., Han, S.K., et al.: Developing student-athlete school satisfaction and psychological well-being: the effects of academic psychological capital and engagement. J. Sport Manag. **34**(4), 1–13 (2020)
9. Li, N., Zhai, H., Seetharam, T.G., et al.: Psychological health analysis based on fuzzy assisted neural network model for sports person. J. Intell. Fuzzy Syst. **2021**(5), 1–15 (2021)
10. Calvo, H., Gutiérrez-Hinojosa, S.J., Rocha-Ramírez, A.P., et al.: Psychological attachment style prediction based on short biographies. J. Intell. Fuzzy Syst. **39**(2), 2189–2199 (2020)
11. Karasawa, Y., Uemae, M., Yoshida, H., et al.: Prediction of clothing comfort sensation of an undershirt using artificial neural networks with psychophysiological responses as input data. Text. Res. J. **92**(3–4), 330–345 (2022)
12. Kobayashi, K., Matsushita, S., Shimizu, N., et al.: Automated detection of mouse scratching behaviour using convolutional recurrent neural network. Sci. Rep. **11**, 658 (2021)

Chapter 3
Research and Application of Distribution Network Digital Twin Spatial Layout Technology

Gang Wang, Lin Peng, Yanfang Mao, Aihua Zhou, Min Xu, He Wang, Zhujian Ou, and Jianguang Yao

Abstract This paper studies the layout of the electrical topological relationship of the digital twin space. Based on the geographical resource map and the electrical topological map, the distribution network topological nodes are extracted using the single-line diagram and the ring network diagram tree structure. According to the hierarchical mark of the equipment, the single-line diagram and the ring network diagram structure, combined with the point cloud model, the distribution network digital twin space is rapidly constructed using the automatic generation algorithm of the distribution network three-dimensional model diagram. In this paper, based on the point cloud model, we carry out the research on the three-dimensional correction technology of distribution network equipment, build the point cloud recognition model library of fixed equipment, automatically identify and correct the outdoor overhead line tower, station building and other equipment through the point cloud model, identify the contents of the objects in the station building and underground cable line according to the prior knowledge, and correct the three-dimensional model position based on the three-dimensional model and the point cloud model. This paper develops the spatial layout component of the distribution network digital twins that integrates geographic, physical, management and business information, and realizes the spatial layout and rapid reconstruction of the distribution network digital twins based on electrical topology.

G. Wang (✉) · L. Peng · A. Zhou · M. Xu · H. Wang
State Grid Smart Grid Research Institute CO., LTD., Nanjing 210003, China
e-mail: 26261045@qq.com

State Grid Laboratory of Power Cyber-Security Protection and Monitoring Technology, Nanjing 210003, China

Y. Mao · Z. Ou · J. Yao
Nantong Power Supply Company, State Grid Jiangsu Electric Power Co., Ltd., Nantong 226006, Jiangsu, China

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_3

3.1 Introduction

Under the background of the construction of new power system, the pressure of the distribution side consumption and the distributed new energy grid connection service is increasing, the operation control form is evolving to the active distribution network, the topology analysis and model construction is difficult, inefficient, and costly, and the distribution network digital resource sharing and maintenance, multi-dimensional data simulation and analysis capabilities are insufficient [1, 2]. It is urgent to improve the level of intelligent operation, lean operation and maintenance, and efficient operation of the distribution network by digital twinning technology. The distribution network facilities are characterized by diversified types, massive scale and complex structure [3, 4]. The traditional digital twins of distribution network facilities tend to static three-dimensional models, while the conventional construction methods such as point cloud measurement and CAD design have problems of high cost and slow speed. High-cost local area construction modeling is difficult to support large-scale system-level applications [5, 6]. The distribution network digital twin spatial layout technology needs to be studied to solve the above problems.

3.2 Automatic Layout Technology of Power Grid Diagram

At home and abroad, there is little research on the automatic layout technology of power grid diagram, and it is only limited to the transmission network with relatively simple structure [7, 8]. However, the use of GIS (geographic information system) and PMS (generation management system) systems as the source of diagram data for power grid dispatching services is also at the experimental stage in the industry, and there is no mature and unified standard to follow [9, 10]. In recent years, GIS-based topology analysis can not only realize traditional topology analysis, but also realize power point tracking analysis, power supply radius analysis, optimal repair path analysis, dynamic coloring, etc. [11, 12]. However, at present, the automatic layout technology of power grid diagrams and models lack three-dimensional space support and point cloud correction, which leads to the inability to achieve accurate spatial analysis function of distribution network.

3.3 Layout of Electrical Topological Relationship in Digital Twin Space

Under the dual-carbon background, the new power system needs to support the two-way interaction between power supply and load, and needs to build a digital twin-space covering the electrical topology to support the electrical analysis, spatial analysis, comprehensive analysis and other functions of the power grid. The digital twin space needs to be built on the basis of the distribution network electrical topology map, combined with geographical map, 3D model and point cloud model to build a digital twin space with mutual integration of models. The line type of the distribution network electrical topology diagram system is divided into overhead line, ground cable line and overhead ground cable hybrid line. There are many multi-circuit station building equipment in the underground cable line, such as switching post, ring network cabinet, distribution room, box transformer, cable branch box, high voltage user, etc. These station building equipment are independent elements on the electrical topology diagram, and there are detailed electrical wiring diagrams inside these independent elements.

In the traditional electrical topology diagram system, each power equipment is associated with a point on the topology description, and the relevant power equipment includes distribution transformer, switch, grounding switch, etc. The relationship between equipment is connected by lines. The line includes the power line and the connecting line inside the station building. In order to build a digital twin space that integrates geographic model map, electrical topology map, 3D model and point cloud model, it is necessary to automatically generate a 3D model based on the electrical topology map, geographic resource map and point cloud model. The automatic generation of 3D models requires searching the electrical topology map and geographical resource map, and constructing the distribution network single-line diagram and ring network diagram topology structure based on the digital twin space.

The search algorithm is shown in Fig. 3.1. The search process for a distribution network line is based on the node corresponding to the outgoing line switch (for switching station) or distribution line outgoing line (for substation) in the electrical topology map or geographical resource map. On the electrical topology map, search all relevant lines, equipment and topology data from the load side (including distributed power supply) along the dotted line, and stop the search of this line until it meets the end of the line or the contact switch, Search the electrical topology data of all selected distribution line circuits. During the search process, the data information of tie switch is recorded in order to generate a digital twin space for multiple tie associated feeders. According to the electrical topology data of the distribution line that has been searched, the single-line diagram and ring network diagram tree structure of the line are established.

After establishing the tree structure of single-line diagram and ring network diagram, it is necessary to set the coordinates of each device in the tree object in 3D space. To set the coordinates of each equipment in three-dimensional space, it is necessary to determine whether the equipment is a trunk line or a branch line in

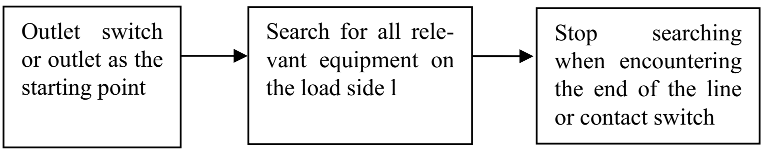


Fig. 3.1 Flow chart of distribution network route search

the electrical topology diagram, and whether it is an equipment outside the station building or an equipment inside the station building. Therefore, it is necessary to integrate the equipment in the established tree model. The final result of integration is to establish the level of each element in the equipment list. The integration basis is the search order of the equipment in the in-depth search from the distribution line outlet to the load side (including distributed power supply) based on the topological relationship.

The layout of the 3D model needs to solve two problems: the routing mode of critical path nodes and the layout of branch line equipment. Determine the latest digital twin space layout according to all equipment on the critical path in combination with the geographical resource map and point cloud scanning map. After determining the routing mode, the coordinates of all equipment and lines on the critical path of the geographical resource map in the digital twin space can be determined. According to the direction from left to right, from top to bottom, and according to the topological relationship between the devices, the device nodes on the critical path are determined according to the needs of digital twinning analysis. If the trunk line contains resources in the station building, then when the specific layout is carried out, a station building area is cut out from the trunk line by combining the point cloud model, and placed below or above the trunk line, and connected with the digital twin space of the station building. When drawing the branch line part, set the 3D position of the branch line equipment based on the spatial layout relationship of the point cloud model. Before drawing the branch line equipment, it is necessary to clarify the digital twin space position of the trunk line. Poll the branch equipment to be drawn in the single-line diagram, calculate the maximum number of nodes of the branch, and confirm that the branch is drawn above or below the trunk line. Set the coordinates of each device on the branch in order from parent to child.

3.4 Automatic Generation Algorithm of Distribution Network 3D Model Drawing

The electrical topology constructed based on the layout of electrical topology relationships in digital twin spaces, combined with point cloud models, 3D models, and geographic models, is automatically generated into a distribution network 3D model diagram based on feature matching, layout, and routing algorithms. The processing logic of the 3D model graph generation algorithm is shown in Fig. 3.2.

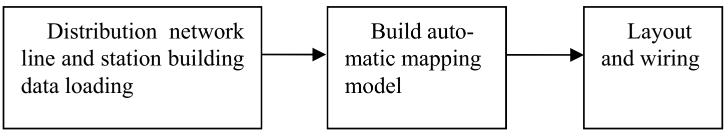


Fig. 3.2 Algorithm diagram of 3D model drawing generation

Distribution network line and station building data loading: first, the initial loading process of the distribution network line, station building data, point cloud model, 3D model, and geographic model is used to build a memory object model for subsequent automatic mapping. In this process, the topology fusion and simplification of the distribution network equipment 3D resources, line model, and station building wiring diagram and other models are required to support efficient mapping processing in the subsequent automatic generation process. Specifically, match 3D model resources based on the line model and equipment types and models in the station building wiring diagram, and establish a mapping relationship between equipment and 3D models on the topology diagram.

Establish automatic mapping model: realize the creation of 3D model space and all 3D equipment entity objects of the distribution network based on the data structure in the middle memory, and establish the association relationship with the source objects of the distribution network electrical topology diagram and internal wiring diagram. Based on the logical relationship of distribution business, point cloud model, geographic model and electrical topology model matching model are established, mainly to achieve feature matching. Neural network is used for model training, and finally the distribution network digital twin mapping model is constructed.

Layout and wiring: based on the geographical model, electrical topology model and point cloud model, the distribution network digital twin mapping model matching algorithm is adopted to realize the initial layout and wiring of the 3D model diagram through the layout and wiring algorithm. After completion, the 3D model diagram is saved to realize the complete initialization and generation process of the 3D model diagram.

3.5 3D Correction Technology of Distribution Network Equipment Based on Point Cloud Mode

The original 3D point cloud only contains geometric information, which can only meet the simple measurement and analysis functions, and is difficult to meet the deeper level of intelligent services and applications such as distribution network planning and spatial analysis. The semantic 3D model of the entity not only contains necessary geometric information, but also has rich semantic information and topological association information. The automatic extraction of distribution network equipment is a necessary and important step to automatically reconstruct the semantic rich entity model from the point cloud. Based on the constraints of distribution network topology prior knowledge, this paper extracts the semantics of the distribution network point cloud model, and further extracts the geometric information of the identified semantic objects. The partition and extraction of space is the process of dividing space into semantic regions. Each partition space represents a distribution network device. The partition and extraction of space can determine the ownership

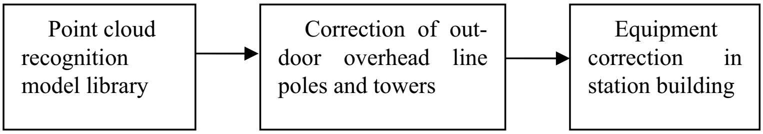


Fig. 3.3 Point cloud modification algorithm diagram

of each component, and determine the semantic relationship between distribution network devices by judging the connectivity of each space.

The method used in this paper is mainly to build a point cloud identification model library for fixed distribution network equipment, automatically identify the shape of the cable in the distribution network room and the ground through the point cloud model, correct the outdoor overhead line tower based on the 3D model and the point cloud model, identify the content of the object in the station room and the ground cable based on the distribution network topology prior knowledge, and correct the 3D model position based on the distribution network business knowledge, 3D model and the point cloud model, as shown in Fig. 3.3:

1. Take all the overhead line, underground cable and overhead cable point cloud data as samples, use the 3D depth learning network to extract depth features from each point cloud data, cluster all the sample depth features through K-mean (K-mean algorithm) algorithm, and obtain corresponding cluster centers. At the same time, regard them as words, and combine these words together to form an overhead line, underground cable and overhead cable point cloud dictionary.
2. Point cloud correction of indoor and underground cables and outdoor overhead lines: The point cloud model of overhead lines, underground cables and overhead cables is the point cloud information scanned by UAV, and the frame of the point cloud model of the station building and underground cables is defined in advance, and then the approximate point cloud information of the station building, underground cables and overhead lines is determined by combining the 3D model; The main model of the overhead line tower is determined according to the three-dimensional model, and then the point cloud model is used to correct the conductor position, circuit model, and tower type. Through this step, the specific three-dimensional model of the station building, ground cable shape, and the overhead line tower is determined.
3. Object correction of indoor and underground cables: this paper uses the algorithm based on region growth to segment and extract the plane of point cloud. The algorithm first sorts the points according to the curvature of each point in the point set, and takes the points corresponding to the minimum curvature as the initial seed points of region growth; Define the K neighborhood search range, search for points within the neighborhood of the seed point, if the angle between the normal direction of the point in the neighborhood and the normal direction of the current seed point is less than the angle threshold θt. Then add the neighborhood point to the current plane area and check the curvature value of all points in the neighborhood. If the curvature value of the neighborhood point is less than

the curvature threshold *Ct*, the neighborhood point will be grown as a new seed point; Repeat the above steps until the seed point set is empty, and complete the region growth algorithm.

The indoor plane point sets are extracted by the region growth algorithm, and the plane parameters need to be estimated according to the plane point sets. The commonly used plane fitting algorithms include the least square fitting algorithm and the random sampling consistent plane fitting algorithm.

The following prior constraints are mainly involved in the process of plane marking of indoor components (ceiling, floor, wall):

(1) The normal direction is divided by calculating the angle between the normal direction of the fitting plane and the gravity direction. Generally, the normal direction of the indoor ceiling and floor plane is collinear with the gravity direction, and the normal direction of the wall plane is perpendicular to the gravity direction;
(2) Area and size: in the case of the same floor height in the multi-indoor space scene, the floor point cloud plane is usually the largest plane, and the area, height and width of the ceiling plane and the wall plane also meet certain constraints;
(3) Plane position: the interior ceiling and floor plane are generally located at the highest and lowest plane of each room, and the wall plane is perpendicular to the ceiling and floor plane, forming a closed space.
(4) The transformer is located in the middle of the room, and the main body is composed of several cuboids. There are several types of cylindrical bushings on the top, which are connected with conductors;
(5) The switch body is composed of several cuboids and connected with wires.

With the help of the above priori knowledge, the extracted plane is marked and identified, and the point cloud of transformer, wire, switch, reactive compensation box, fuse and DC screen is proposed. The point cloud space needs to be further divided to determine the attribution of each component.

3.6 Distribution Network Digital Twin Spatial Layout Component

Integrate geographic, physical, management and business information, and develop distribution network digital twin spatial layout components. Support the spatial layout and scene integration of twin objects based on semantic understanding, realize the spatial layout and rapid reconstruction of distribution network digital twins based on electrical topology, and support the interaction between distribution network equipment and topology network.

The specific functions are as follows: distribution network data acquisition tool, which carries out the collection and regulation of distribution network account and measurement data, including basic data, operation monitoring data, maintenance test data, automation data, information communication data, environmental data, etc. Two-dimensional and three-dimension model association, using the configuration mapping association model of two-dimensional topology and three-dimensional model, realizes the interaction and deduction simulation between the distribution network equipment twins and the topology network. The space constraint rule base mainly includes the space constraint of the transformer, fixed switch cabinet, channel width, single row assembled capacitor bank, double row assembled capacitor bank, and a complete set of capacitor cabinet. The object topology linkage rules, the mapping linkage rules of 2D entities and 3D models, realize the linkage operation and deduction simulation of 2D topology wiring diagram elements and 3D twin models. Geospatial engines, build 3D GIS engine, and use open source framework to realize grid 2D topology, 3D rendering and spatial interaction. The 3D twin engine adopts a lightweight cross platform H5 architecture, supporting browser based 3D models and a fast rendering visualization engine for large scenes. The holographic topology engine realizes the construction technology and interactive mapping of standard entity libraries, graphic editing and topology generation, real-time verification of line topology and station topology business rules, association of topology nodes and ledger data, interactive analysis and access of power points, open capacity calculation, economic analysis of wiring, and load characteristic analysis. The topology constraint rule library mainly includes two types: business specification and data specification. The business protocol library includes islands, no power supply, connectivity, etc.; The data specification library includes constraint rules such as null value constraints, duplicate constraints, including relationship constraints, excluding relationship constraints, threshold constraints, temporal data anomalies, etc.,

The overall architecture of the system is mainly divided into three parts: browser client, cloud storage, and high-performance service cluster.

The browser client local cache and cloud cache constitute a multi-level cache system. Browser client is an important user-oriented interactive window for data visualization, mainly composed of user interface, Cesium rendering engine and local data cache. The user interface is implemented in HTML, JavaScript and other programming languages, and Cesium is used for distribution network digital twin space visualization and interaction. The user triggers the client to send a request to the server by clicking or checking the layer resource and other interactive operations, and renders the data responded by the server in the browser. In addition, the local data cache is mainly the data cache pool set based on the browser's cache mechanism, which is an important part of the multi-level cache system. According to the user's interaction, the browser first requests the local cache data, and then sends the data request to the server if the data does not exist in the local cache. After receiving the server response data, add it to the local cache to reduce the response time of multiple requests for data.

Cloud storage and browser local storage form a multi-level cache system. Multi-level cache is an effective method to improve the efficiency of large-scale data loading.

The high-performance server cluster mainly has the following two functions.

(a) Data preprocessing: The main function is to convert data formats that are not supported or recommended by the rendering engine into data formats that the rendering engine can load and render.
(b) Data storage: including the file system and database system. The file system mainly stores electrical topology map, point cloud model, 3D model resources, etc.; The database system mainly stores the detailed information corresponding to 3D model resources.

3.7 Conclusion

Under the background of a new power system construction, the distribution network facilities have the characteristics of diversified types, massive scale and complex structure. The traditional digital twin of distribution network facilities tends to static three-dimensional models, while the conventional construction methods such as point cloud measurement and CAD design have the problems of high cost and slow speed, and the high-cost local area facility modeling is difficult to support large-scale system-level applications. This paper studies the layout of the electrical topological relationship of the digital twin space. Based on the geographical resource map and the electrical topological map, the distribution network topological nodes are extracted using the single-line diagram and the ring network diagram tree structure. According to the hierarchical mark of the equipment, the single-line diagram and the ring network diagram structure, combined with the point cloud model, the distribution network digital twin space is rapidly constructed using the automatic generation algorithm of the distribution network three-dimensional model diagram. In this paper, based on the point cloud model, we carry out the research on the three-dimensional correction technology of distribution network equipment, build the point cloud recognition model library of fixed equipment, automatically identify and correct the outdoor overhead line tower, station building and other equipment through the point cloud model, identify the contents of the objects in the station building and underground cable line according to the prior knowledge, and correct the three-dimensional model position based on the three-dimensional model and the point cloud model. This paper develops the spatial layout component of the distribution network digital twins that integrates geographic, physical, management and business information, and realizes the spatial layout and rapid reconstruction of the distribution network digital twins based on electrical topology.

Acknowledgements This work was financially supported by the science and technology project to State Grid Corporation "Research on rapid construction and linkage analysis technology of distribution network facilities digital twins integrated with electrical topology (5108-202218280A-2-396-XG)".

References

1. Yan, Y.H., Zhou, R.X., Yu, W., et al.: Station topology identification based on carrier wave and power frequency distortion technology. Autom. Technol. Appl. **40**(10), 68–71 (2021) (in Chinese)
2. Yan, W.G., Wang, G.M., Lin, J.K., et al.: Verification method of low-voltage distribution network topology based on AMI measurement information. China Power **52**(2), 125–133 (2019) (in Chinese)
3. Que, P.X., Gong, X.R., Song, Z.Y, et al.: A Complex Network Topology Editing and Visualization System Based on WEB: cn110929201a (2020)
4. Feng, R.H., Zhao, Z., Xie, S., Huang, J.L., Wang, W.: Low voltage distribution network topology identification method based on principal component analysis and convex optimization. J. Tianjin Univ. **54**(7), 746–753 (2021) (in Chinese)
5. Li, Y.Z., Zhao, Q.S., Liang, D.K., et al.: Construction of low-voltage distribution network impedance topology model based on measurement information. Power Grid Clean Energy **37**(4), 15–22 (2021) (in Chinese)
6. Weng, Y.X., Ma, W.Z., Shi, J., et al.: Auxiliary decision-making of power grid operation mode based on semi dynamic topology optimization. Power Eng. Technol. **40**(4), 8 (2021)
7. Li, Z., Chang, V., Ge, J., Pan, L., Hu, H., Huang, B.: Energy-aware task offloading with deadline constraint in mobile edge computing. EURASIP J. Wirel. Commun. Netw. **1**, 1–24 (2021)
8. Zhu, Y.H., Liu, S.X., Ji, L.X., et al.: A sequential link prediction algorithm integrating local topological influence. J. Electron. Inf. **44**(4), 13 (2022)
9. Lei, Y.: Feudal lake topological optimization of material nonlinear substructure based on parametric level set method. Appl. Math. Mech. **42**(11), 11 (2021)
10. Lou, W.J., Wang, L., Jiao, T.C., et al.: Voltage stability preventive control method based on network topology optimization. China South. Power Grid Technol. **16**(1), 8 (2022)
11. Yao, Y., Xu, J.Y., Cui, Q.Y.: Communication topology optimization of distributed economic dispatching in active distribution network. Energy Eng. (2), 5 (2020)
12. Shan, S.R., Wang, J.L.: Topology optimization method of wireless monitoring network based on GSO algorithm. Smart Power **48**(6), 7 (2020)

Chapter 4
Enterprise Accounting Decision Support System Based on Deep Learning Algorithm

Jiangyan Cheng

Abstract In the traditional financial decision analysis system, these unstructured data cannot be obtained and processed, and the data cannot be closely related and continuously improved. This article proposes a financial data mining (DM) algorithm based on deep learning (DL), which provides technical support for the construction of the enterprise accounting decision support system (DSS). Today, with the wide popularization of network informatization, the supervision and management functions of financial and auditing departments are becoming more and more important. Only seamless inspection by the financial department can ensure the accurate performance of the management process, only non-isolated statistics can generate a true description of the financial operation of each unit, and only the implementation without time difference can produce the efficiency significance of control in the matter.

4.1 Introduction

With the informationization and economic marketization of enterprises, enterprises are facing huge market competition, which inevitably increases the pressure of enterprise operation and survival [1]. Enterprises want to grow and develop without the participation of financial management, and the specific contents of financial management work by management mainly include analyzing enterprise financial indicators, monitoring and reducing enterprise financial risks, and making financial decisions [2]. To construct the technical principle of the intelligent financial application scenario, we can discuss it from two dimensions: commercial application principle and data technology principle [3, 4]. Because enterprises are facing increasingly fierce competition, the traditional decision-making methods can no longer meet

J. Cheng (✉)
Fujian Institute of Information Technology, Fuzhou 350002, Fujian, China
e-mail: 790693032@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_4

the requirements, and leaders demand to develop high-level computer application systems, and it is imperative to develop a decision-supporting financial software [5].

In the field of financial management, due to the form of information provided, the timeliness of information provided and the support of information provided for decision-making, the usual accounting treatment methods and financial analysis methods can not well support the analysis and decision-making of complex problems [6, 7]. At present, the increasingly fierce market competition and the rapid growth of information technology, internet and e-business constitute the macro environment for the survival and growth of enterprises, which makes the traditional financial management face new challenges, so it is urgent to improve the decision-making level of enterprise financial management as soon as possible [8, 9]. This article puts forward a financial DM algorithm based on DL, which provides technical support for the construction of enterprise accounting DSS.

4.2 Methodology

4.2.1 Intelligent Financial Decision Support System (IFDSS) and Data Mining

The function of an intelligent financial decision support system should be based on the function of a traditional financial decision support system, and deeply analyze and mine historical, realistic, clear, fuzzy, internal, customer and other multi-level and multi-angle information by applying data mining technology. The overall design of the system is based on the analysis of the system function, dividing the system theme according to the system objectives, functions and environmental conditions, building the model, and then determining the overall functional module structure of the system, and defining the sub-objectives and sub-functions of each subsystem. Accounting behavior itself is a kind of data behavior, and its essence is the process of transforming an accounting phenomenon into a quantifiable form. The traditional double-entry bookkeeping method directly reflects the profit and loss of the business entity through data, which lays the foundation for recording with standard data [10]. Different levels of dataization have different emphases. The front-end intelligent financial robot and the middle-level intelligent financial assistant emphasize full and complete data collection, and continuously pass it back, while achieving a certain degree of analysis. The process of decision-making using the theory of management and advanced management ideas is through computer technology. Because the documents received by the finance department do not all have a unified format and reporting standards, the data collection cannot be based on the predefined logic of specific locations, but also needs to play the role of artificial intelligence (AI) cognitive technology to extract specific data independently [11]. Only in this way can the source of data collection be extended to unstructured or semi-structured documents, providing resources for higher-level machine learning, analysis and application.

The theme of financial analysis is to diagnose the financial situation of the enterprise according to the data of the enterprise, so as to grasp the lifeline of the economic operation of the enterprise, and it is an important basis for the relevant stakeholders, countries, investors, creditors and enterprises themselves to make relevant decisions. The function of DSS is to evaluate and analyse the business data according to the historical data accumulated by the daily basic business processing of enterprises, and to estimate and predict the development trend of the future business situation of enterprises by using mathematical, statistical or other intelligent technologies and methods [12]. The system needs to start with the business environment and business conditions of the enterprise, analyze the business and financial benchmark indicators and investment opportunities, find out the key factors that affect the investment efficiency of the enterprise, and then search for key projects and related methods to improve the investment efficiency of the enterprise according to the key factors, and finally realize the preparation and management of the investment business plan of the enterprise.

4.2.2 DL-Based Financial DM Algorithm

IFDSS is a concrete application of DSS in the financial field, and it is an inevitable trend of financial informatization after the rapid growth of market environment and computer information technology. Under the framework of enterprise data strategy and management policy, finance needs to conduct governance consultation on data architecture, data standards and master data, and conduct data demand management, data warehouse platform construction and application process establishment and assessment. The questions raised by users can be transformed into information language, forming data information that can be interpreted by computers [13]. Everyone who logs into the system will have a unique account number and password, as well as their own authority, and the artificial falsification of information can be curbed. IFDSS is a special intelligent DSS based on modern management theory and computer technology. It uses statistics, quantitative economics, data warehouse technology, AI and fuzzy mathematics and model technology to provide decision support through human–computer interaction. For investment decision support system, it is necessary to sort out and summarize the information systems related to investment decision, including personnel composition information, technical level information and performance information of the human resources information system, cost information and profit information of financial system, fixed assets and depreciation information of fixed assets system, production information, logistics information and personnel performance information.

As an assistant application for business personnel and individual financial personnel, intelligent financial assistant emphasizes rapidity and flexibility, which cannot be achieved by previous data processing [14]. This is the embodiment of the advantages of the data center, and it is also the fundamental reason why the establishment of intelligent finance must start from the digital transformation of financial

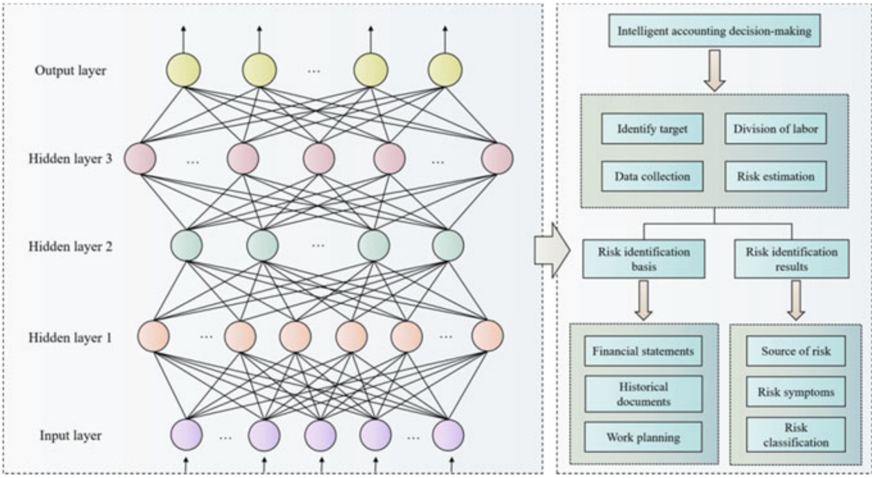


Fig. 4.1 DL model of enterprise accounting DSS

sharing services and the establishment of the data center. With the growth of business activities, financial data has been in the process of dynamic changes, so how to collect data efficiently when the management makes financial decisions is an urgent problem to be solved. The DL model of enterprise accounting DSS is shown in Fig. 4.1.

Enterprise investment decision-making is to use the raise funds to determine better or best investment projects in order to obtain investment income. When making investment decisions, enterprises need to analyze and evaluate the feasibility of investment schemes with specific indicators, including financial evaluation indicators, environmental evaluation indicators, risk evaluation indicators and the knowledge and experience of evaluation experts. The selection of models in the enterprise accounting decision-making system is to select from many models according to the characteristics of financial decision-making problems, and then link them as needed to form a composite model, so as to solve complex decision-making problems. Let the time series be $x_1, x_2, \ldots$, the moving average method can be expressed as:

$$F_{t+1} = (x_t + x_{t-1} + \cdots x_{t-N+1})/N \tag{4.1}$$

It is further derived as:

$$F_{t+1} = \frac{1}{N} \sum_{t-N+1}^{t} x_i \tag{4.2}$$

In the formula: x_t is the latest observed value, F_{t+1} is the predicted value of the next period. In order to calculate a moving average, it is needed to have N past values. After finishing, the calculation formula of the moving average method can be simplified as:

$$F_{t+1} = \frac{x_t}{N} - \frac{x_{t-N}}{N} + F_t \tag{4.3}$$

In order to make the system serve decision-making more effectively, enterprises should make use of some related emerging technologies in the computer field, such as DSS, neural network, data warehouse technology, DM, reasoning, etc., and further study and design functional financial DSS that meets the actual needs of enterprises to assist managers in making financial strategic decisions. Through data marts, users' behavior can be combined with the value created by finance through data, which is convenient for the financial department to clarify which services can be converted into benefits, thus providing more accurate performance assessment and feedback. Determine the data source by analyzing the decision-making requirements. Select the required data from the operable database, summarize and sort it out, and store it in different information layers of the data warehouse.

The amount of information in accounting decision-making is inversely proportional to the frequency of information, and the amount of information contained in different information can be accumulated. The Set sample set is written as:

$$S = \{s_1, s_2, \ldots, s_m\} \tag{4.4}$$

The sample categories are:

$$C = \{c_1, c_2, \ldots, c_k\} \tag{4.5}$$

Then the calculation formula of the sample information entropy is as follows:

$$\begin{aligned} H(S) &= \sum_{j=1}^{k} \sum_{i=1}^{m} p(s_{ij}) \log \frac{1}{p(s_{ij})} \\ &= -\sum_{j=1}^{k} \sum_{i=1}^{m} p(s_{ij}) \log p(s_{ij}) \end{aligned} \tag{4.6}$$

where $p(s_{ij})$ represents the probability that the sample point in the sample set belongs to the category c_j. Let the sample attribute set:

$$A = \{a_1, a_2, \ldots, a_t\}^T \tag{4.7}$$

Which has t different values. The sample set is divided into t sample subsets according to the attribute A, and the information entropy of the sample set is as follows:

$$H(S|A) = -\sum_{j=1}^{k} \sum_{i=1}^{t} p(a_{ij}) \log p(a_{ij}) \tag{4.8}$$

Among them, $p(a_{ij})$ represents the probability that the sample point belongs to the category c_i when the attribute A is a_i.

The establishment of the system completes the optimization and transformation of the traditional financial process, which makes the financial management mode more centralized, improves the efficiency of enterprise financial management and makes the decision more scientific and reasonable.

4.3 Result Analysis and Discussion

In the actual operation process, people find that it is not as easy as imagined to obtain useful information in the data ocean. Users of information use the data warehouse in a predictable and repetitive way. Information users instruct them what they want to know before using the data warehouse. They often visit the data warehouse regularly every day. In the process of visiting, they often only visit a small part of the data, and the results can often be obtained by visiting the data. The simulation operating system is Windows 11, the processor is Core i5 12400F, the graphics card is RTX 3060, the memory is 16 GB, and the hard disk capacity is 1 TB. After the data is stored in the data warehouse according to the unified format and different themes, the data can be expressed and accessed. According to the specific needs of enterprises for investment decision information, users can express and access the results in different ways through the purchased products and their supporting components. Analysts can directly calculate the investment payback period of the project in the non-discounted cash flow indicators, the net present value, the profitability index and the internal rate of return in the discounted cash flow indicators according to the needs of data analysis of specific different projects. The data outlier removal process is shown in Fig. 4.2.

The training samples are used to train the network and learn the weights after training. Then, the trained network is used to input the test network to verify the learning effect of the network. In the financial big data environment, the input sample is a full sample space, that is, all the data accumulated in production, circulation, consumption and management during the construction of financial big data. By training the whole sample space, the training results can be more accurate. Compare the output data of this model with the real accounting decision data, as shown in Fig. 4.3.

It is not difficult to see that the result of financial data forecast is convergent, which can approximate the original data well and has the basis for making decisions on accounting behavior. Different levels of applications have different algorithms for improving their capabilities and selecting them, which need to be selected and designed based on specific financial work scenarios. Knowledge diggers then use the data warehouse to explore why this effect occurs, and dig out the internal knowledge that produces these phenomena from the data warehouse. Intelligent financial applications need to learn and improve their performance by observing and adapting to their own environment, just like human beings in the process of dealing with

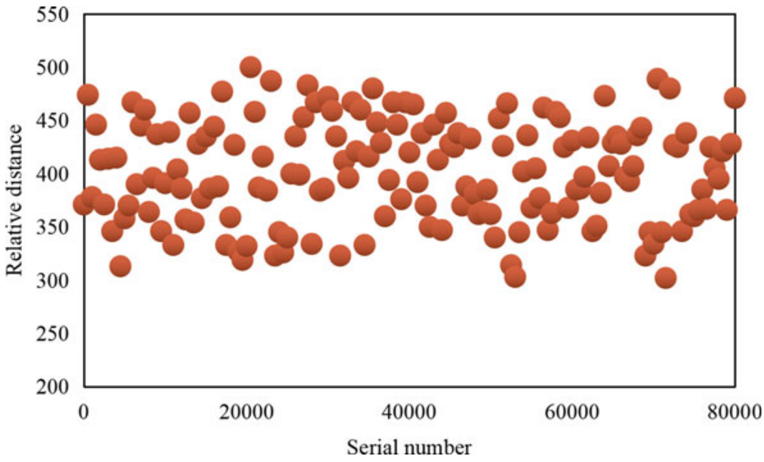


Fig. 4.2 Financial data processing to remove outliers

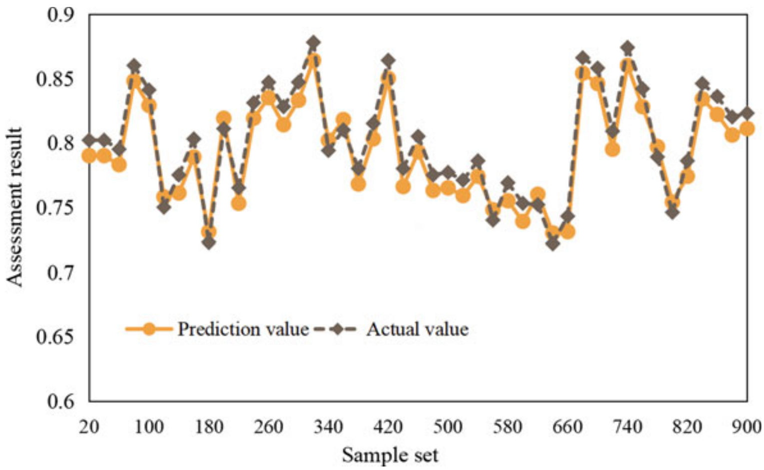


Fig. 4.3 Learning results of this model

repeated business and collecting data. Its data can be obtained from heterogeneous external data sources, online practical processing systems and other offline business data to support decision support and online analysis applications. Information users usually observe some general data or aggregate data, and seldom use some metadata or detailed data. From the nature of information users' work, they are often salesmen, who use some predefined queries to operate on general data and perform some simple processing. Compare the recall and accuracy of the algorithm in enterprise accounting decision, as shown in Figs. 4.4 and 4.5.

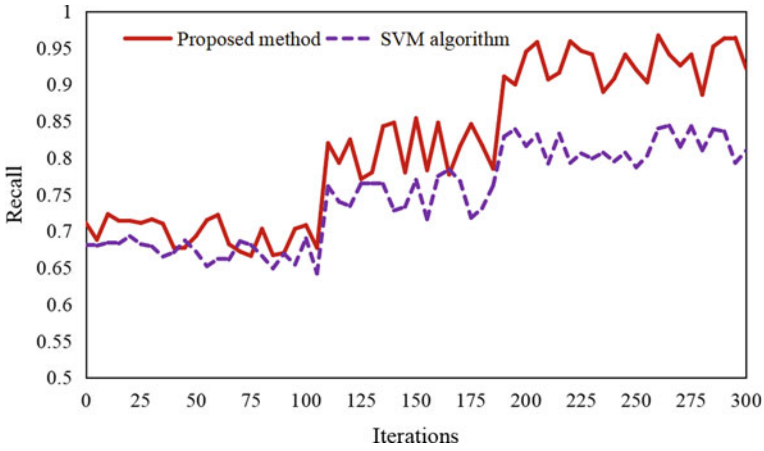


Fig. 4.4 Comparison of recall of enterprise accounting decision-making

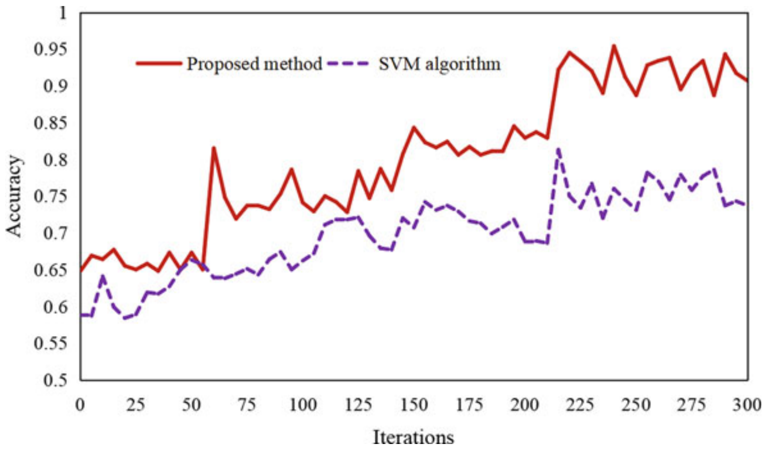


Fig. 4.5 Comparison of forecasting accuracy of enterprise accounting decision-making

When using data warehouse, information users often only involve the operation field of data warehouse. After submitting the query request, they often hope to get a response in a short time. Most of the queries submitted by these users are predefined queries. In order to complete these predefined queries, the designer of data warehouse must know the requirements, data structure and data content of these predefined queries during the design of data warehouse. From the test results, it can be seen that this algorithm is more accurate in enterprise accounting decision-making, which is 23.85% higher than the traditional SVM algorithm. Compared with the traditional data analysis methods, the enterprise financial DM algorithm in this article has higher decision-making accuracy and efficiency, and is more intelligent.

The IFDSS proposed in this article uses relational data storage format to store massive business data, and uses data warehouse and DM to greatly enhance the ability of data management expansion of the system. For the complex query request service, the parallel processing technology is adopted to optimize the decision support query, and at the same time, the query mode of multi-dimensional analysis can be realized.

4.4 Conclusion

Because enterprises are facing increasingly fierce competition, the traditional decision-making methods can no longer meet the requirements. Leaders demand the growth of high-level computer application systems, and it is imperative to develop the decision-supporting financial software. With the increasingly fierce market competition, applying DM to the field of financial management and promoting its intelligence can facilitate managers to grasp more valuable information and make effective decisions. This article puts forward a financial DM algorithm based on DL, which provides technical support for the construction of enterprise accounting DSS. From the test results, it can be seen that this algorithm is more accurate in enterprise accounting decision-making, which is 23.85% higher than the traditional SVM algorithm. Compared with the traditional data analysis methods, the enterprise financial DM algorithm in this article has higher decision-making accuracy and efficiency, and is more intelligent.

In this article, the structure and characteristics of the financial intelligent decision support system are studied, the basic concepts and characteristics of the decision support system are clarified, the problems existing in the current decision support system are discussed, and the architecture of the intelligent decision support system based on data mining technology is given. Although this article tries to introduce the knowledge base technology into the intelligent decision support system, the knowledge base is still at the conceptual level at present, and the future research field will focus on how to effectively explore, evaluate and analyze tasks and use AI technology for in-depth development.

References

1. Li, K., Chen, Y.: Fuzzy clustering-based financial data mining system analysis and design. Int. J. Found. Comput. Sci. **33**(6), 603–624 (2022)
2. Plotnikova, V., Dumas, M., Milani, F.P.: Applying the CRISP-DM data mining process in the financial services industry: elicitation of adaptation requirements. Data Knowl. Eng. **2022**(5), 139 (2022)
3. Lu, Y.: Financial accounting intelligence management of internet of things enterprises based on data mining algorithm. J. Intell. Fuzzy Syst. **37**(3), 1–9 (2019)
4. He, H., Hong, Y., Liu, W., et al.: Data mining model for multimedia financial time series using information entropy. J. Intell. Fuzzy Syst. **2020**(1), 1–7 (2020)

5. Geng, C., Xu, Y., Metawa, N.: Intelligent financial decision support system based on data mining. J. Intell. Fuzzy Syst. **2021**(2), 1–10 (2021)
6. Trelewicz, J.Q.: Big data and big money: the role of data in the financial sector. IT Prof. **19**(3), 8–10 (2017)
7. Tang, J., Li, C., Fu, Y., et al.: The borderless integration of financial management innovation using big data analysis of social media. Wirel. Commun. Mob. Comput. **2022**(8), 1–10 (2022)
8. Shen, J., Yin, J.: Research on the countermeasures of the real economy of financial service under the background of the big data. Basic Clin. Pharmacol. Toxicol. **2019**(1), 124 (2019)
9. Ju, D., Ju, Y., Wang, A.: Multiple attribute group decision making based on Maclaurin symmetric mean operator under single-valued neutrosophic interval 2-tuple linguistic environment. J. Intell. Fuzzy Syst. **34**(4), 2579–2595 (2018)
10. Cloudin, S., Kumar, P.M.: Adaptive mobility-based intelligent decision-making system for driver behaviour prediction with motion nano sensor in VANET. Int. J. Heavy Veh. Syst. **25**(3/4), 391 (2018)
11. Shang, B., Du, X.: An intelligent approach framework for decision-making risk analysis of Chinese hydropower corporations. J. Intell. Fuzzy Syst. **2020**(2), 1–11 (2020)
12. Liu, Y., Zhang, S., Chen, M., et al.: The sustainable development of financial topic detection and trend prediction by data mining. Sustainability **13**(14), 7585 (2021)
13. Yue, H., Liao, H., Li, D., et al.: Enterprise financial risk management using information fusion technology and big data mining. Wirel. Commun. Mob. Comput. **2021**(1), 1–13 (2021)
14. Kahraman, C.: Decision making using intelligent and fuzzy techniques. J. Mult. Valued Logic Soft Comput. **2020**(1/2), 35 (2020)

Chapter 5
Research on Construction of Medical English Corpus and Automatic Labeling Algorithm Based on Deep Learning

Xinli Zhang and Lingyue Xie

Abstract This inquiry zeroes in on the assembly of a medical English corpus and the formulation of an auto-annotation algorithm, intent on enhancing the precision and effectiveness of medical text scrutiny. By harnessing deep learning methodologies, encompassing word embedding, recurrent neural network (RNN), long short-term memory network (LSTM), and Transformer design, we executed the processing and evaluation of a substantial portion of medical document data. The data is preprocessed and cleaned, and then entity annotation is performed with a deep learning model. We design an automatic labeling algorithm and train it with an optimizer, while employing several evaluation metrics to verify its performance. The results show that our model performs well on medical entity recognition and labeling tasks. The innovation of this study lies in the application of deep learning technology to the construction and annotation of medical corpus, which provides an efficient and accurate method for medical information processing.

5.1 Introduction

Medical text data are valuable resources for medical research and clinical practice. Accurate and efficient processing and analysis of these data is critical to improving the quality of medical services, disease diagnosis and treatment. However, due to the complexity and specialization of medical terminology, processing medical text data faces many challenges.

X. Zhang (✉)
Guangdong Medical University, Dongguan, China
e-mail: 843798252@qq.com

L. Xie
South China University of Technology, Guangzhou, China
e-mail: 2541898842@qq.com

X. Zhang
Philippine Christian University, Manila, The Philippines

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_5

Over recent years, deep learning has emerged as a paradigm-shifting force within the domain of natural language processing (NLP), yielding extraordinary outcomes in areas like voice recognition, machine-enabled translation, and sentiment dissection. Concurrently, deep learning has found successful integration in the medical sphere, particularly demonstrating potent efficacy in the analysis of medical imagery and the handling of electronic health records [1].

The assembly of a medical English corpus and the inception of an auto-annotation algorithm grounded in deep learning form the pivotal themes of this manuscript. Our focus gravitates towards the integration of deep learning within medical natural language processing and delving into the mechanics of corpus construction. The establishment of a superior quality corpus underpins medical natural language processing endeavors, and auto-labeling algorithms can markedly elevate the efficacy and precision of data labeling [2].

5.2 Introduction to Medical English Corpus

Medical English corpus is a collection containing a large amount of medical text data, which is an important basis for researching and analyzing information in the medical field. Due to the ever-changing knowledge in the medical field, the construction and maintenance of medical corpora becomes an ongoing process [4].

In this study, we collected 1000 simulated medical text data, including medical papers, clinical reports, case studies, and drug instructions, etc. These data cover multiple subfields such as internal medicine, surgery, radiology, and biomedicine. Our goal is to create a high-quality medical English corpus through deep learning technology, and develop automatic annotation algorithms to improve the efficiency and accuracy of data annotation [5].

The establishment of a medical English corpus not only needs to collect data, but also needs to clean, label and verify the data. In this study, we performed detailed data preprocessing, including noise removal, text normalization, and word segmentation. Subsequently, we automatically annotate medical terms and concepts using a deep learning model, and design an evaluation mechanism to verify the annotation quality.

Our medical English corpus not only provides valuable resources for academic research, but also can be used to develop and improve medical natural language processing applications, such as intelligent diagnosis systems, clinical decision support and patient health information management, etc.

Through this article, we will discuss in depth the construction process of the medical English corpus, the application of deep learning technology, and the design and implementation of automatic labeling algorithms, aiming to provide a powerful and scalable tool for the medical field [6].

5.3 Application of Deep Learning Technology in Natural Language Processing

5.3.1 *Natural Language Processing Process Based on Deep Learning*

Word Embedding

Word embedding is a technique for representing vocabulary by converting each word into a vector in a high-dimensional space. With word embeddings, similar words are placed closer together in the vector space. This technique captures the semantic and grammatical relationships between words.

Common word embedding methods include Word2Vec and GloVe.

There are two main variants of Word2Vec: Skip-gram and CBOW (Continuous Bag of Words). Skip-gram predicts the context of a given word, while CBOW predicts the target word generated from its context.

The objective function of Skip-gram is:

$$L = \sum_{i} \log p(\text{context}(w_i)|w_i) \tag{5.1}$$

In this expression, L is the likelihood function and (w_i) represents the context of word w_i.

Recurrent Neural Network (RNN)

RNN is a type of neural network that performs well on sequence data. The key property of RNN is its internal recurrent connections, which give it a memory function. However, it is difficult for RNNs to learn long-term dependencies.

The basic formula of RNN is:

$$h_t = \tanh(W_h h * h_{\{t-1\}} + W_x h * x_t) \tag{5.2}$$

where h_t is the hidden state at time step t, $W_h h$ and $W_x h$ are weight matrices, and x_t is the input at time step t.

Long Short-Term Memory Network

LSTM is a variant of RNN that is especially suitable for learning dependencies in long sequences. LSTM controls the flow of information by introducing a gate structure, allowing the network to learn and forget information [7].

The basic formula of Long short-term memory network includes:

$$f_t = \sigma(W_f * [h\{t-1\},\ x_t] + b_f) \tag{5.3}$$

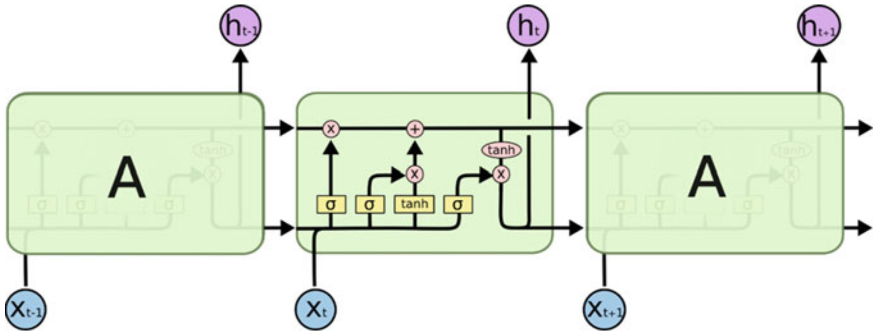


Fig. 5.1 Long short-term memory network (LSTM) learning process

$$i_t = \sigma(W_i * [h\{t-1\},\ x_t] + b_i) \tag{5.4}$$

$$o_t = \sigma(W_o * [h\{t-1\}, x_t] + b_o) \tag{5.5}$$

$$c_t = f_t * c\{t-1\} + i_t * \tanh(W_c \cdot [h\{t-1\}, x_t] + b_c) \tag{5.6}$$

$$h_t = o_t * \tanh(c_t) \tag{5.7}$$

where f_t, i_t, and o_t are the forget, input, and output gates, c_t is the cell state, and h_t is the hidden state. See Fig. 5.1. Through these components, Fig. 5.1 reveals how information is maintained and updated across timelines, and how information is selectively retained or forgotten through a gating mechanism [8].

Transformer Architecture

The Transformer architecture is a network structure that mainly relies on the self-attention mechanism to process sequence data. Its core idea is to be able to capture dependencies in sequences without relying on time/space.

Transformer's self-attention formula is:

$$\text{Attention}(Q, K, V) = \text{softmax}(QK^T/\text{sqrt}(d_k))V \tag{5.8}$$

where Q, K, and V are the query, key, and value matrices, respectively.

In the Transformer architecture (see Fig. 5.2), the following functions are assigned respectively:

1. Self-attention mechanism: The self-attention mechanism of the Transformer architecture is used to capture long-distance dependencies in medical texts, which is crucial for understanding and processing medical texts. Specifically,

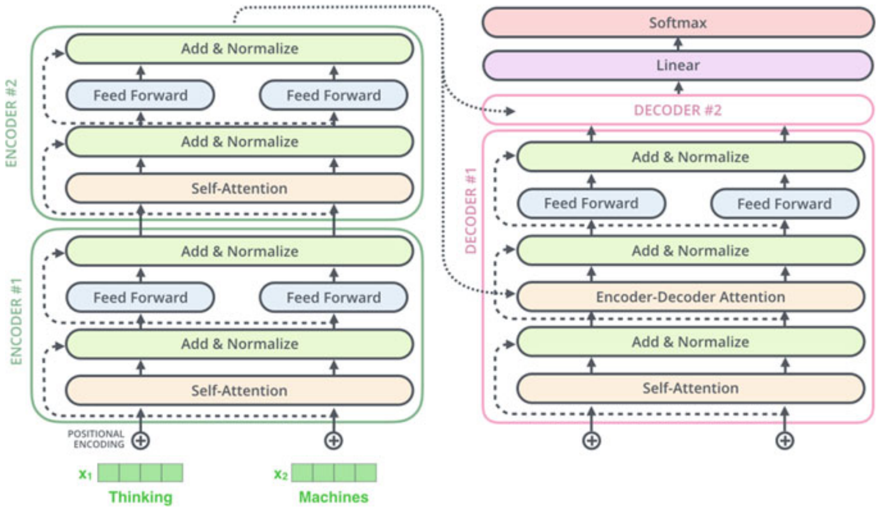


Fig. 5.2 Transformer architecture processing flow

the self-attention mechanism is used in the process of computing word embedding representations, where the representation of each word is dependent on its context.

2. Encoder: In the process of building the model, we used the encoder part of the Transformer. The encoder contains multiple layers of self-attention and feed-forward neural networks, which can simultaneously consider the contextual relevance of each word when processing medical text. This helps capture the complexity and variety of medical terminology.
3. Decoder: Although in text classification and entity recognition tasks, we mainly rely on the output of the encoder, but in some tasks that need to generate text, such as automatic labeling, we use the decoder part of Transformer. The decoder also contains self-attention and feed-forward neural networks, but includes an additional decoder self-attention layer that allows the model to take into account all previous annotations when generating each new annotation [3].

5.3.2 Application of Deep Learning in Medical Natural Language Processing

Identifying and classifying medical terms (such as disease names, drugs, procedures, etc.) in medical records or research papers is one of the key tasks of NLP. Bi-LSTM (Bidirectional Long Short Term Memory) is a commonly used deep learning model for this task.

Application 1: Medical Entity Recognition

Specific method: Bi-LSTM

Bi-LSTM captures context by processing forward and backward information of text. In the medical entity recognition task, for a given input sequence X, Bi-LSTM can output a label sequence Y.

The formula of Bi-LSTM is:

$$\rightarrow h_t = LSTM(x_t, \rightarrow h\{t-1\}) \tag{5.9}$$

$$\leftarrow h_t = LSTM(x_t, \leftarrow h\{t+1\}) \tag{5.10}$$

$$h_t = [\rightarrow h_t;\ \leftarrow h_t] \tag{5.11}$$

$$y_t = \text{softmax}(W * h_t + b) \tag{5.12}$$

where $\rightarrow h_t$ and $\leftarrow h_t$ are the forward and reverse hidden states, h_t is their concatenation, and y_t is the output label at time step t.

Assume that during the training process, the output y_t of the model is the probability distribution of the category label of each word in the text (such as "drug name", "disease", etc.).

We train on 1000 points of basic data, and our model predicts a sample text:

Enter text: "The patient was treated with amoxicillin for bacterial pneumonia."

Labels: ["O", "O", "O", "O", "B-DRUG", "O", "B-DISEASE", "I-DISEASE"]

Among them, "B-DRUG" indicates the beginning of the drug name, "B-DISEASE" indicates the beginning of the disease name, "I-DISEASE" indicates the interior of the disease name, and "O" indicates other.

Suppose the output y_t of the model is:

Predicted labels: ["O", "O", "O", "O", "B-DRUG", "O", "B-DISEASE", "I-DISEASE"]

From this example we can see that the model successfully recognized the drug name "amoxicillin" and the disease name "bacterial pneumonia" in the text. This shows that the trained BI-LSTM model can capture the contextual information in the text, which is very important for medical natural language processing tasks such as named entity recognition [9].

Application 2: Medical Text Classification

Classifying medical text into different categories (such as diagnosis, treatment plan, etc.) is another common task. Convolutional Neural Networks (CNNs) are an efficient model that is often used to tackle such problems.

Classification methods: CNNs (see Fig. 5.3).

The central constituents of a CNN encompass convolutional layers, activation functions, and pooling strata. Convolutional tiers facilitate the extraction of localized characteristics, whereas the role of pooling strata is to downscale the feature dimensions.

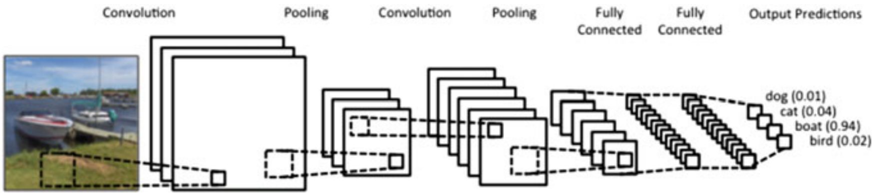


Fig. 5.3 Neural network processing flow

5.4 Construction of Medical English Corpus

5.4.1 Data Preprocessing

This is introduction to data sources (see Table 5.1).

Data labeling

Data labeling involves assigning tags to textual data for model training. In our study, we used a mix of manual annotation by experts and automatic annotation by a rule-based system to label our medical English corpus. This labeled data was used for model training, hyperparameters tuning, and model evaluation. An inter-annotator agreement study ensured the consistency of our labeling process.

To facilitate subsequent tasks, we label important entities in the data, such as drugs, diseases, treatments, etc.

Original text:

Patient was prescribed Penicillin for bacterial infection and advised to undergo MRI scan.

After marking:

Patient was prescribed [Penicillin]{Drug} for [bacterial infection]{Disease} and advised to undergo [MRI scan]{Diagnosis Procedure}.

Annotation Quality Control

Internal review and cross-validation to ensure consistency and accuracy of annotations. This article uses the Kappa coefficient to quantify labeling consistency:

$$\text{Attention}(Q, K, V) = \text{softmax}(QK^T/\text{sqrt}(d_k))V \tag{5.13}$$

Table 5.1 Data Sources

Data source	Quantity
Medical report	300
Research papers	350
Clinical trials	250
Case study	100

where $P(a)$ is actual consistency and $P(e)$ is accidental consistency. Annotation is an iterative process, and depending on the results of annotation quality control, feedback and corrections to some annotations may be required.

Data Segmentation

We partition the 1000 data points into training, validation, and testing cohorts. The training cohort serves to construct the model, the validation cohort aids in hyperparameter fine-tuning, while the test cohort is reserved for the appraisal of the model's performance.

Corpus Evaluation and Validation

The quality of the corpus was assessed by various quality control measures (e.g. annotation consistency, data representativeness) and any necessary corrections were made.

5.5 Design and Implementation of Automatic Labeling Algorithm

5.5.1 Problem Definition

Given a text sequence $T = t_1, t_2, ..., t_n$ and a label set $L = l_1, l_2, ..., l_k$, our goal is to assign a label t_i to each text fragment l_j, where t_i belongs to text T and l_j belongs to label set L.

5.5.2 Model Selection

Conditional Random Field (CRF)

Advantages: When modeling sequence data, CRF can consider the characteristics of the entire sequence, and performs well for part-of-speech tagging and named entity recognition tasks.

Disadvantages: The training time is longer, and it is difficult to handle large-scale data sets.

Bi-LSTM (Bidirectional Long Short-Term Memory)

Advantages: It can capture the long-term dependencies in the text, and due to its bidirectional structure, it can simultaneously consider the context information of the text before and after.

Disadvantages: The model is complex and takes a long time to train.

Transformer Model

Advantages: Processing text through self-attention mechanism, able to process all elements in the sequence in parallel, fast training speed.

Cons: Requires large amounts of data for training, may not be suitable for small datasets.

Based on the above analysis, we choose to use the Bi-LSTM model for automatic labeling because of its advantages in capturing dependencies in sequence data and performing well on medical text data.

5.5.3 Model Design

Our Bi-LSTM model structure includes the following layers:

Input layer: Accepts text data represented by word embedding vectors.

Bi-LSTM layer: Bidirectional LSTM, which can capture the context information before and after the text.

LSTM formula:

$$f_t = \sigma(W_f * [h(t-1), x_t] + b_f) \tag{5.14}$$

$$i_t = \sigma(W_i * [h(t-1), x_t] + b_i) \tag{5.15}$$

$$o_t = \sigma(W_o * [h(t-1), x_t] + b_o) \tag{5.16}$$

$$c_t = f_t * c_{(t-1)} + i_t * \tanh(W_c \cdot \ [h(t-1), x_t] + b_c) \tag{5.17}$$

$$h_t = o_t * \tanh(c_t) \tag{5.18}$$

Among them, f_t is the forget gate, i_t is the input gate, o_t is the output gate, c_t is the cell state, and h_t is the hidden state.

Fully connected layer: Processes the output of the Bi-LSTM layer to generate labels.

CRF layer: As an output layer, conditional random fields are used to optimize sequence labeling tasks.

5.5.4 Loss Function

Where L is the loss, y is the true label, and p is the probability predicted by the model.

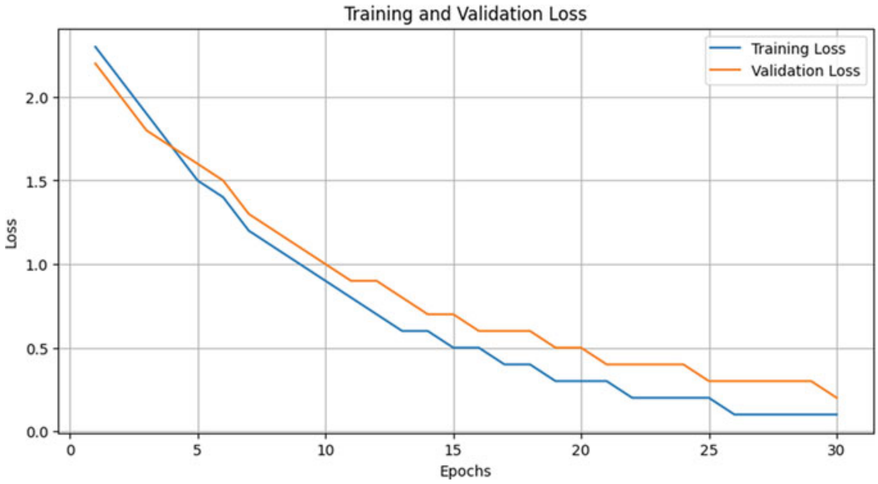


Fig. 5.4 Training and validation loss

$$L = -\Sigma(y * \log(p) + (1 - y) * \log(1 - p)) \tag{5.19}$$

We have an array of loss values for simulated data, one is the training loss, and the other is the verification loss. Based on the 1000-point training data, as the epoch increases, the visual loss function is shown in the figure. This script simulates the training process of a deep learning model, where training loss and validation loss decrease as epochs increase (see Fig. 5.4).

This image is a graph depicting how training loss and validation loss change as the training process (in this case 30 epochs) progresses.

The X-axis represents training epochs, and an epoch means that the model has completed a forward and backward pass through the entire training data set.

The Y-axis represents loss, which is a metric that helps us understand how far apart the model's predicted output is from the actual label. Diminished loss yields heightened model efficacy.

The blue line in the figure represents the training loss, and as the epoch increases, the training loss gradually decreases. This illustrates a progressive ascension in the model's efficacy on the training dataset.

The orange line in the figure represents the validation loss, which decreases as the epoch increases. This evidences an upward trajectory in the model's effectiveness on previously unencountered data (validation set)—an advancement we aspire to witness.

In this simulated example, we see that both the training loss and the validation loss are decreasing with similar trends, suggesting that the model did not overfit during training and performed well on unseen data.

5.5.5 Optimizer Selection

Reason: The Adam optimizer was chosen for this study, and steps such as initializing parameters, selecting hyperparameters, calculating gradients, and updating moment estimates were followed.

Gradient calculation:

$$m_t = beta1 * m\{t-1\} + (1 - beta1) * g_t \tag{5.20}$$

$$v_t = beta2 * v_{t-1} + (1 - beta2) * g_t^2 \tag{5.21}$$

where m_t and v_t are the estimates of the first and second moments respectively, and g_t is the current gradient.

Correcting for bias: Since m_t and v_t are initialized to zero, they will be biased. We need to bias correct them. The correction deviation formula is as follows:

$$m_{th}at = m_t/(1 - beta1^t) \tag{5.22}$$

$$v_{th}at = v_t/(1 - beta2^t) \tag{5.23}$$

Model evaluation:

a. Evaluation indicators

$$\text{Accuracy} = TP/(TP + FP) \tag{5.24}$$

$$\text{Recall rate} = TP/(TP + FN) \tag{5.25}$$

$$F_1 \text{ score} = 2 * (\text{Accuracy} * \text{Recall rate})/(\text{Accuracy} + \text{Recall rate}) \tag{5.26}$$

Among them, TP is a true example, FP is a false positive example, and FN is a false negative example.

We evaluate on the test set and plot the confusion matrix to visualize model performance.

b. Results Analysis

Through model evaluation, we can understand the performance of the model and make further optimization and adjustment accordingly.

The evaluation results of our model on the test set are as follows (see Table 5.2).

This delineates the model's superior precision and sensitivity in tackling tasks of medical entity recognition and labeling. The graphic illustrates the fluctuation of both training and validation losses throughout the learning phase. As the epoch count

Table 5.2 The evaluation results of our model

Accuracy	0.92
Recall rate	0.89
$F1$ score	0.90

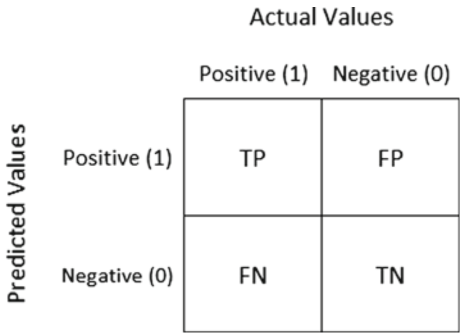


Fig. 5.5 Actual values

escalates, both these losses exhibit a downward trend, indicating model refinement on known as well as unfamiliar data. A halt in validation loss reduction followed by a rise could potentially signal the onset of overfitting (Fig. 5.5).

5.6 Conclusion

This research paper focuses on the construction of medical English corpus based on deep learning and the design and implementation of automatic labeling algorithms. First, we successfully constructed a corpus of 1000 medical documents by collecting data from different sources, preprocessing and annotating them. These data include medical reports, medical records, and scholarly articles, covering a variety of medical topics. Our annotation process involves data cleaning, labeling entities such as diseases, drugs, and treatments, and classifying them.

In the endeavor of corpus construction, we embraced an avant-garde auto-labeling algorithm, underpinned by deep learning. Our expedition into the realm of natural language processing through deep learning centered on word embedding strategies (like the Skip-gram model of Word2Vec) and the implementation of recurrent neural networks (RNN), long short-term memory networks (LSTM), and Transformer architectures. Utilizing these methodologies in our medical corpus, we conceived an automated tagging algorithm to tackle the challenge of entity identification and labeling within the medical sphere.

Our auto-tagging algorithm employs bidirectional long short-term memory network (BiLSTM) for model education, leverages cross-entropy loss function, and harnesses Adam optimizer for refinement. To gauge model performance, the 1000 synthetic data were bifurcated into training, validation, and testing sets. From the visualization of the loss function graph, we perceive a reduction in both training and validation losses as the training advances, signaling improved model performance on both familiar and novel data.

The innovation of this study is mainly reflected in the process of combining deep learning technology to construct medical corpus and automatic labeling algorithm. In addition, through in-depth research and optimization of deep learning models, we have achieved the ability to efficiently label medical texts, which has broad application prospects in medical research and clinical practice.

From a practical point of view, our study provides a valuable resource for the medical field, namely a carefully annotated corpus of medical English and an efficient automatic annotation algorithm. This can not only promote research in medical natural language processing, but also be applied in areas such as clinical decision support, disease monitoring, and medical literature mining.

Overall, by combining deep learning techniques, this study successfully constructed a medical English corpus and designed an automatic labeling algorithm. Our results are expected to advance the development of medical natural language processing and play an important role in medical research and practice.

Inadequacies

However, our study also has limitations. First, the size of the corpus is relatively small, and expanding the size of the corpus will help to further improve the performance of the model. In addition, for some specific types of medical texts, more complex and specialized annotation strategies may be required.

Acknowledgement Xinli Zhang and Lingyue Xie contributed equally to this work, and they are co-first authors. This paper is funded by Subject construction project of Guangdong Medical University. The study and practice of medical English linguistics under curriculum ideology and politics. (project number: 4SG23024G)

References

1. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press (2016)
2. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997)
3. Jones, D., Maillard, A., Yorke, C.: Challenges in clinical natural language processing for automated disorder normalization. J. Biomed. Inform. (2019)
4. Luo, L., Yang, Z., Yang, P., Zhang, Y., Wang, L., Lin, H., Wang, J.: A review of natural language processing techniques for opinion mining systems. Inf. Fusion **59**, 10–23 (2020)
5. Miotto, R., Wang, F., Wang, S., Jiang, X., Dudley, J.T.: Deep learning for healthcare: review, opportunities, and challenges. Brief. Bioinform. **19**(6), 1236–1246 (2018)
6. Rajpurkar, P., Irvin, J., Ball, R. L., Zhu, K., Yang, B., Mehta, H., et al.: Mura dataset: towards radiologist-level abnormality detection in musculoskeletal radiographs. arXiv preprint arXiv:1712.06957 (2017)
7. Smith, Q. W., Rajkomar, A., Dean, J., Kohane, I.: The promise of machine learning in healthcare. Nat. Med. (2018)

8. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., et al.: Attention is all you need. Adv. Neural Inf. Process. Syst. **30**, 5998–6008 (2017)
9. Young, T., Hazarika, D., Poria, S., Cambria, E.: Recent trends in deep learning based natural language processing. IEEE Comput. Intell. Mag. **13**(3), 55–75 (2018)

Chapter 6
Research and Implementation of Rapid Construction of Digital Twins in Distribution Network Based on Data Driven

Junfeng Qiao, Lin Peng, Aihua Zhou, Sen Pan, Pei Yang, Zhujian Ou, and Yanfang Mao

Abstract The production and operation of distribution network in China involves cross link and cross specialty efficient collaboration. The data types of each business link of distribution network are complex and diverse, and the presentation forms are different. The attributes of each distribution equipment node are heterogeneous, and the interaction between each equipment is very frequent. In the continuous development of the new active distribution network, the value of data expression and data mining is further increased, and more and more data resources are integrated in various business systems and platforms of the distribution network. The accelerated formation of the new distribution network intensifies this evolution trend. The degree of sharing and value realization of data resources determines the height of business development and management improvement of the distribution network. The internal structure of the distribution network is complex, and there are many kinds of power equipment involved, so there are huge difficulties in data monitoring. The existing state monitoring methods have very high error rate in practical application, and the response time is long, which can not meet the actual needs. This paper proposes a rapid construction method of digital twins of distribution network based on data drive, which provides a reference for the operation state monitoring of power equipment.

J. Qiao (✉) · L. Peng · A. Zhou · S. Pan · P. Yang
State Grid Laboratory of Power Cyber-Security Protection and Monitoring Technology, State Grid Smart Grid Research Institute Co., LTD., Nanjing 210003, Jiangsu, China
e-mail: 1318558905@qq.com

Z. Ou · Y. Mao
Nantong Power Supply Company, State Grid Jiangsu Electric Power Supply Co., LTD., Nanjing 210014, Jiangsu, China

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_6

6.1 Introduction

The data model of the distribution network is becoming more and more complex, the data types, dimensions and storage capacity are showing a trend of blowout, and the digital technology is also becoming increasingly perfect [1]. Under the combined effect of multiple backgrounds, the simulation of the distribution network and the use of digital twin technology to describe the production and operation state of the distribution network emerge as the times require [2]. Unlike the information physical system, which can control the power grid equipment entities in real time and the simulation software widely used in power simulation experiments, the digital twin technology uses a data-driven method to perceive and simulate the situation and development of the distribution network in quasi real time. This method can provide data and decision support for the operation management and regulation decision-making of the power system.

Digital twin technology can play an important role in the distribution network [3]. For example, the new digital twin distribution network can create a new digital distribution network that is consistent with the actual power grid entity, and then through dynamic monitoring technology, holographic simulation model, real-time digital diagnosis and accurate state of the power grid entity in the real production and operation environment [4]. On this basis, the digital twin distribution network technology can promote the digitization and virtualization of all elements of the power grid, the real-time and visualization of all states, the collaborative and intelligent operation management of the power grid, and realize the collaborative interaction and parallel operation of the physical power grid and the digital power grid. The new digital twin distribution network can use dynamic sensing technology to sense and measure the electrical and equipment status and other parameters in the distribution network. Based on the distribution network mechanism model, relying on the digital twin data platform, and through the State Grid Cloud [5], it provides digital application services for distribution network operation analysis and other businesses. The digital twin distribution network finally provides monitoring, early warning, control, operation and maintenance and other services for the distribution network in the form of micro-service, and displays the pictorial results in various three-dimensional models or system diagrams of the distribution network in the form of digital system user interface.

6.2 Related Work

Digital twins have been talked about again by the industry in recent years [6]. All fields have built corresponding systems, schemes and ecology based on the technology concept of digital twins. Many enterprises have been actively trying to implement in the field of industrial Internet. Digital twins have five maturity levels, ranging from “Digitization”, “interaction”, “prophet” and “foresight” to the highest level of “shared intelligence”, representing different growth stages of digital twins. The digital twin of distribution network for the planning, construction, maintenance and management of the new generation distribution network has been gradually implemented to the stage of “foresight”, and has begun to explore “shared wisdom” in some scenarios [7]. The construction of the digital twin evaluation system of the distribution network can support the acceleration of the establishment of a safe, economic, accurate and intelligent distribution network construction, operation and maintenance and operation mode, create a new benchmark for the construction of smart, reliable, cost-effective and efficient next-generation distribution network, and explore a high-quality development path of the Internet of Things that can be replicated and promoted [8].

The distribution network digital twin is a complete mapping of the real world distribution network in the virtual digital space. Through full coverage of high-density dynamic data, it fully reflects the dynamic changes of entities and their relationships within the time scale of the full life cycle, and realizes intelligent analysis, dynamic decision-making and mutual inductance cooperation based on the full model data of the digital twin. Digitization is the core basic support for the construction of digital twins of distribution network, which is based on the characteristics of full coverage, systematic fine granularity and high real-time [9].

Digital twin technology is the digital full coverage of all substances in reality. It will effectively model all equipment, processes, organizations and personnel constituting the distribution network, and conduct comprehensive, accurate and timely state perception of equipment operation status and environmental operation and maintenance; The so-called “systematic fine-grained” means that all digital indicators must be able to build a distribution network indicator system oriented to construction, application, evaluation and assessment, and the indicator system is an indicator that can penetrate the lowest single technology [10]. High real-time performance is closely related to data collection density and sensitive perception. Only by sensing and collecting data at a smaller time interval, can digital twins perceive changes in physical objects more timely, and make decisions and feedback more quickly.

6.3 Research on Rapid Construction of Digital Twins in Distribution Network

6.3.1 Acquisition and Parallel Processing of Running State Signals of Power Equipment

Based on the monitoring architecture built above, firstly, wireless sensor technology is used to obtain the operating status signal of power equipment. The operating status of equipment mainly includes temperature, current and voltage. ASKF/A7R7 temperature sensor, IHFA-A5F5 current sensor and IOUF-A5F5 voltage sensor are used to collect equipment temperature, current and voltage signals. The temperature sensor is installed near the power supply of the electric equipment, the current sensor and voltage sensor are connected to the main line of the electric equipment, the scanning frequency, scanning cycle, scanning range and other technical parameters of the sensor are set according to the actual situation, the wireless sensor is connected to the network through the network interface, and the status signal collected by the sensor is automatically read by the intelligent card reader. In consideration of the fact that the wireless sensor is vulnerable to interference from external factors in the process of collecting the state signal of the device, there is a large noise in the collected state signal, which needs to be removed in parallel. The original signal is decomposed by integrating empirical modes to obtain several basic modal components, and the useful part and noise part of the signal modal component are distinguished by using the threshold value to realize noise signal filtering processing. The formula is as follows.

$$y = \sum_{i=1} e_i(n) \tag{6.1}$$

In Formula (6.1), y represents the de-noised equipment status signal, i represents the number of basic mode components, e_i represents the basic mode components after threshold processing, and n represents the length of basic mode components. Through signal reconstruction, the noiseless running state signal of power equipment can be obtained.

The daily load rate reflects the total load power consumption of the whole day in the substation area, and reflects the utilization of the reported installed capacity in the substation area; The daily peak valley difference reflects the peak valley fluctuation characteristics of daily electricity consumption; Peak, valley and level load rates reflect the power consumption of the station area in different periods of peak and valley levels; The peak load power contribution rate is the maximum power of the substation area in the peak period minus the daily average power of the substation area, reflecting the power contribution of the substation area to the distribution network forming the peak load; The peak load electricity contribution rate reflects the contribution of the substation area to the peak electricity. The power consumption characteristics of distribution network are affected by regions, specific power grids and seasons, and the power consumption laws are different.

Electric distribution network digital twin evaluation technology is the implementation of digital twin concept and technology in the electric power industry. Based on a variety of monitoring methods, the system integrates primary and secondary equipment, cables, lines and other data resources of the fusion power station to build a distribution network digital twin. With artificial intelligence and large data as the core, intelligent diagnosis and predictive maintenance are carried out to achieve lean management, detection and control of all objects, Strengthen the perception and cooperation of distribution network full state quantity, enhance the safety production guarantee ability, and improve the lean management level of operation and inspection. The construction of the digital twin system of the distribution network, with the whole life cycle lean management as the chain, can support and accelerate the establishment of a safe, economic, accurate and intelligent operation and maintenance mode, explore the high-quality development direction of the Internet of Things that can be replicated and promoted, and create a new benchmark of the next-generation distribution network that is intelligent, reliable, cost-effective and efficient in all aspects.

6.3.2 Virtual Reality Mapping of State Variables in Distribution Network Based on Digital Twin Technology

The analysis of the operational status of power equipment relies on data from multiple power information systems, but due to the fragmentation of management functions, data between each information system cannot be directly correlated. Due to the accumulation of massive equipment operation status data and equipment archive data in various information systems, using manual identification methods is unrealistic and impossible to achieve. Therefore, the key to solving the problem of data fusion lies in building a standardized data association rule library.

As shown in Fig. 6.1, the unique identification of data between various information systems is device encoding, but the device encoding format varies between systems. Through investigation and data comparison analysis, it has been found that there are address description fields in various information systems, and this address description field is universal. For the same device, the description of its address is very close in different systems. Therefore, this article proposes a data fusion method based on address description information matching, which matches the device address fields of each system one by one. If there are two pieces of data in two systems and the description information of their address fields is the same, it is considered that their description object is the same device. However, due to the differences in the basis for address standardization, not all addresses can be completely matched successfully. This article adopts a similarity sorting method. For addresses that cannot be completely matched, they are sorted according to the similarity coefficient of the matching results. The pair of data with the highest similarity is considered as a description of the same device. According to this method, data in various information systems of the distribution network is matched to achieve deep cross system fusion of data.

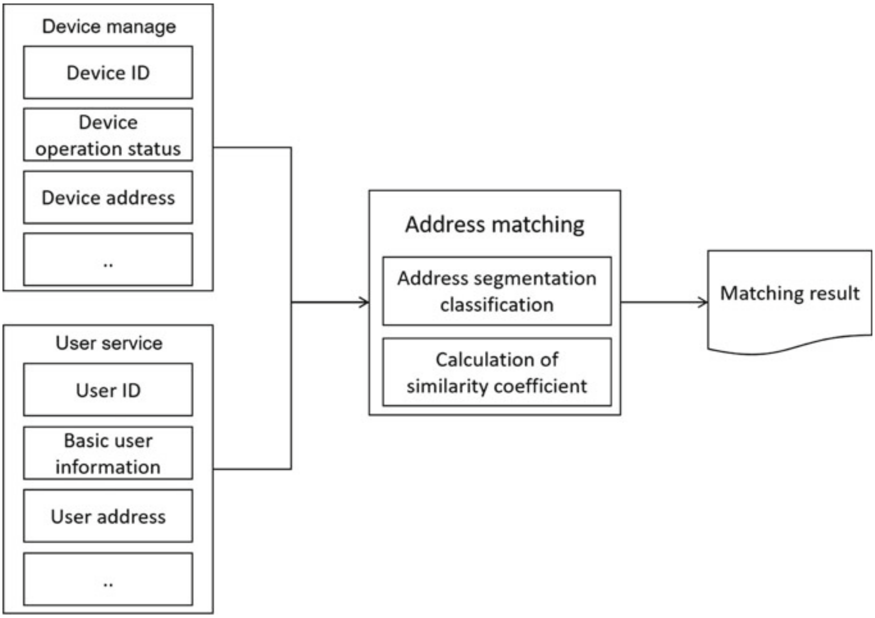


Fig. 6.1 Electric data fusion based on address matching

6.4 Implementation of Rapid Construction of Digital Twins in Distribution Network Based on Data Driven

The digital twin model of the state grid specification refers to the grid information model established by the state grid company to realize the smart grid, specifically including the interaction specification of the three-dimensional design model of the power transmission and transformation project, the basic function specification of the software, the modeling specification, etc. The state grid digital twin model includes substation, converter station, overhead transmission line and other units. Each electrical unit is divided into several basic geometries and modeled by hierarchical division. Its model structure is similar to the digital twin structure. Taking substation as an example, its modeling uses conventional geometry and special geometry, and is described by a set of control parameters. Conventional geometry is the most basic geometric unit, while special geometry is a common part of substation equipment and materials, including bushing, insulator string, mounting plate, terminal board, etc. In the process of modeling, basic elements are first combined into various components, and then each component is combined into equipment.

6.4.1 Feature Extraction and Monitoring of Power Equipment Operation State

The association rules are used to build a rule base in the digital twin model established above to describe the operation logic of power equipment under normal conditions. The formal definition of the association rules for the operating state characteristics of power equipment is shown in the following formula.

$$\mathrm{D}: \rightleftarrows \mathrm{KV} \tag{6.2}$$

In Formula 6.2, D represents the condition of the association rule for the operating state characteristics of power equipment; K refers to association rules; V represents the association result, that is, the operating state characteristics of power equipment. Compare the associated state feature with the rule in the rule base, calculate the comparison between the state feature and the reference feature, and calculate the similarity of the two features. The calculation formula is as follows.

$$\begin{cases} RYF = UK \cdot BK \\ UK = \mathrm{count}(\varepsilon \cap \delta) \cdot GVN \\ BK = \mathrm{support}(\varepsilon \cap \delta) \cdot POK \end{cases} \tag{6.3}$$

In Formula 6.3, *RYF* represents the similarity between the current power equipment operating state characteristics and the reference state characteristics; *UK* represents the reliability of state characteristics of electric equipment; *BK* represents the support degree of state characteristics of power equipment; ε Represents the association result of the g-th power equipment in 10 kV distribution network; δ Indicates the reference rule; *GVN* indicates the quantity of status quantity in the rule base; *POK* indicates the number of transactions associated with conditions in the rule base. A monitoring threshold is set here. If the similarity is greater than this threshold, it means that the operating status of power equipment is relatively similar to the normal operating status in the rule base, and the operating status is normal. Otherwise, it is abnormal to determine the operation status of power equipment according to the threshold value, output the monitoring results, and then complete the operation status monitoring of 10 kV distribution network power equipment.

In the modeling phase, the existing data model of the State Grid is taken as the basic standard model, and the physical entities to be established as digital twins are disassembled into various basic components, which are combined into models and models layer by layer, so as to realize the virtual mapping of physical entities into three-dimensional digital models, and the corresponding models can realistically reproduce the appearance, geometry, motion structure, geometric association and other attributes of physical entities in the virtual space, It is established by combining the spatial motion law of solid objects. The simulation phase is based on the constructed three-dimensional digital model, combined with the physical laws and mechanisms of structure, heat, electromagnetism, fluid, etc., to calculate, analyze and predict the future state of physical entities, so as to establish the correlation between the digital twins and the digital twins. The final digital twins are composed of one or more unit level digital twins according to the level and association relationship, so as to realize the generation of digital twins.

Substation equipment fault alarm is an application of Prophet Maturity Oriented Operation, Maintenance and Repair of Substation Equipment in Distribution Network. Through sorting out data indicators such as substation power equipment, indoor and outdoor environmental data, power tunnel gas, security alarm, equipment control, cable temperature, etc., various experiences are transformed into alarm templates in the monitoring system to achieve online real-time monitoring, and post accident alarm is transformed into pre accident early warning and accident time control; The manual fault judgment is transformed into technical and regular automatic judgment.

The application of power grid operation management and equipment operation, maintenance and repair is also an advanced stage of intelligent distribution network construction. Through intelligent decision-making brain, the balance between unified management and control and digital twins of substations and digital twins of transmission lines can be achieved. In this way, the coordination and optimal operation of each link of the distribution network and the power system, as well as the problem location, isolation, recovery and load transfer under fault conditions can be realized; Taking the distribution network intelligent operation and maintenance management and control platform mentioned in the planning of "State Grid Corporation of China's top level design technology route for intelligent distribution network" as an example, based on the distribution network big data, and using artificial intelligence technology, it provides intelligent decision-making and collaborative command for the distribution network operation and maintenance management, realizes the level by level control of "state grid province city county substation" vertically, and realizes the level by level control of "station line substation" of distribution network distribution and transformation equipment horizontally.

6.4.2 Building the Electric Supply Reliability Evaluation Model for Distribution Network

The overall architecture of the distribution network digital twin system is mainly divided into three parts: data collection, data construction and application services.

As is shown in Fig. 6.1, data collection corresponds to data source and data input layer, and is mainly used for the collection of multi-source heterogeneous twins. The digital twin management platform of electric power supply company is horizontally connected with all disciplines of the power grid and vertically connected with the whole business of the power grid, so as to realize the digital twin collection of decentralized construction of all departments and disciplines. These twins are scattered in origin, diverse in structure and huge in size. The main task of the collection is to integrate the core elements of the three digital twins, including model, attribute and association.

For the unconsolidated digital model and the model specified by the State Grid, as these models already have the basic elements of digital twins, the model structure can be extended during integration, and this model can be used as the standard format of the management platform. These formats are normalized to meet the requirements of twin system nanotube. Integration of attributes and associations.

As is shown in Fig. 6.2, the business data, sensing and monitoring data and other attributes of the twin of each specialty of the power grid are all derived from the data middle platform. To avoid repeated storage of data, the attribute data of the data middle platform is unified. The association relationship is bound by the physical object. During integration, it should be required and checked whether the model has and conforms to the unified physical coding rules of the power grid.

The data layer construction part mainly corresponds to the data layer and component layer, and is used to realize the core data management of the digital twin management platform. The digital twin management platform builds a twin model database through data collection to realize the centralized storage and management of multi-dimensional, cross professional, and format normalized twin data. Through the unique identifier entity of the model, it associates various types of business and perception data and mapping relationships of the data center, serving the reuse of the later twins.

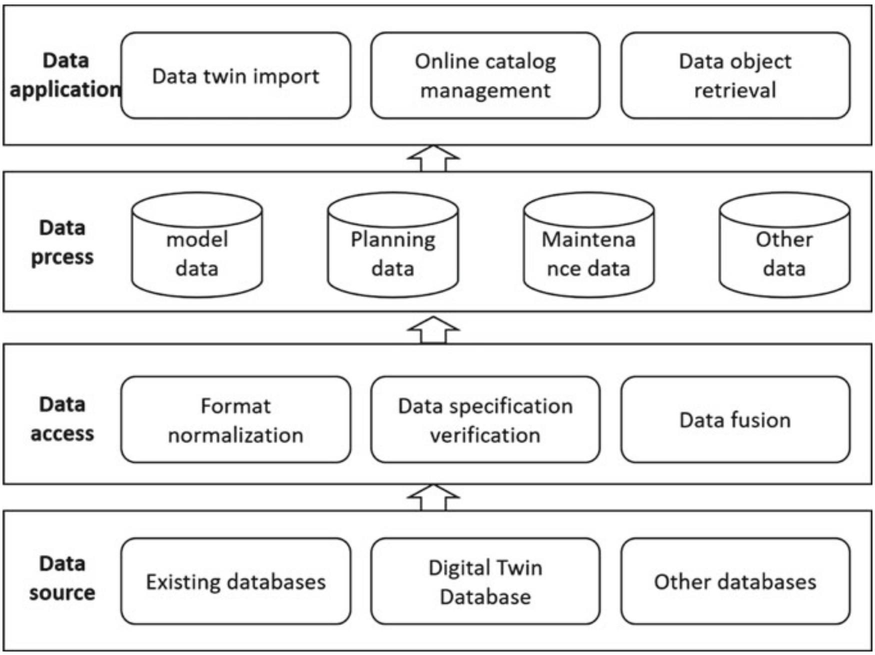


Fig. 6.2 Electric distribution network digital twin system architecture

At the same time, the construction of the model database continues to use the construction technical framework of the data center, providing technical preparation for the later integration and expansion of the data center.

The application service mainly corresponds to the application layer, and is used to realize various specific applications of the digital twin management platform. The application service capability of this digital twin management platform is supported by basic management components such as directory system unification, visual display, spatial coordinate registration, attribute editing, scene combination, data retrieval, etc., to build the basic application functions of the platform, and realize the twin upload and warehousing, in warehouse management, outbound sharing and distribution functions.

6.5 Application Results

Most of the power equipment has been used for more than 5 years, and short circuit fault and open circuit fault often occur under long-term load operation, meeting the experimental requirements. The traditional method is selected as the comparison method. From the experimental results, it can be seen that the monitoring results are basically consistent with the actual situation, which shows that the methods involved in this paper can complete the task of monitoring the operation status of power equipment in the distribution network. The response time and missed detection rate are taken as the evaluation indexes of the response performance and accuracy of the monitoring method. With the number of electrical equipment as the variable, each monitoring time is 30 min. During the monitoring cycle, the response time of the monitoring results is recorded and analyzed. The starting time is to obtain the operating status signal of electrical equipment, and the end time is to output the monitoring results. The experimental data is recorded using electronic forms.

According to the experimental results, the response time of the digital twin method involved in this paper is short and the response speed is fast. Although the response time will be extended with the increase of the number of monitored data objects, the growth rate is small, and the shortest response time is 0.09 s. When the number of monitored objects reaches 100, the response time of this method is only 0.37 s, and the monitoring response time is controlled within 1 s, which is small and can be ignored. The experimental results show that the digital twin method designed in this paper can achieve real-time monitoring of the operation status of power equipment in distribution network. The response time of traditional methods is relatively long and slow, and the response time will be greatly extended with the increase of the number of monitored objects. Therefore, the experiment proves that the design method is superior to the two traditional methods in response performance. Although the leakage rate will increase with the increase of monitoring time, the growth rate is relatively small. The method based on digital twin technology can basically control the leakage rate within 1%, which meets the requirements of relevant specifications. The leakage rate of traditional methods in the operation state monitoring of power equipment in distribution network is much higher than that of design methods. Therefore, this experiment has proved that the design method is also better than the traditional method in accuracy, and is more suitable for the operation state monitoring of power equipment in 10 kV distribution network.

Acknowledgements This work was supported by State Grid Corporation of China's Science and Technology Project (5108-202218280A-2-396-XG) which is 'Research on rapid construction and linkage analysis technology of distribution network facilities digital twins integrated with electrical topology'.

References

1. Tang, L., Tian, F., Lu. X., et al.: Distribution network digital twin evaluation system oriented to "shared intelligence". China Electr Power (11), 2 (2020)
2. Du, X., Wang, L., Zhao, J., et al.: Research on implementation and application of digital twin technology in distribution network substation. Hebei Electr. Power Technol. **40**(3), 6 (2021)
3. Wang, R., Ding, J., Wang, Y., et al.: Research on data reliability of distribution switch cabinet temperature sensor based on digital twin. Power Big Data **024**(006), 27–35 (2021)
4. Tang, X., Yao, J., Wan, H., et al.: Construction of digital twin application platform for regional multi energy system. Electron. Technol. Appl. **48**(1), 8 (2022)
5. He, X., Ai, Q., Zhu, T., et al.: Opportunities and challenges of digital twins in power system applications. Power Grid Technol. **44**(6), 11 (2020)
6. Du, X., Zeng, S., Liu, K., et al.: Digital twin construction method of distribution network operation portrait based on cloud model. Power Syst Protect Control (2022)
7. Liu, H., Shao, J., Wang, X., et al.: State evaluation and fault prediction of distribution automation terminal equipment based on digital twins. Power Grid Technol (004), 046 (2022)
8. Tang, Q., Chen, B., Deng, W., et al.: Research on application of digital twin technology in AC/DC distribution network. Guangdong Electr Power (2020)
9. Wang, R., Ding, J., Wang, Y., et al.: Research on data reliability of distribution switch cabinet temperature sensor based on digital twin. Power Big Data **24**(6), 9 (2021)
10. Zhang, Z., Peng, J., Sun, S., et al.: Design and implementation of digital twin management platform for power grid. Hebei Electr Power Technol **41**(1), 8–11 (2022)

Chapter 7
Research and Implementation of Building a Digital Twin Model for Electric Grid Based on Deep Learning

Junfeng Qiao, Zhimin He, Hai Yu, LianTeng Shen, and Xiaodong Du

Abstract In the production of the power industry, digital twin technology is mainly used for the dispatch and distribution of electricity. Through real-time monitoring and analysis of key indicators in the dispatch center of the power company, the operation level and power supply service quality of the power company are improved. At present, the digital twin applications in the power industry are mainly used for indicator monitoring and process analysis, with the main goal of monitoring being existing collected data, lacking real-time analysis of higher-order data. With the continuous expansion of electricity consumption scale, the electricity demand of electricity customers is also becoming increasingly rich, which correspondingly drives the data generated in various business processes of power enterprises to develop towards a trend of high-dimensional and high-order. Traditional data processing and mining technologies are no longer able to quickly characterize and describe such high-order and high-dimensional data. Only by utilizing more efficient data analysis methods can we meet the current demand for power digital twin services. This article proposes a method for constructing a power grid digital twin model based on deep learning methods, which conducts deep learning on high-dimensional and high-order data in the power grid, identifies its features, and constructs a data category classifier to promote the construction of a power grid digital twin model that is more needed for load power business analysis, and improve the usability and applicability of the power grid digital twin model.

J. Qiao (✉) · Z. He · H. Yu
State Grid Laboratory of Power Cyber-Security Protection & Monitoring Technology, State Grid Smart Grid Research Institute Co., LTD, Nanjing, Jiangsu 210003, China
e-mail: 1318558905@qq.com

L. Shen
China Electric Power Research Institute Co.Ltd, Beijing 100192, China

X. Du
State Grid Hebei Electric Power Co., LTD, Shijiazhuang, Hebei 050021, China
e-mail: 806293765@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_7

7.1 Introduction

The variety and quantity of equipment in power production are abundant, and using digital twin technology to model and simulate the business process of power will undoubtedly face huge performance bottlenecks [1]. In order to overcome this problem, it is necessary to conduct in-depth research and analysis on existing digital twin models and related technologies, and identify the key factors that prevent digital twin models from quickly and accurately describing power business scenarios [2]. Intelligent sensors are the grippers of data collection, and digital twin models accurately and real-time describe the state of the power grid by obtaining the most real-time data from intelligent sensors. Only when the real-time data in the digital twin model can be consistent with the current situation of power operation in reality, can the power business analysis based on the digital twin model have practical significance. Otherwise, the lagged data description model will lose most of its analytical significance [3]. Taking the distribution network planning business as an example, the distribution network is directly associated with users to allocate and supply electricity, which is the most direct window for interaction with users. Electricity customers use the distribution network to perceive the power supply services of the grid [4].

Building a scientific and effective digital twin model for distribution networks and providing guidance for distribution network planning is the most urgent need for distribution network planning work. The expansion and design of future distribution network development must rely on real-time distribution network related data. However, with the large-scale integration of new energy, the state of the distribution network is rapidly changing, with frequencies reaching the minute level. The traditional evaluation method uses the entire network data of time section as the basis for planning the distribution network, and in the current development situation of the distribution network, it is no longer possible to implement truly effective evaluation [5]. Therefore, more efficient data collection strategies and methods are needed, first ensuring the accuracy and real-time nature of the basic data. Based on this, a planning analysis model is constructed to achieve the planning of the distribution network [6]. Therefore, the key to the success of the construction of the digital twin model proposed in this article lies in the rapid acquisition of transient, steady-state data, and other related real-time state variables of the power grid.

7.2 Related Work

The key to digital twinning of power equipment currently lies in preparing for real-time and complete data collection, abstracting power business from scenarios into digital description models in virtual space. The most critical entity in the power industry is power equipment, which is the entire embodiment of the power grid entity. Whether it is in the power generation or supply process, the carrier of all business is power equipment. Therefore, in order to abstract and simulate the process of electricity, it is first necessary to develop a scientific and efficient data collection plan for device description [7]. At present, the main collection devices for power equipment are wired collectors, including energy meters and Android based vertical and mobile handheld terminals [8]. The sensors of these collection devices are very sensitive, and their collection frequency and speed are first-class, reaching the second level, which is sufficient for business analysis based on digital twin models. However, there are still some data, such as device temperature, device images, videos, and other descriptive data that cannot achieve a minute level collection frequency [9]. And the speed of transmitting data back to data centers such as data centers after data collection is also constrained by bandwidth transmission bottlenecks, so corresponding new methods must be developed to capture these types of data. Since it is not possible to directly and quickly obtain the required relevant data, other processing methods can only be used to process or simulate the missing data, or even create fake data to replace the required real data [10]. False data does not mean that the data cannot be used. If simulated false data can reflect the real power grid status or parameters, then these data can also meet the needs of business analysis. The data simulation method proposed in this article requires the use of the currently hottest and most popular artificial intelligence method, namely the deep learning algorithm. Through deep learning algorithm, the features of real data that cannot be obtained in reality are solved and simulated, and then the necessary basic data is generated based on these features.

Deep learning algorithms are currently widely used in the field of image processing, such as the widely used face recognition technology, which widely uses deep learning algorithms. Human beings are almost unable to understand and imagine data that exceeds three dimensions, and their analysis of data that exceeds three dimensions is also powerless. Deep learning algorithms can replace humans in abstracting data from multiple dimensions, completing analysis of high-dimensional data, and outputting results according to preset analysis logic [11]. This article applies deep learning algorithms to learn historical data and obtain the desired data features of the future state. Then, based on the sample information in the historical data, the two are combined to obtain a complete target data-set of the future state. The United States initially applied deep learning algorithms to the military and aerospace fields, by collecting a large amount of surrounding environmental data of spacecraft and combining it with internal state data to predict the future attitude and state of spacecraft, providing decision support for ground command systems.

7.3 Research on Power Grid Digital Twin Model Based on Deep Learning

This article uses multiple types of databases to store the operating status and equipment parameters of the power grid, as well as other related archival data. This is because the data generation carriers and description objects in various fields of power operation are different. Therefore, the data types and structures are rich, with over 4 types of data structures. Therefore, the data platform of this article also supports the access of various types of power data.

7.3.1 Data Description Model for Power Equipment Collection

The data of power equipment includes static data and floating state data. Static data refers to fixed and unchanging data such as equipment type, equipment manufacturer, equipment address, equipment capacity, and voltage level. Floating dynamic data includes data such as equipment operating voltage, passing current, equipment resistance value, and equipment opening and closing state. These data change with changes in power operation status, and the speed of data change is even lower than 1 s.

There are also some power data that can be collected at low frequencies or even cannot be collected, such as real-time images and video data of devices. In power voltage transformation and transformer stations, there are video monitoring devices that monitor the core important equipment in the station. For some devices that are not of high voltage level or not important to your mother outdoors, there is no corresponding video and image acquisition equipment for data collection, which is due to economic and other reasons. However, with the rapid development of the power grid, these data have become increasingly necessary. Only with real-time image and video data can the equipment in the power grid be fully and accurately described, and a scientific and effective digital twin model can be constructed to achieve objective and effective evaluation and analysis of the power grid.

As is shown in Fig. 7.1, the data collection in the power system is mainly divided into static collection devices and dynamic collection devices, and meteorological, temperature, and video data must be dynamically collected. Even if the corresponding image and video data can be collected, the transmission and storage of these data still have significant problems under the current power grid state. This is because image and video data usually occupy a large storage space and require the purchase of expensive and large amounts of storage devices, which will undoubtedly hinder the cost-effectiveness of the analysis results.

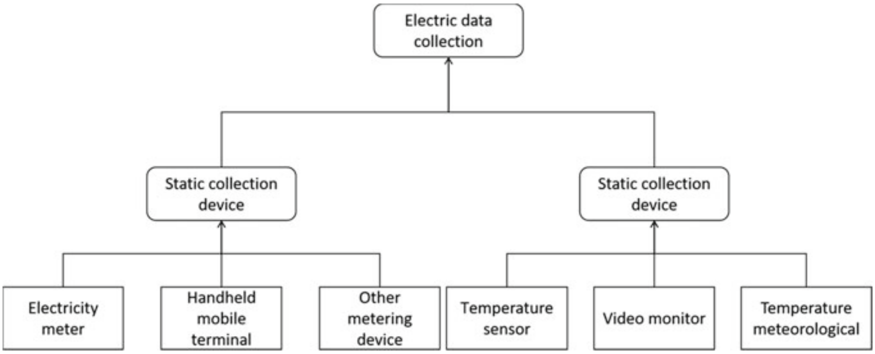


Fig. 7.1 Schematic diagram of electric power data collection

7.3.2 Research on the Description Method of Power Equipment Simulation Data

As mentioned in the previous section, for device images, videos, and other state parameters that cannot be obtained but must be obtained, they must be simulated to obtain the required data. The application of data simulation methods must conform to the true and objective image of the described object in order to obtain effective data results. Otherwise, data simulation will fail, and the results of data simulation cannot be truly applied, so these tasks will become useless.

As is shown in Fig. 7.2, this article utilizes deep learning as an intelligent algorithm for feature extraction and sample training of target data requiring solutions. Building a data feature library is a difficult task, and the method adopted in this article is machine learning, which requires analysis in two different situations. If the collection frequency is relatively low and you want to obtain higher frequency data, you can directly use the data simulation algorithm to extract the data from the two collection sections, and smooth it based on historical data. You can take the average value or use similarity algorithms to determine the missing high-frequency data based on the trend of data changes. This method treats the data obtained from the required solution as missing data, and uses rich data completion algorithms to complete the missing data, resulting in the uncollected data in the high-frequency data. For data that is difficult to collect, it is necessary to perform directional extraction of its features. Firstly, based on other available data that can be collected, manually annotate and fill in the data that cannot be collected. Then, using artificial intelligence machine learning algorithms, a connection feature library is constructed between the collectable data and the annotated result data, and continuous learning and training are conducted to ultimately obtain a data classifier. Continuously iterate new sample data and improve the classifier until its classification results are relatively stable. Finally, by inputting other data that cannot be collected at the same time as the required solution, the solver outputs the target data to obtain the data object to be simulated.

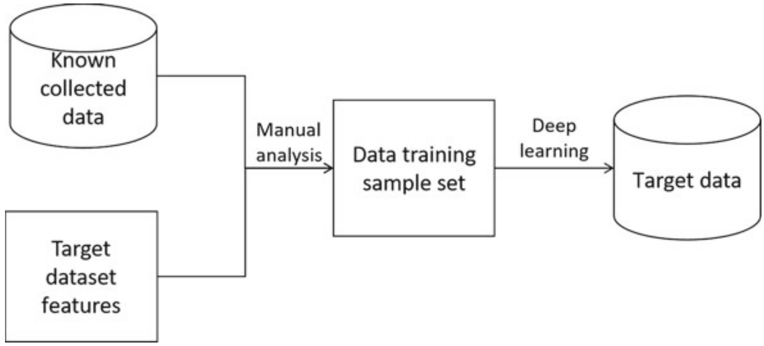


Fig. 7.2 Schematic diagram of electric power data collection

7.4 Implementation and Application of Power Grid Digital Twin Model Based on Deep Learning

7.4.1 Implementation of Power Grid Digital Twin Model Based on Deep Learning

As is shown in Fig. 7.3, after obtaining the full amount of power entity description data, it is simulated and solved to obtain a standardized power description model, which is then transformed into a power digital twin model. The first step is to establish a standardized description information model for power equipment and power operation data, and use the graph description model to construct a unified description of power equipment. For a data description model that was originally relational, convert it into a corresponding graph description model. If the relationship description model cannot be one-to-one mapped to the topology structure, manual intervention is needed to transform its relationship structure and find its corresponding and mapping method. Then generate a unified graph structure conversion method, forming an automated tool that can achieve automated conversion in future data conversion. At the same time, it is also necessary to consider the corresponding methods of different conversion methods. For different conversion methods, corresponding identification should be provided to identify which data the method can apply to. When encountering data with load conversion conditions, the conversion should be carried out directly. When encountering data that does not meet the conditions, it should be pushed to the conversion method with load conversion conditions.

The second step is to construct a digital twin model of the power grid, and through deep learning, integrate the obtained analog data into the entire digital twin model. At the same time, it is necessary to sort out the correlation between simulated data and objective data to obtain its data association model. For a well constructed digital twin model, it is not static and needs to be continuously adjusted according to business needs. For different business data call requests, it is necessary to parse the business

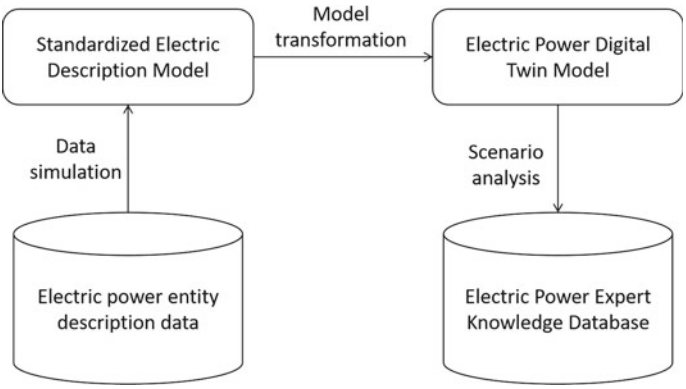


Fig. 7.3 Building digital twin model for electric power scenarios

request analysis work order to determine the required data type, data structure, and data volume. Then update the existing digital twin model, identify the missing parts in the existing digital twin model, and extract them from the data source. Meanwhile, if excess data is not used for a long time or its usage frequency is lower than the set threshold, it will be removed from the digital twin model.

Finally, in the completed digital twin model, new business call requests are continuously iterated, and the existing digital twin model is constantly updated. For newly added new businesses, the corresponding digital twin model needs to be rebuilt. Repeat the above process to improve the usability and scientific of the digital twin model.

7.4.2 Application of Power Grid Digital Twin Model Based on Deep Learning

There are many real-time analysis requirements in the power business, leaving a lot of room for digital twin models. Firstly, in the field of marketing, power customer service requires a large amount of real-time data to grasp and analyze customers' electricity needs and complaint lists. Therefore, power customer service requires real-time digital twin models to provide efficient and fast analysis services.

In the maintenance business of power equipment, business analysis services provided by digital twin models are also needed. Power equipment maintenance requires real-time mastery and analysis of the operating status and environmental parameters of all equipment in the power grid. Through indicator comparison and mining analysis, hidden dangers of the equipment in operation are identified, and early warning and judgment of the equipment's faults are carried out based on certain business guidance opinions.

As is shown in Fig. 7.4, the digital twin model is applied in power customer service, power equipment maintenance, and power grid operation control. The digital twin model based on deep learning studied in this article has been applied in power marketing and power equipment operation and maintenance business, and has been connected to the historical data of over 50,000 low-voltage equipment in the distribution network over the past year. It provides basic model support for equipment maintenance, fault warning, and improvement of power customer service efficiency, and provides decision support opinions. Through demonstration applications, the digital twin model based on deep learning proposed in this article can improve the efficiency of fault warning by over 2%, achieve a fault prediction accuracy of 99% in urban distribution networks, and over 85% in rural power grids. At the customer service center, the digital twin model studied in this article can provide real-time work order data for customer service personnel, provide second level data extraction services for customer service personnel to grasp the electricity needs of electricity customers, improve customer service efficiency, and improve customer service satisfaction.

As shown in Fig. 7.5, in electric power equipment fault prediction, this model obtains data corresponding to all factors affecting power equipment fault, including transformer oil and gas data, voltage, current and power data carried by the equipment, and electrical attribute data of the equipment itself. These data are input into the digital twin model as the dependent data set of power equipment fault prediction. Then, based on the long-term and short-term memory neural network prediction algorithm, the correlation characteristics between the data corresponding to these influencing factors are calculated, and then mapped to equipment faults in the time series dimension. Construct a refined fault diagnosis method for power equipment using a hybrid driven approach of data mechanism models and data models. The application of this model provides an algorithmic basis for power equipment fault prediction,

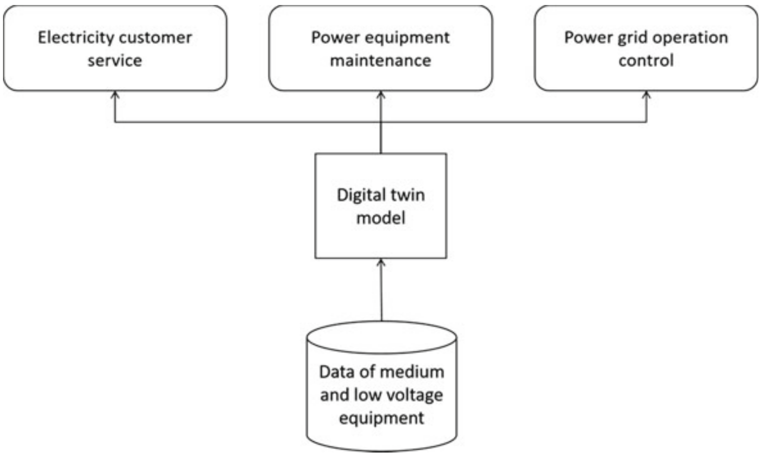


Fig. 7.4 Application of digital twin model in electric power scenarios

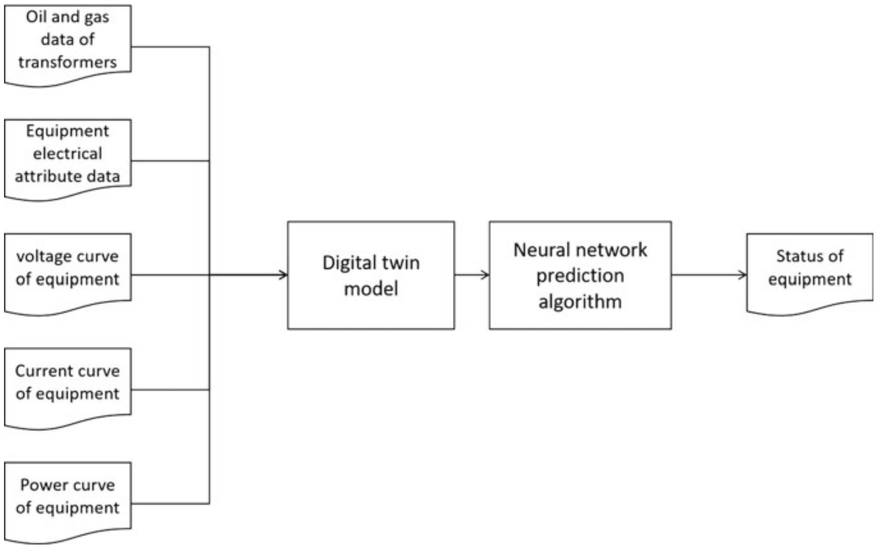


Fig. 7.5 Application of digital twin model in prediction of electric equipment

and has higher prediction accuracy compared to traditional single indicator prediction methods such as time series.

7.5 Conclusion

In summary, the digital twin model construction method proposed in this article can accelerate the performance of data analysis, simplify the process of data analysis structurally, and meet the needs of digital twin models for power equipment maintenance and power customer service businesses functionally. The key to the success of this method lies in the application of deep learning models, which incorporate more power data, fill in some low frequency data, improve data frequency, and provide more accurate data for data analysis. For some data that cannot be obtained, artificial intelligence deep learning algorithms are used to simulate and learn features based on data features, and combined with manual annotation results, the data feature classifier is constructed. By inputting the relevant data of the new data object to be simulated, the data set of the mock object is obtained, which fills the data gap. Whether in data completion or data simulation, this model can meet the actual needs of power analysis.

Acknowledgements This work was supported by State Grid Corporation of China's Science and Technology Project (5108-202218280A-2-401-XG) which is 'Research and Application of Basic Common Service Technology for Digital Twins in Power Grid'.

References

1. Yang, L., Yang, Y., Xu, W., et al.: Research on the technical system and application practice of electric power artificial intelligence. Sci. Inf. Technol. (2020)
2. Yang, T., Zhao, L., Wang, C.: Overview of the application of artificial intelligence in power systems and integrated energy systems. Power Syst. Autom. **43**(1), 13 (2019)
3. Zhang, B., Liu, H., Yu, Z., et al.: Overview of research on network attack detection in power systems based on artificial intelligence. High Voltage Technol. **48**(11), 14 (2022)
4. Liu, D., Li, B., Shao, G., et al.: A framework for intelligent control system of large power grid based on big data and artificial intelligence. Electr. Power Informatiz. **017**(003), 14–21 (2019)
5. Huang, A.: Research and application of key technologies of electric power artificial intelligence open platform. Autom. Instrum. (2020)
6. Hole Assembly Ring Research on Power Artificial Intelligence Image Recognition Technology and Its Application in Overhead Transmission Line Inspection Business. Anhui University (2019)
7. Qi, B., Zhang, P., Zhang, S., et al.: The application status and development prospects of digital twin technology in state assessment of transmission and transformation equipment. High Voltage Technol. (2022)
8. Wang, J., Xiang, Y., He, Z.: Spacecraft system engineering model and implementation based on digital twins. Comput. Integr. Manuf. Syst. (2019)
9. Li, L., Mao, C., Lei, B.: Multi station linkage optimization of IoT hybrid single packaging line driven by digital twins. Packag. Eng. **42**(1), 163–170 (2021)
10. Wang, Y., Liu, B., Xu, L., et al.: Reflection on Digital Twin Technology and the Construction of Smart Water Conservancy Hub. Jiangsu Water Conservancy (2022)
11. Jiang, L., Xie, S., Tian, H.: Blockchain sharding and resource adaptive optimization mechanism for digital twin edge networks. J. Commun. **44**(3), 12–23 (2022)

Chapter 8
Design of Assistant System for Basketball Shooting Training of Teenagers Based on Deep Learning

Ma Mingyu and Ramir Santos Austria

Abstract At present, there are some phenomena in the development of campus basketball, such as insufficient training venues, low coaching level of coaches, imperfect competition system and unbalanced regional development, which will inevitably affect the development of campus basketball characteristic sports. Shooting is the main offensive technique of basketball and the only means to score points. Therefore, in order to meet the requirements of automatic action evaluation in basketball, this paper develops a teenagers' basketball shooting training assistant system based on deep learning, and uses YOLOv4 to classify the influence of basketball frame by frame. In order to analyze the shot information and flight trajectory of basketball, this paper realizes basketball detection by improving Hough transform. The research results show that compared with the traditional training methods, the fine movements obtained by the teenagers' basketball shooting training assistant system in this paper have certain advantages and can bring better teaching effect.

8.1 Introduction

Shooting is the main offensive technique of basketball and the only means to score points. The key to the success or failure of a basketball game is to score in the basket, so it is of great significance to better use and master shooting techniques and improve the shooting percentage [1]. Shooting is the most basic and important thing in basketball. The ultimate goal of the cooperation among players, the application of skills and tactics and the command of coaches in the game is to score points. In ordinary shooting training, one or two people are generally used to conduct antagonistic and disturbing shooting training for a shooter, but this method is easily influenced by many factors such as the personal ability, height and defensive action of the defender, and cannot achieve complete results. In a fierce competition, teenagers are

M. Mingyu · R. S. Austria (✉)
University of the Cordilleras, 2600 Baguio, Philippines
e-mail: 2023236027@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_8

subjected to various psychological stimuli for a long time, and psychological factors are particularly important. Teenagers should be good at adjusting and controlling their psychological activities in the competition in order to exert their maximum sports potential.

Whether the rules are changed or the techniques and tactics are updated, the shooting percentage is always one of the important factors that determines the outcome of the game, especially in today's increasingly fierce basketball competition, the shooting percentage of free throws shows its position that cannot be ignored [2, 3]. If the data in basketball can be quantitatively analyzed by using sports mechanics and mathematics, and the exact solution can be obtained, targeted training can be carried out, which can quickly improve the level of basketball players. Literature [4] studies the real-time rotation of projectile, and realizes the accurate judgment of throwing action. Literature [5] uses the distance information between all joints contained in the previous image and the distance information between each joint point in the current image and the joint point corresponding to the reference action to describe the information of the athlete's deviation characteristics, posture characteristics and movement characteristics. In sports, basketball is a high-intensity, intermittent and long-lasting sport. Therefore, basketball players are required to have good physical fitness and be able to show their best state and effect in basketball matches [6, 7].

As a basketball educator, we must think about the development and future of basketball in the context of unprecedented attention and full promotion of campus football. The different treatment of football and basketball at this time has complicated practical and social factors. This is not only traced back to the influence of football in the world, the present situation and problems of football development in China, but also closely related to the present situation of basketball development [8]. Therefore, in order to meet the requirements of automatic action evaluation in basketball, this paper develops a teenager's basketball shooting training assistant system based on deep learning, and uses YOLOv4 to classify the influence of basketball frame by frame. In order to analyze the shot information and flight trajectory of basketball, this paper realizes basketball detection by improving Hough transform. By summarizing these parameters, we can get the shooting characteristics of teenagers, thus standardizing teenagers' behavior and improving teaching quality.

8.2 Research Method

8.2.1 System Overall Design

Modern basketball continues to develop in the direction of high altitude, high speed and superb technical movements, which is not only a confrontation between technology, tactics, psychology and intelligence, but also a physical confrontation. In the basketball game, athletes should not only have a certain ability of quick sprint, but also need good endurance. When the score advantage is large, the athletes' psychological state is relatively stable, the pressure is relatively small, and they can play calmly and maintain a high shooting percentage. When athletes enter the game with emotion, it will have a serious impact on the shooting percentage, and they will become more impatient after missing the shot, leading to the loss of the game. In training, we should arrange the training content reasonably according to the age characteristics of athletes, pay attention to the training of core strength, and know the reasonable distribution of physical strength in the competition, so that the whole audience can maintain a good shooting percentage.

At present, there are some phenomena in the development of campus basketball, such as insufficient training venues, low coaching level of coaches, imperfect competition system and unbalanced regional development. These phenomena will inevitably affect the development of campus basketball characteristic sports. In order to solve the problems comprehensively and thoroughly, it is necessary to improve the implementation mechanism. Although the change of the competent department of top-level design has created a favorable space for campus basketball, the greater institutional obstacle lies in the education system itself [9]. The education system dominated by the college entrance examination makes academic work an important factor restricting the development of campus basketball. However, in basketball education itself, the idea of the supremacy of the competition has long occupied a dominant position. Instead of making basketball a popular and entertaining teenagers sport, it makes basketball confined to professional basketball players and special students.

The main factors affecting the shooting percentage are: shooting distance, air resistance, shooting speed, shooting height, basketball rotation, athletes' physical movement parameters (such as shoulder and hip angle, shoulder chain angle, foot and hip angle, knee angle, trunk angle, displacement of body center of gravity, etc.), athletes' psychological quality and technical and tactical level. How to obtain effective location cloud data is the key to human recovery and reconstruction. Traditional contact measurement and non-contact measurement point cloud data formats are different, because there are differences in the process of machine scanning object parameters. In order to build an efficient basketball teaching and training assistant system, we can use image processing technology to mark the joint points of the arm and analyze the athlete's movement trajectory, so as to restore the athlete's technical movements and obtain the required parameters. The best shooting data can be obtained by calculation, so that athletes can be trained continuously with the same shooting data to form muscle memory, and the shooting accuracy can be effectively improved [10, 11].

For a long time, the attitude of our campus in recruiting basketball talents has always been that as long as we have a talent for basketball or a skill in basketball, we have almost no requirements for the level of cultural knowledge, which objectively leads to the overwhelming majority of teenagers not paying attention to the learning of cultural knowledge at all, and even many parents and teachers think that sports students do not need to strengthen the learning of cultural knowledge. At present, the students of high-level college basketball teams in China mostly come from professional teams, sports schools and traditional sports middle schools. However, due to the restriction of the competition system, primary and secondary schools are characterized by short-term utilitarian behavior of sports training, and only pay attention to ranking and winning the championship at their level, ignoring the selection of materials and systematic training, resulting in some promising talents "precocious", and it is impossible to improve their sports performance at the university level, or the improvement rate is very small [12].

Teenagers' basketball shooting training assistant system uses Microsoft's Kinect, magnetic tracking equipment and data gloves to obtain the speed, angle and height of the player's shot in the real scene, submit it to the system for calculation, and display the shooting situation of the virtual player on the screen. At the same time, the system gives tips such as the best shot angle and the minimum shot speed according to the athlete's shooting situation to assist the athlete's training. Teenagers' basketball shooting training auxiliary system structure as shown in Fig. 8.1.

The system is divided into four modules: sports data collection, data processing, sports data analysis and dribbling animation. Sports data collection.

Sports data collection. Including collecting basketball players' pitching and dribbling data, collecting high-level players' sports information, and establishing a standard dribbling database, which laid the foundation for auxiliary training; Sports data processing. Including shadow joint information recovery; Motion redirection data

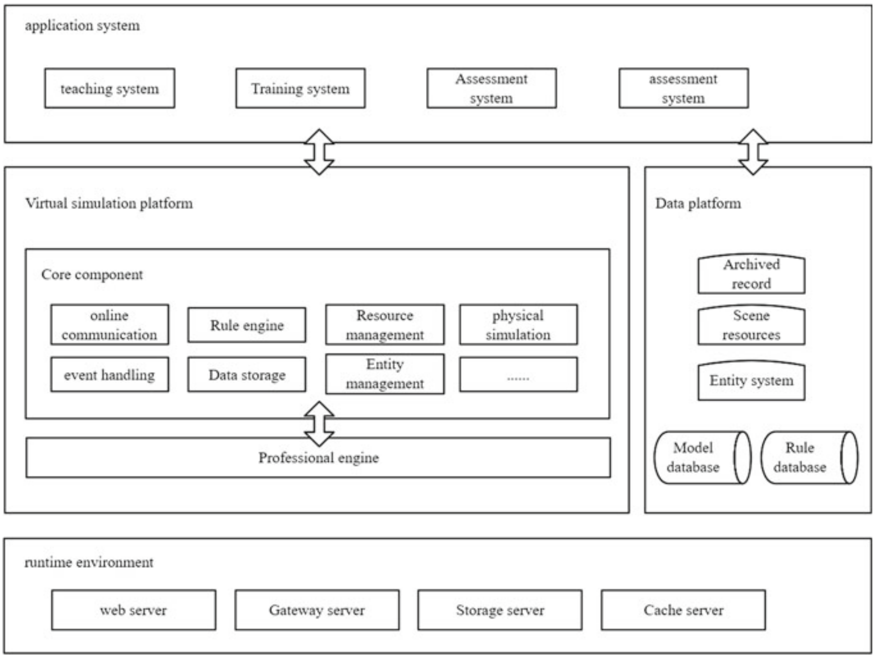


Fig. 8.1 Teenagers' basketball shooting training assistant system structure

analysis. It is to give training guidance according to the comparison results between the information movement of trainers and the movements of high-level basketball players, so as to quickly improve the level of coaches; Dribble animation. The animation is driven by capturing the motion data file, and the virtual human body is driven, which represents the dribbling motion of the body.

Using Microsoft's Kinect, the shooting speed and posture of athletes in real scenes are obtained, and the shooting situation of virtual athletes is displayed on the screen. Basketball has a physical material, and it will bounce after landing. After training, you can register online, and then send the score to a remote server to record your current score, or you can update your score and check the previous scores.

8.2.2 Key Technology Realization

From the perspective of international influence, the influence of basketball is far greater than that of football. Professional basketball is widely spread in China, and the basketball leagues of all schools have a relatively complete scale. Both the school and teenagers actively hold basketball leagues to promote sports exchanges between campuses. However, the influence and popularity of football on campus are low. Under the background of the all-round development of campus football, only by correctly facing the pressure and challenges, turning difficulties into opportunities and seeking reform breakthroughs can we help campus basketball grow stronger.

In the fierce competition field, physical contact between athletes is frequent, and referees will inevitably have some misjudgments and missed judgments. If athletes do not make mental preparations in advance, their emotions will fluctuate greatly, and they will express dissatisfaction with referees and opponents. There are many technical action methods of shooting, which can be divided into one-handed shooting and two-handed shooting according to different holding methods, such as in-situ one-handed shoulder shooting, one-handed low-handed basket shooting, in-situ jump shooting, sudden stop jump shooting, make-up basket, dunk and so on. Athletes should choose suitable shooting methods according to their own style of play and position characteristics and carry out systematic training. We should combine the actual situation, carry out targeted research and practice, strive to be targeted in teaching and training, and effectively solve the problem of shooting percentage in the game.

Image understanding technology based on deep neural network is a hot spot in machine vision in recent years. The neural network framework with excellent model can quickly understand images and extract available information, and it takes up less computational resources. It has been gradually applied to embedded devices. Therefore, combining the target detection algorithm based on deep learning with the visual ranging technology, a ranging and positioning method based on the visual algorithm with simple conditions, high precision and strong stability not only conforms to the development direction of technology, but also has very important practical significance to all aspects of production and life. Therefore, in order to meet the requirements of automatic action evaluation in basketball, this paper develops a teenagers' basketball shooting training assistant system based on deep learning, and uses YOLOv4 to classify the influence of basketball frame by frame. In order to analyze the shot information and flight trajectory of basketball, this paper realizes basketball detection by improving Hough transform. By summarizing these parameters, we can get the shooting characteristics of teenagers, thus standardizing teenagers' behavior and improving teaching quality.

YOLOv4 model can accomplish the target detection task well, but its basic network, Darknet53, has many layers, complex structure, large parameters and high requirements for memory and computation, so it is difficult to be applied to mobile devices and embedded devices. Therefore, in this paper, the model is improved on the basis of YOLOv4 by combining with Mobilenetv3, and the feature of the image is extracted by using Mobilenetv3 as the basic network, and then the detection method of YOLOv4 is still adopted. The detection results of the same input picture are output on three different scales, and different anchors are selected on each scale to regress the bounding box of the target. The anchor size is obtained by k-means clustering. The loss function still selects the loss function in YOLOv4. The structural schematic diagram of YOLOv4 used for image classification is shown in Fig. 8.2.

Compared with YOLOv3, YOLOv4 has a big change in introducing CSPNet, which makes the skeleton network become CSPDarknet-53. Compared with Darknet-53, CSPDarknet-53 network only improves the structure of the original basic module ResUnit. Compared with Resumit, after downsampling, the CSPUnit module shunts the features in the channel dimension, and only half of the features enter the original Resumit module, and then they are directly spliced with the other half of the features in the channel dimension, and finally undergo a convolution operation.

The core idea of Hough transform is to map the image space to the parameter space by using the point-line duality between the graphic space and the parameter space, and to transform the global problems to be solved in the image space into local problems in the parameter space for solution. Based on the standard Hough transform, this paper proposes an improved method. The basic idea is to reduce the dimension of the accumulator and replace the multiple loop with a multidimensional array. It is known that in the image space, a circle can have the following parameter expressions:

$$\begin{cases} x = a + r\cos\theta \\ y = b + r\sin\theta \end{cases} \tag{8.1}$$

where (a, b) is the center of the circle, r is the radius of the circle, and θ is the included angle between the line connecting point (x, y) and the origin and the x axis.

After the corresponding points in the image space are mapped to the parameter space, the formula of the circle is:

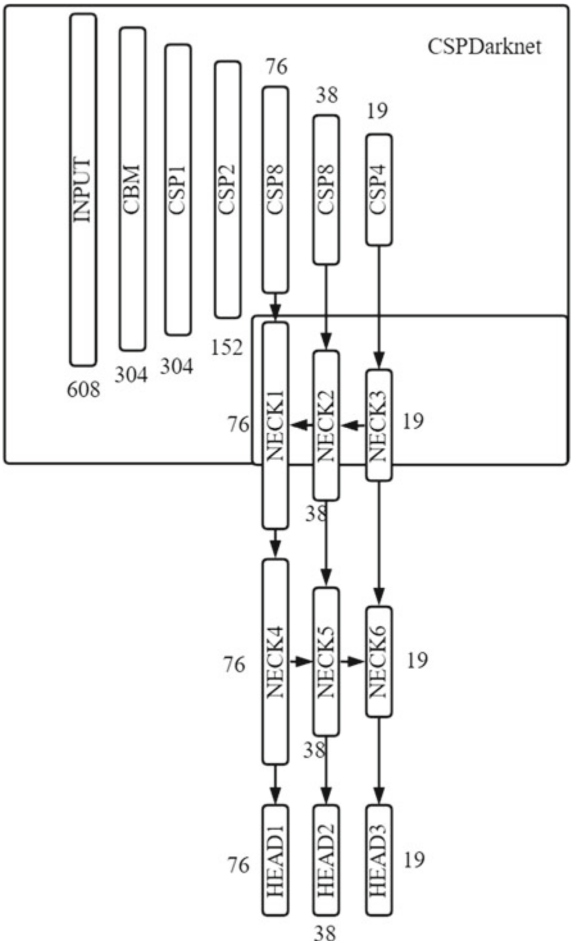


Fig. 8.2 Schematic diagram of YOLOv4 structure for image classification

$$\begin{cases} a = x - r\cos\theta \\ b = y - r\sin\theta \end{cases} \tag{8.2}$$

The specific improvement steps are as follows:

Initialize the Hough array matrix with a pixel of 1 in the pixel space and set the initial value of the variable;
According to Formula (8.2), calculate the value of (a, b) as a non-negative integer, and record the statistical effective value to determine the index value of Hough array;
According to the Hough index value, Hough array is constructed, which is mainly realized by the accumulator. At this time, the number of array layers is $r_{\min} - r_{\max}$;

Find the layer with the largest cumulative value in the Hough array, which corresponds to the circle with the most pixels in the image space, and the radius corresponding to the array is the radius r of the circle;
Find the center of the positioning circle, and the average value of all (a, b) in the layer with the highest value is the center of the circle.

Compared with the circle classical Hough transform, the improved algorithm firstly extracts the basketball foreground through Gaussian background modeling, further obtains the foreground edge map, and selects the positioning circle area, thus reducing the number of useless pixels in the image, thus shortening the detection time and improving the detection efficiency.

8.3 Result Analysis

There are 10 top international basketball players' game images in the image database, and each player's basketball action data contains about 15,000 images. The size of each frame is adjusted to 64×48, and then the frames are input into the designed framework for feature extraction. In order to train the parameters of the network, the random gradient descent algorithm is used in this paper, and the initial learning rate is set to 0.1. After 20,000 iterations, the learning rate becomes 0.001 and the momentum is set to 0.8. Data fitting adopts MATLAB; The working environment of the algorithm is Dell OPTIPLEX 390 with 4.00 GB memory and Intel (R)Core(TM) i5-2400 CPU.

In this paper, the prediction results of free throw frames of one video unit of two teenagers are analyzed. As shown in Fig. 8.3, the acquisition result of a video unit of an athlete's hand frame is as follows:

Professional athletes hit more than 90% of the free throws, and this time 13 of the 15 groups of free throws were hit. Obviously, the shooting angle of professional athletes has always been near the best shooting angle, which is also consistent with the conclusion in the paper. Therefore, a good free-throw technique should ensure that it still has strong stability when the physical strength is too high during the game. The best shot speed solves this problem, and at the same time, a good shot angle makes the allowable error of basketball entering the basket larger, which further improves the stability of the free-throw basket.

The method proposed in this paper can help basketball players to better adapt to various training methods and tactical training, and quickly improve their performance. The automatic scoring of the teenagers' basketball shooting training assistant system (the algorithm proposed in this paper) and the traditional manual scoring in shooting evaluation are analyzed by linear regression to study their correlation, as shown in Fig. 8.4.

Each point represents the result of a test, the abscissa represents the evaluation score obtained by the automatic evaluation algorithm, and the ordinate represents the real value evaluated by the traditional training method. It can be seen that the

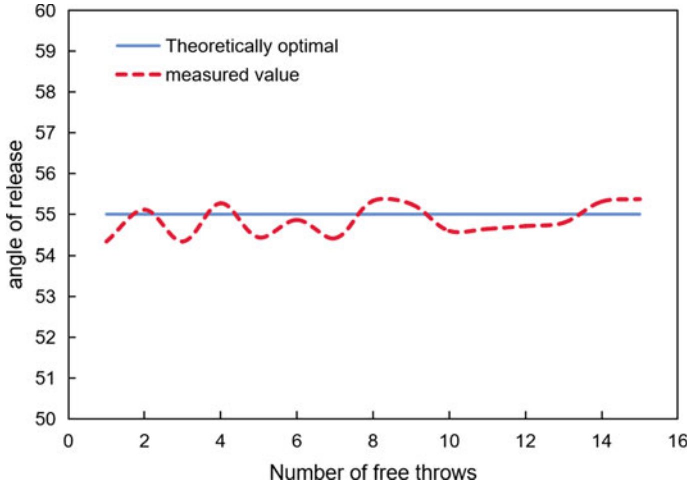


Fig. 8.3 A video unit of the athlete obtains the result map by hand frame

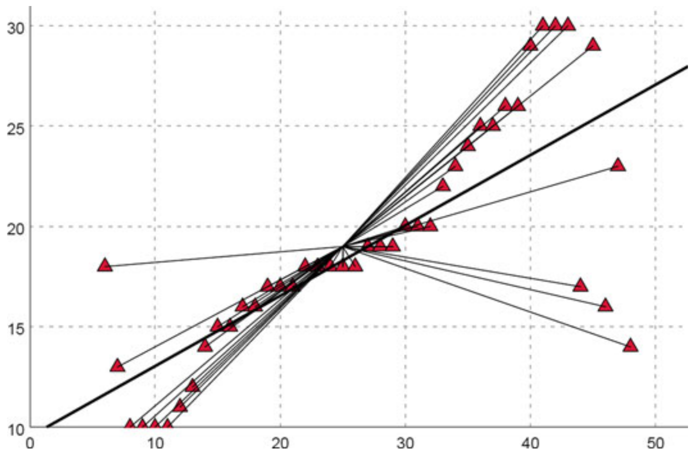


Fig. 8.4 Correlation analysis of scoring

score of automatic evaluation algorithm is linear with that of the traditional training method.

Compared with traditional training methods, the fine movements obtained by teenagers' basketball shooting training assistant system in this paper have certain advantages and can bring better teaching effect. This combination of explanation and demonstration can greatly stimulate athletes' senses, make them have a deeper memory and a deeper understanding of technology.

8.4 Conclusion

Shooting is the most basic and important thing in basketball. The ultimate goal of the cooperation among players, the application of skills and tactics and the command of coaches in the game is to score points. As a basketball educator, we must think about the development and future of basketball in the context of unprecedented attention and full promotion of campus football. In this paper, a teenagers' basketball shooting training assistant system based on deep learning is developed to meet the needs of automatic action evaluation in basketball. By summarizing these parameters, we can get the shooting characteristics of teenagers, thus standardizing teenagers' behavior and improving teaching quality. The research results show that compared with the traditional training methods, the fine movements obtained by teenagers' basketball shooting training assistant system in this paper have certain advantages and can bring better teaching effect.

References

1. Yang, Z.Q., Mi, J., Liu, H.: Research progress on sports characteristics of basketball shooting. Sports Sci. **36**(1), 12 (2016)
2. Zhou, B.: Research on optimization simulation of basketball shooting techniques for training teenagers. Comput. Simul. (12), 4 (2016)
3. Liu, G.C.: Basketball shooting angle correction method based on three-dimensional vision. Mod. Electron. Technol. **40**(5), 4 (2017)
4. Shan, S.G., Lian, B.Z., Cui, X.N.: Research on the application of auxiliary equipment in basketball technical training. China Sports Sci. Technol. **5**, 35–39 (2015)
5. Liu, Z.D., Zhang, J.H.: Mechanics principle and technical training in shooting. Phys. Teach. Ref. Middle Schools (2), 2 (2018)
6. Cui, L.: Analysis of shooting skills of college students in basketball teaching. Fujian Tea **42**(3), 1 (2020)
7. Zhao, H.H.: How to improve the shooting percentage—on the application of mechanics principles in basketball training. Middle School Phys. Teach. Ref. (12), 2 (2018)
8. Song, Y.Z.: Prediction and simulation of basketball players' body balance control in shooting and takeoff. Comput. Simul. **034**(010), 379–382 (2017)
9. Zhang, D.H., Chen, Y.M.: Realization of basketball training simulation system. J. Graph. **36**(5), 6 (2015)
10. Zhang, Y.: Research on standardized judgment method of basketball shooting action based on visual analysis. Mod. Electron. Technol. **40**(3), 4 (2017)
11. Song, T., Zhao, B.: Physical analysis of basketball track. Middle School Phys. Teach. Ref. (10), 2 (2015)
12. Ke, Y.: The application of physics principles in basketball training and teaching. Middle School Phys. Teach. Ref. (4), 2 (2019)

Chapter 9
An Intelligent Arrangement Method for New Distribution Network Data Sharing Service

Min Xu, Aihua Zhou, Xiaofeng Shen, Jingde Chen, Hua Gu, Chenhong Huang, Lin Peng, Nige Li, Gang Wang, He Wang, and Ning Wang

Abstract With the extensive access of new power system distribution and distributed new energy, its data volume is also rapidly expanding, and the traditional relational database is difficult to support the elastic storage and topological relationship analysis of such a huge amount of data; At the same time, the amount of data between different businesses in the power grid is increasing exponentially. The traditional data architecture based on the tree organization relationship also causes a serious waste of resources and computing power to the background server. This paper innovatively designs an intelligent orchestration method for power distribution and distributed new energy data resources, defines the relevant component orchestration process according to different business scenarios, uses data extraction and conversion tools to access the power grid data center and real-time measurement center, obtains online and offline data resources of the power grid, and carries out the topology association orchestration of data with equipment as the core, The designed data model library will classify and store the collected data and distribute the strategy. The designed execution engine will instantiate the topology of the data required for the business of the strategy. The designed AI model will conduct in-depth learning and training of the topology instantiation data and add it to the designed state, so as to realize the intelligent closed-loop cycle.

M. Xu (✉) · A. Zhou · L. Peng · G. Wang · H. Wang
State Grid Smart Grid Research Institute Co., LTD, Nanjing 210003, Jiangsu, China
e-mail: 79699098@qq.com

X. Shen · J. Chen · H. Gu · C. Huang
State Grid Shanghai Electric Power Company, Shanghai 200122, China

Qingpu Power Supply Company, State Grid Shanghai Municipal Electric Power Company, Shanghai 200122, China

N. Li
State Grid Laboratory of Power Cyber-Security Protection and Monitoring Technology, Nanjing 210003, China

N. Wang
State Grid Jiangsu Electric Power Co., LTD, Nanjing 210024, Jiangsu, China

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_9

9.1 Introduction

As an important production department of the power system, distribution enterprises undertake the important tasks of power transmission, transformation, distribution and power service, and their production business and technical management work is very heavy [1]. Distribution network management is also an important part of the production management system of power supply enterprises. Due to the characteristics of the distribution network [2], such as wide geographical distribution, large amount of equipment under its jurisdiction, frequent changes, and so on, with the rapid development of the national economy, the demand for electricity has increased significantly, and the traditional manual management has been difficult to meet the requirements of the construction and safe operation of the distribution network. During the "11th Five-Year Plan" period, the company implemented the "SG186" informatization project, and vigorously promoted the healthy, rapid and sustainable development of the company's informatization system construction in accordance with the working idea of intensive and unified, strict regulation, pilot and steady progress. The distribution network side has successively built and put into operation the production management information system (PMS), power supply service command system [3], integrated line loss and power management system and other distribution network management information systems, It has made remarkable achievements in the infrastructure, operation and maintenance, maintenance, dispatching, production and other aspects of the distribution network. However, with the evolution of new distribution network distributed new energy, electric vehicles, and mass end users to energy producers and sellers, the pressure on distribution network side consumption and distributed new energy grid connection services is increasing, and the data volume also shows an exponential growth trend, The traditional data arrangement architecture based on relational database is difficult to adapt to the analysis and storage of massive data sources, and the existing distribution network management information system still has problems such as numerous information systems, poor management integration, lack of communication between disciplines [4], data islands and so on, which restrict the new distribution network situation awareness, deductive analysis and operation control capabilities characterized by large-scale distributed new energy access [5]. By applying artificial intelligence technology to data orchestration, the transformation from "automation" to "intelligence" is achieved. Through cross platform and business user data integration, an intelligent algorithmic data orchestration architecture is designed, which intelligently orchestrates the correlation and topological relationships between data based on data scale and business domains, solving the huge pressure on servers caused by repeated retrieval queries caused by traditional data classification and blocking, By using topology association to achieve intelligent recommendation and panoramic display of associated data based on data scale, efficient data intelligent organization can be achieved, greatly avoiding duplicate retrieval and calculation of data, and reducing server operating pressure. The data retrieval of traditional business system relational database is based on the Tree structure data of organization relationship. Taking equipment query as an example, when querying the account, you need to click the organization relationship, equipment name, account query and other buttons in turn. When querying the equipment status monitoring data, you need to repeat the above steps, that is, completing the equipment monitoring and account query requires more than 6 data retrieval operations, not only the complicated operation process, It also causes serious waste of resources and computing power to the backend server.

With the extensive access of new power system distribution and distributed new energy, its data volume is also rapidly expanding. The data volume generated by a single provincial power grid company's optical statistical meter will reach the PB level within 5–10 years. The traditional relational database is difficult to support such a large amount of data for elastic storage and topological relationship analysis; At the same time, the amount of data between different businesses of the power grid is increasing exponentially. The way of unprocessed direct data sharing will bring huge pressure on the server resources. The traditional data architecture based on the tree organization relationship, simple query and retrieval operations require multiple clicks to select, not only causing the complexity of the operation, but also causing a serious waste of resources and computing power to the background server [6].

9.2 A New Energy Distribution Data Configuration Method

This paper designs an intelligent closed loop business data orchestration method. By applying artificial intelligence technology to data orchestration, the transformation from "automation" to "intelligence" is realized. Through data integration across platforms and business users, the data orchestration architecture based on intelligent algorithms is designed [7], and the association and topological relationship between data is intelligently orchestrated according to data scale and business domain, solve the huge pressure on the server caused by the repeated retrieval and query of data caused by the traditional data classification and blocking, and realize the intelligent recommendation and panoramic display of the associated data according to the data scale through the topological association method, so as to achieve efficient data intelligent arrangement, greatly avoid the repeated retrieval and calculation of data, and reduce the pressure on the server operation. The specific algorithm is described in the following steps:

Step 1: Use the data extraction and conversion tool to access the power grid data center and the real-time measurement center to obtain the power grid data resources. The power grid data is divided into offline data and online data. Data extraction tools can achieve intelligent extraction of temporal topology data in distribution networks. By collecting timestamp metadata stored in the topology inventory data of distribution and consumption through the power dedicated equipment management system, time information elements and their association with the topology model are automatically extracted. During topology data extraction, time dimension data is intelligently added to form temporal topology data, so we use data extraction tools to extract the offline data at the frequency of synchronization once an hour for the data center and use the conversion tool to convert the data format into the data format required by the system for easy storage, The data format mainly includes but is not limited to serial number, equipment ID, data content, responsible person, extraction time, etc.,The number of data types should be controlled within 100, the length of ID data is generally defined as an int integer, and the length is generally controlled within 32 characters, and the subsequent associated query is performed through equipment ID. Online data uses the data extraction tool to extract real-time data from the real-time measurement center at the frequency of synchronous acquisition once a minute and store it in the time series database after converting the format. The data format mainly includes but is not limited to serial number, equipment ID, status content, extraction time, etc.

Step 2: Design the closed-loop business data intelligent orchestration method, as shown in Fig. 9.1. The topology association orchestration of the data is carried out with the device as the core. The method is to quickly build the data chain through the instantiation of the model according to the business needs, while monitoring and learning the implementation of the strategy, constantly optimizing the design of the model and strategy, and achieving efficient and robust data intelligent orchestration. The orchestration method is mainly divided into three modules: model library, execution engine and artificial intelligence module.

Step 3: Classify the data collected in Step 1 into the model inventory storage, and store the data according to the data type, so as to improve the efficiency of data storage and analysis, as shown in Fig. 9.2. The model library is a tool library for orchestration methods, which realizes the modeling of resource, business, and system operation and maintenance. At the same time, it designs the entire process of business deployment and creates standardized strategies. The model base framework can reuse the models of resources, services and system operation and maintenance,

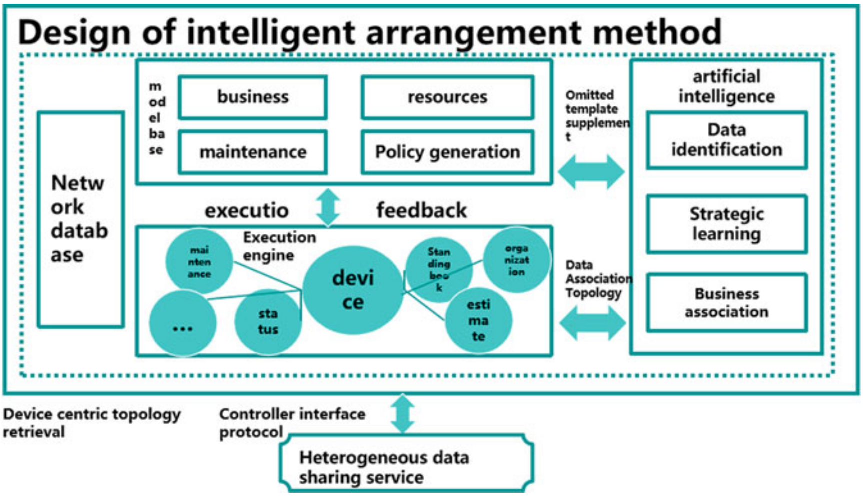


Fig. 9.1 Technical architecture of intelligent arrangement of power grid data sharing service

and design specific data formats and types according to different grid business fields, avoiding the repeated design of similar business requests, and improving the efficiency of large-scale business arrangement. At the same time, the standardization strategy can make the design-state module shield the heterogeneity of various data sources in the execution engine, and improve the compatibility of the orchestration method. The model base is used to realize the modeling of resources, business, operation and maintenance and the generation of strategies. Based on the data center, it is reclassified according to the new business characteristics, and classified according to the division principle of "strong internal correlation and loose external coupling" of the data subject domain to form the digital platform data subject plan. At present, 16 system data subject domains are planned, and 65 types of data entities are as shown in Fig. 9.2, covering SGCIM personnel, materials, projects, assets, power grids 8 company-level data topics, such as customer, security and integration. Among them, structured offline data includes 52 types of data entities in data subject areas such as topology graph, equipment topology, equipment abnormality, operation and inspection, operation operation, power grid resources, engineering projects, basic management, model management, graphic entities, location environment, security control, etc. All structured offline data are classified and stored in network graph database; Structured online real-time data includes 4 types of real-time data entities in the subject area of collected measurement data. After docking and integration from the measurement center, they are classified and stored in the TDEngine timing database to improve the efficiency of high-speed real-time data collection and response. Unstructured data are offline data, mainly including 9 types of data entities in three types of data subject areas, including 3D model, multimedia and GIS. Distributed file storage database is used for classified storage, and the file name rule is GUID_ Document type abbreviation_The file name is abbreviated, for example, the point cloud model file name of X substation is GUID_ SMDY_ XBDDYMX. Through the standardized naming method and data classification storage mechanism, data retrieval can automatically conduct holographic topology and fast retrieval of the equipment's related account, operation and maintenance, status, evaluation, organization relationship and other data throughout the life cycle of the equipment through the equipment object, greatly improving the retrieval efficiency and reducing the pressure on the server database and computing resources [8].

Step 4: The model base generates the strategy according to the business request and sends it to the execution engine to instantiate the topology of the data required by the business. The execution engine is used to implement the strategy. The execution engine is the executor of the data arrangement. The strategy issued by the model base is parsed according to the standardized template to realize the instantiation of the business data. Based on the network graph database, the flexible expansion storage and fast topology analysis of the power grid data resources are realized, With the equipment ID described in step 1 as the index core, holographic topology and fast retrieval of the whole life cycle data associated with the equipment such as the equipment associated account, operation and maintenance, monitoring, evaluation, status, etc. are realized, which greatly improves the retrieval efficiency and reduces the pressure on the server database and computing power.

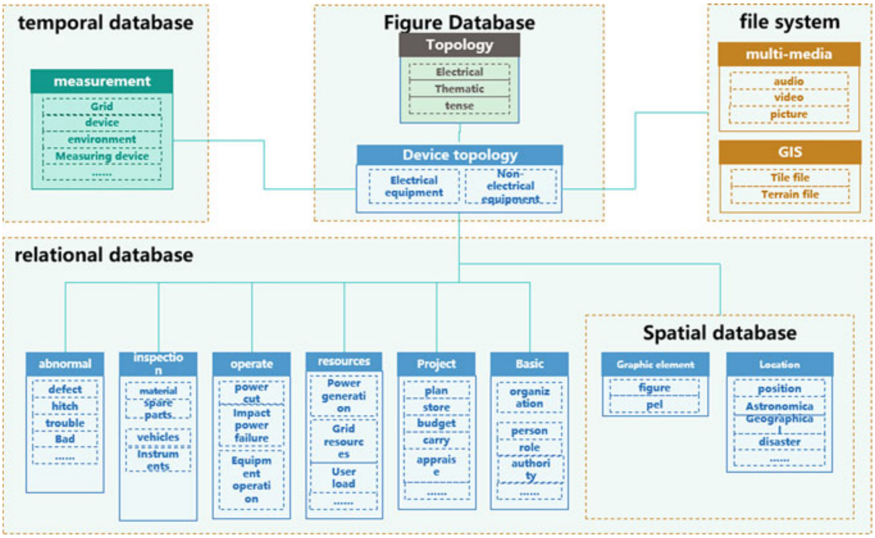


Fig. 9.2 Data classification storage architecture

Step 5: Import the topology instantiation data after the analysis of the execution engine strategy in Step 4 into the AI module for training, output the training results to optimize the execution of the controller. After the training effect reaches stability, the model can be formed and added to the design state, and the cycle is repeated to realize the intelligent closed-loop. The strategy learning and optimization is realized by using the AI module, including many components carried on the cloud platform such as modeling, strategy design, data monitoring, data association, topology analysis, etc. The topology instantiation data after the execution engine strategy analysis is used as input, and various types of machine learning technologies are used to train strategies in data recognition, business strategy classification learning, business association analysis, etc., The topology association data after the optimization strategy is output to realize the rapid retrieval and association of data with the device as the core. The AI module is based on the micro-service architecture and standardized component interfaces. Each component can evolve independently and accelerate the iteration process. When deploying business, users can select the required components to integrate into their own business to speed up business deployment [9].

Step 6: Repeat steps 3–5 to form an intelligent closed-loop data classification storage, strategy generation, strategy implementation, data instantiation, training optimization, and finally realize the intelligent arrangement of power distribution and distributed new energy data resources, and improve the classification storage, efficient calculation and intelligent arrangement of the system for massive power distribution and distributed new energy data resources.

Through the complex network full node association mode and breadth/depth traversal method, improve the efficiency of accurate data search under massive data, achieve optimization and design of data interaction, data processing, data services, operation management, and node management, and significantly improve the efficiency of data resource sharing services. Data interaction services are not just a data interaction interface, but should start from the user experience dimension and meet the requirements of providing services such as reasonable page layout, general operating specifications, error handling, and humanized feedback and prompts. Developers should be able to visualize the configuration and development of data interaction, data processing, data services, operation management, and node management through development tools. Therefore, a data interaction application architecture based on graph databases is proposed, as shown in the figure. Utilize the data access services deployed in companies in various cities to collect and obtain raw data from various heterogeneous business systems, build a data model, and apply for a service thread. Call the database collection plug-in through the plug-in scheduler to perform preliminary data verification. At the same time, use the message queue delivery plug-in to achieve asynchronous data transmission services; After the data receiving service deployed in the provincial company obtains the data sent by the prefecture/city company, it creates a resource thread, asynchronously verifies and delivers the data through the plug-in scheduler, and ultimately stores it in the provincial company server graph database in a topological relationship manner.

(1) Data interaction mode. When business systems interact with data, the timeliness requirements, data volume, and initiators of data interaction vary. These requirements should be met and flexible configurations can be made. The response modes of interaction are divided into synchronous and asynchronous. When business systems interact with data, different data interaction processes are required due to the size of the data volume and the form of interaction between applications. When the amount of data is large and the data being interacted does not have dependencies such as sequence and reference, asynchronous methods are used for processing; Use synchronous interaction when data is small, requires immediate response, and has interdependent relationships. The timeliness requirements for interactions are divided into immediate and automatic. Some business data interactions occur immediately upon triggering, while others occur periodically. Some data interactions need to consider server performance issues and occur periodically at a certain point in time. Therefore, multiple data interaction triggering methods need to be defined. The interaction initiator is divided into sending and extracting. During data interaction between business systems, there are both active push and passive data retrieval. Data interaction specifications need to define different methods [10].
(2) Data interaction format. The foundation of data interaction is a mutually recognized set of data and interaction protocols. Extensible Markup Language (XML) is the best choice for data bearer and transmission due to its scalability, fidelity, strong interoperability, and standardization. XML uses a set of tags to depict data elements. Each element encapsulates simple or complex data, defining an unlimited set of XML tags.
(3) Data interface mode. Using Web Service for data interface design, Web Service is a service-oriented architecture technology that provides services through

standard Web protocols, ensuring interoperability of application services from different platforms, and achieving loose coupling with business systems. The service interface design of the data bus follows the message model defined in IEC 61,968-1, and the message structure should include three elements: verb, noun, and message body. A noun indicates the type of message body requesting or responding. The service interface is defined using the Web Services Description Language (WSDL) standard, defining request, response, and error messages in one or more operations, and defining structured message bodies using XML Schemas (XSDs). The data bus implements a universal data service, which provides a universal service interface for receiving and parsing service request messages from data synchronization service callers. The routing of the data service is implemented based on the verbs and nouns contained in the header; Finally, the request message is routed to the data synchronization service interface of the specified service provider to invoke the data service. The data interaction service caller interacts with the general service interface of the data bus general data service, sends a service request message in accordance with the IEC 61,968-1 message model, and parses the service response message. The data interaction service publisher implements specific data services, provides data interaction functions, receives request messages input from the data bus, and returns a response message to the data bus after the service is completed.

(4) Frequency of data interaction. According to the frequency rule of data interaction, data interaction can be generally divided into periodic and non-periodic interactions. Periodic interactions refer to data interactions that occur over time periods, such as by day. Non-periodic interaction enables on-demand interaction and can be incremental.

(5) Data interaction quality monitoring. Since interaction occurs between different platforms, whether data can arrive completely and consistently, how the process of data interaction is, the stability of data interaction, and the resource utilization of the data platform are all concerns of users on both sides of the data interaction platform. Therefore, interaction quality monitoring and management components must be designed to monitor the interaction process.

9.3 Conclusions

This paper innovatively reads the corresponding program files automatically according to the document configurations of different database manufacturers, defines the relevant component orchestration process according to different business scenarios, builds a virtual data link network through data integration across platforms and business users, establishes a power grid data directory, understands the data and its business meaning, and establishes a knowledge map, and builds an independently evolving power grid knowledge base based on artificial intelligence technology, In combination with terminal capabilities, business characteristics and other factors, intelligent orchestration is introduced into the data sharing interaction, so that data can be easily located and classified. Topological elastic storage and efficient association retrieval of the association between massive data are realized by using topological computing and network-based databases, and the intelligent orchestration and management of data from multiple sources in various business areas are realized, and the data value is realized by different areas. At the same time, only necessary data will be transmitted and collected to reduce the cost of data replication and calculation and improve the quality and efficiency of data sharing. However, the data orchestration method mentioned in this article has only been tested on small-scale data of tens of thousands of levels, and has not been validated in large-scale distribution network data of over a million levels. It is necessary to further expand the practical scope of the method proposed in this article to improve the efficiency and quality of orchestration in the face of large-scale multidimensional data.

Acknowledgements This work was financially supported by the science and technology project to State Grid Corporation "Research on new distribution network data deep fusion and resource sharing technology(5400-202258431A-2-0-ZN)".

References

1. Ferraiolo, D.F., Kuhn, D.R., Chandramouli, R.: Role-Based Access Control. Artech House (2003). ISBN:1580533701
2. Zhou, Y., Louis, T.A.: A smoothing approach for masking spatial data. Appl. Stat. **4**(3), 1451–1475 (2010)
3. James, P.A. Co.: Computer security threat monitoring and surveillance. Box 42 Fort Washington, USA, Pa. 19034 (1980)
4. Van Schyndel, R.G.; Tirkel, A.Z.; Osborne, C.F.: A digital watermark. In: Proceedings of the IEEE International Conference on Image Processing, vol. 82, pp. 86–90 (1994)
5. Alam, F., Mehmood, R., Katib, I., et al.: Data fusion and IoT for smart ubiquitous environments: a survey. IEEE Access **5**(0), 1 (2017)
6. Huang, M., Liu, Z., Tao, Y.: Mechanical fault diagnosis and prediction in IoT based on multi-source sensing data fusion. Simul. Model. Pract. Theory **102**, 1 (2020)
7. Zhang, L., Gao, H., Wen, J., Li, S., Liu, Q.: A deep learning-based recognition method forde gradation monitoring of ball screw with multi-sensor data fusion. Microelectron. Reliab. **75**, 215–222 (2017)
8. Bijarbooneh, F.H., Du, W., Ngai, C.H., et al.: Cloud-assisted data fusion and sensor selection for internet of things. IEEE Internet Things J. **3**(3):257–268 (2016)
9. Beliakov, G., Gagolewski, M., James, S. Hierarchical data fusion processes involving the Mbius representation of capacities. Fuzzy Sets Syst. (2) (2021)
10. Rumelhart, D.E., McCLelland, J.L.: Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Language (1986)

Part II
Information Security and Power Management Control

Chapter 10
Design of Information Security Protection System for Cloud Business System

Yanbin Zhang, Lingyun Wang, Dan Liu, Yang Su, Yafeng Zhu, and Xike Zhang

Abstract As one of the contemporary information infrastructures, the cloud platform has become a vital backbone for the digital transformation of businesses, and has disrupted the conventional methods of information architecture. With an increasing number of business services being transferred to the cloud, the digitization and interconnectivity of information systems have witnessed a continual surge, thereby exposing these systems to mounting network security threats. Consequently, it has become imperative to prioritize the security of such systems. To ensure the information security of cloud-based business systems, this paper proposes an information security protection system designed for cloud-based business systems. The design of this system must consider various factors, such as the security design of the cloud environment, including the security of the physical environment, the communication environment, and the computing environment. The design should also account for data security classification and management, data security protection, security management strategies and systems, security management institutions and personnel, security construction management, security operation and maintenance management, network topology, access control, intrusion detection and prevention, data protection, and disaster recovery. In order to create a comprehensive and effective security strategy, a multi-layered approach that combines physical and virtual security measures is required. This includes the implementation of firewalls, VPNs, encryption, and regular security audits and updates. Proper training and education of system administrators and users are also critical to ensure the success of the security

Y. Zhang · Y. Su · Y. Zhu
Guangzhou Customs Technology Center, Guangzhou, China

L. Wang
Guangzhou Customs, Guangzhou, China

D. Liu
China Electronic Port Data Center Branch Guangzhou, Guangzhou, China

X. Zhang (✉)
Shenzhen National Climate Observatory, Shenzhen, China
e-mail: kokocheung@hotmail.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_10

system. The implementation of a robust network security protection system can help safeguard sensitive data and prevent cyber-attacks on cloud-based business systems.

10.1 Introduction

In the current era of digital information, the development of various business systems has brought convenience to enterprises, but at the same time, it has also brought significant information security risks [1]. With the continuous development of network attack technology and the upgrading of hacker attack methods, the security situation of enterprise business systems is becoming increasingly severe, and incidents of information leakage and data destruction are not uncommon [2]. Therefore, establishing a sound information security protection system for business systems is particularly important [3]. The design of an information security protection system for business systems can not only effectively protect the core data and business systems of enterprises but also provide secure and stable information services, thus enhancing the core competitiveness of enterprises [4]. At the same time, establishing a sound information security protection system for business systems is also a necessary measure for enterprises to fulfill their social responsibilities [5]. Therefore, researching how to design and construct an information security protection system for business systems that is suitable for the characteristics of the enterprise itself is of great significance for safeguarding enterprise information security and enhancing enterprise core competitiveness.

With the development of the Internet and information technology, the application of business systems in enterprises has become increasingly widespread. Ensuring the security of business systems has become an important aspect of enterprise information security. To address this issue, researchers have conducted a great deal of research and have designed various business system information security protection systems. Many researchers at home and abroad have devoted themselves to the design and research of business system information security protection systems. Among them, Chen et al. proposed a business system information security protection scheme based on a multi-level defense system, which includes five levels, data protection, application program protection, network protection, host protection, and security management, and details the security measures at each level [6]. Fan et al. have developed a design and implementation plan for business system information security protection based on the previous scheme, combined with actual situations, and strengthened network security protection and encrypted data transmission measures to achieve secure and controllable systems [7]. Chen et al. proposed a business system information security protection design scheme based on access control technology, which combines Role-Based Access Control (RBAC) and Mandatory Access Control (MAC) access control technologies to effectively control user access permissions and safeguard business system information security [8]. Wang et al. proposed a business system information security protection method based on artificial intelligence, which uses machine learning technology to analyze the system, and enhances the intelligence level of the protection system [9]. Zhu et al. proposed a business system information security protection method based on cloud computing technology, which improves the system's security and scalability by migrating system data to the cloud [10]. The design of business system information security protection systems also involves the application of new technologies such as blockchain technology and the Internet of Things (IoT). In recent years, some researchers have proposed business system information security protection schemes based on blockchain technology, which improves system security and trustworthiness through blockchain's decentralization and tamper-proof characteristics [11]. At the same time, the application of IoT technology can also realize real-time monitoring and early warning of enterprise information security [12]. Foreign researchers, including Christos et al., have conducted extensive security analyses on business process models, proposing a model-checking-based security detection method for business process model security, which has been applied in practical systems and achieved good results [13]. Nizar K. et al. analyzed the challenges and opportunities that business systems face in terms of security requirements, threats, and mechanisms, and proposed some countermeasures, providing a certain reference value for business system security protection [14].

In summary, the design of a business system network security protection system is a complex system engineering task that requires the combination of multiple technical and management methods to improve the security and reliability of business systems. Research on the design of business system network security protection systems to improve the security of business systems has become an urgent issue in the construction of informatization. This article aims to explore how to strengthen the security of business systems through the study of the design of business system network security protection systems, ensure the stability, efficiency, and reliability of business systems in the construction of informatization, propose a complete network security protection system design, and provide feasible solutions for the security of business systems.

10.2 Cloud Security Design

The business system is deployed within the Big Data Resource Management Center (BDRMC), and the cloud platform offers security services that conform to the Level 3 security standard, which is a prerequisite for the primary system. The BDRMC bears the responsibility of securing the infrastructure, including the machine room environment, cloud platform, and necessary Level 3 security equipment. Furthermore, the absence of a backup plan in a different location may lead to significant risks of data loss in case of physical damage to storage or sudden natural disasters without any recovery mechanism. Therefore, to address the mandatory requirements of city-level backup services and off-site backup, it is imperative to deploy a backup and recovery system. By applying for a backup center in the BDRMC's machine room and utilizing communication networks, critical data can be backed up in real time to

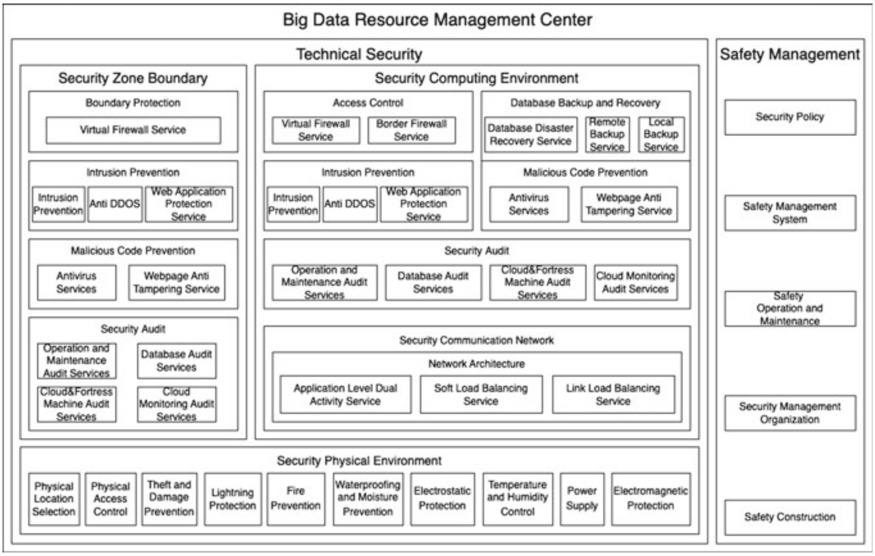


Fig. 10.1 Information security architecture of the cloud in the BDRMC

the backup site for backup and recovery. Furthermore, the disaster recovery system must comply with the Level 3 security standard, as per the protection level standard requirements. The information security architecture of the cloud in the BDRMC is categorized into three components, namely secure physical environment, technical security, and security management, as depicted in Fig. 10.1.

10.2.1 Security Physical Environment

Physical and environmental protection involves safeguarding computer network equipment, facilities, and media against environmental disasters like earthquakes, floods, and fires, as well as human errors, mistakes, and malicious computer crimes that can result in destruction. The business system facilities are deployed on the BDRMC's cloud platform, which must satisfy Level 3 protection requirements. This entails preventing unauthorized access by individuals and criminals who could exploit the network's host to steal information, compromise the integrity and availability of data, and cause damage to the network and host systems. To prevent attacks by illegitimate users, the BDRMC uses effective regional monitoring and anti-theft alarm systems. Strict access control and environmental monitoring systems are also in place to guarantee the smooth functioning of the regional and environmental monitoring systems. Media is categorized and stored in media libraries or archives, while optical and electrical technologies are leveraged to establish anti-theft and monitoring alarm systems for the computer room.

(1) Physical location selection

The computer room should be situated in a structurally sound building that is capable of withstanding earthquakes, high winds, and heavy rain. The building housing the computer room and office should possess verified documentation for seismic fortification. It is advised to refrain from locating the computer room on the topmost or lowermost level of the building. However, if there is no alternative, enhanced waterproofing and moisture-resistant measures must be implemented.

(2) Physical access control

To meet Level 3 security requirements, the computer room entrance and exit must be fitted with an electronic access control system that controls, identifies, and registers personnel accessing the room, thereby enhancing the physical access management of the computer room.

(3) Anti-theft and damage measures

Ensure the proper maintenance of computer room equipment or main components and affix clear and tamper-proof identification marks. Communication cables should be discreetly placed, either underground or inside conduits. Moreover, the computer room should be furnished with an anti-theft alarm system or a video surveillance system, which must be monitored by designated personnel.

(4) Lightning protection

All types of cabinets, facilities, and equipment in the computer room should be adequately grounded through a robust grounding system. Additionally, measures should be implemented to prevent induced lightning, such as the installation of lightning protection devices or overvoltage protection devices.

(5) Fire Prevention

Install an automatic fire suppression system that can detect, notify, and extinguish fires automatically. The data center, along with its associated workspaces and auxiliary rooms, should be constructed with fire-resistant materials that meet the appropriate fire resistance ratings. Furthermore, the data center should be subdivided into sections, and fire isolation measures should be implemented between them to prevent the spread of fire.

(6) Waterproofing and Moisture Prevention

Appropriate measures must be implemented to prevent rainwater from seeping through the windows, roofs, and walls of the data center. Moreover, the data center should be equipped with mechanisms to prevent water vapor condensation, as well as the infiltration and transfer of underground water by deploying water-sensitive monitoring instruments or components. Regular waterproof testing and alarms should be carried out to ensure the water resistance of the data center.

(7) Electrostatic Protection

Install an electrostatic discharge (ESD) flooring system and implement essential grounding and anti-static measures to prevent the accumulation of static electricity. These measures may include the use of electrostatic eliminators and anti-static wristbands.

(8) Temperature and Humidity Control

It is recommended to install an automatic temperature and humidity adjustment system, also known as an air conditioning system, to maintain the temperature and humidity levels within the range permitted for the operation of the equipment.

(9) Power Supply

Incorporate voltage stabilizers and surge protectors, uninterruptible power supply (UPS) systems, as well as redundant or parallel power cable lines to provide reliable power supply to the computer system.

(10) Electromagnetic Protection

To prevent mutual interference, power and communication cables should be isolated from each other. Additionally, electromagnetic shielding should be provided for critical equipment to mitigate the impact of electromagnetic interference.

10.2.2 Security Communication Network

The design and implementation of secure communication networks involve two critical aspects: network architecture and communication transmission security measures.

(1) Network architecture

The network structure provided by the BDRMC constitutes the basis and prerequisite for ensuring network security. To ensure hardware redundancy of communication lines and key network devices, and to guarantee system availability, the information system should rationally plan the network and draw a network topology diagram that aligns with the current operation situation. Considering factors such as information importance and the degree of relevance, different network segments or VLANs are established, and a secure path is created between business terminals and business servers. The network segment where important business systems and data are stored must be partitioned separately and isolated from external systems. Additionally, technical measures such as network device traffic control should be employed to prevent important businesses from being affected by network congestion. To meet the processing capacity requirements during peak periods and ensure that the bandwidth of each part is suitable, the business processing capacity of network devices must be sufficient.

(2) Communication transmission

Employ a data encryption mechanism to guarantee the verification of integrity, and to maintain the confidentiality of the sensitive information fields or the complete messages while transmitting them.

10.2.3 Security Zone Boundary

(1) Boundary Protection

The BDRMC provides security service components to implement access control functions, which ensure that access and data flows crossing the boundary communicate only through controlled interfaces provided by the boundary protection of security service components. Non-authorized devices attempting to connect to the internal network are either restricted or checked by extending admission devices or other security measures through security service components. Additionally, unauthorized behavior of internal users attempting to connect to the external network is restricted.

(2) Access control

The information system boundary represents the clear division of security domains and units of security control. To ensure authorized access and prevent illegal attacks, the firewall services provided by BDRMC should be deployed at the network boundary. All data packets passing through the firewall should be filtered based on strict security rules, and all unsafe or non-compliant data packets should be blocked. Access control should be based on application protocols and content. For access between various network areas, firewall and VLAN partitioning should be utilized. Access control list policies should be set on core switches to restrict end users from directly accessing the security management zone. In addition, IP and MAC address binding should be applied to important network segments and devices.

(3) Intrusion Prevention

The security service components offered by the BDRMC provide intrusion prevention capabilities, which enable dynamic traffic detection within the network through the use of intrusion detection functions. Additionally, the intrusion detection devices' feature library is regularly updated to promptly detect any abnormal behavior present in the network.

(4) Malware Prevention

The BDRMC offers antivirus services to detect and eradicate malware that may exist in the network and hosts. Additionally, it conducts vulnerability scanning and promptly carries out system patch updates. Regular upgrades and updates are implemented to ensure that the system remains secure and protected against potential threats.

(5) Security Audit

The network audit service component conducts security audits on the network boundary and critical network nodes, encompassing all users. The audit log should contain the event date and time, user, event type, whether the event was successful, and other relevant audit information. Audit records should be safeguarded, regularly backed up, and protected from unexpected deletion, modification, or overwriting. The retention period for audit records should be at least 6 months and should not be interrupted.

10.2.4 Security Computing Environment

The computing environment is designed to provide secure protection for servers, network security devices, and classified data objects, safeguarding them from various security threats through identity authentication, access control, security auditing, intrusion prevention, malware prevention, data integrity, data confidentiality, data backup and recovery, and residual information protection measures.

(1) Identity Authentication

Users accessing servers and network security devices must undergo rigorous identity authentication using unique identification and complex authentication information that changes periodically. This may include configuring usernames and passwords with at least three characters and a minimum length of eight digits, enabling login failure processing functions, implementing measures such as session termination, restricting illegal login attempts, and automatic logout. Additionally, necessary precautions must be taken to prevent authentication information from being intercepted during network transmission. This may involve utilizing VPN or cloud bastion machine services to ensure secure protection of remote management.

(2) Access Control

The host and system access control policy for the classified system mandates the implementation of server security reinforcement services provided by the BDRMC. To achieve this, access rights of default accounts should be restricted, system default accounts renamed, default account passwords modified, and expired or redundant accounts deleted from the operating system and database, as well as disabling useless or shared accounts. Permission allocation should be based on the role of the management user to ensure separation of management user permissions, granting only the minimum permissions required for the management user, enabling access control functions, and controlling user access to resources according to security policies. To control access to the server, different network segments and access control permissions should be set on the switch and firewall. Default shares opened by the operating system should be closed, and different access permissions should be set for necessary shares and shared folders. It is also important to set permission requirements for important operating system files and directories.

Multiple administrators should be assigned to manage the server, including system administrators, security administrators, and security auditors, to achieve permission separation for operating system-privileged users, and to set the minimum permissions for each account within their scope of work.

(3) Security Auditing

Security auditing of equipment and computing infrastructure is facilitated by the log auditing and database auditing services provided by BDRMC. Furthermore, the self-auditing feature is enabled on host systems, security devices, and switches to support the on-demand auditing needs. The Network Time Protocol (NTP) is utilized to ensure the accuracy of audit records by aligning the system scope clock. Audit records are securely protected, regularly backed up, and shielded from accidental deletion, modification, or overwriting. The audit logs are retained for at least six months, and the audit process is secured against unauthorized disruption.

(4) Intrusion Prevention

To prevent host system intrusions in the information system, the operating system should follow the principle of minimal installation. This involves only installing the necessary components and applications while disabling any unnecessary system services, default shares, and high-risk ports. Additionally, it is important to keep the system patches up-to-date through the use of an upgrade server or patch distribution service in order to enhance intrusion prevention capabilities. The network antivirus service is also an important tool to detect any intrusion behavior of important nodes, and to provide alarms in the event of a serious intrusion.

(5) Malware Prevention

The network-wide antivirus service is employed across all servers to enhance virus protection capabilities, and to ensure timely updates of software versions and malware libraries for the detection and elimination of malware. Additionally, the antivirus service integrated with the firewall component is utilized to eradicate malware from the network.

(6) Data Integrity and Confidentiality

To ensure the integrity and confidentiality of data, VPN services are used to protect network transmission layer data integrity and confidentiality. This includes using VPN services to encrypt transmissions to same-city or remote backup centers. For highly sensitive data, data encryption systems are employed to guarantee the integrity and confidentiality of critical management data, authentication information, and important business data storage.

Authentication information and important business data are encrypted during transmission and storage to guarantee the encrypted transmission of data. Usernames, passwords, and form data should be encrypted before being stored in the database to prevent sensitive information from being obtained and avoid information leakage.

Link encryption devices are used to ensure data confidentiality. The BDRMC provides a database security service with features such as sensitive data discovery, data desensitization, and database attack prevention using reverse proxy and machine learning mechanisms to ensure cloud database security.

The BDRMC's database auditing service and log auditing service record, analyze, and report user access behavior to databases, assisting users in creating compliance reports and tracking accident investigations. Recording internal and external database network behavior improves data asset security.

The BDRMC's security vulnerability scanning service scans for weak passwords, vulnerabilities, weak security configurations, and backdoors, verifying and repairing database vulnerabilities.

The BDRMC's operations auditing service enables auditability of bastion host operation and centralized management of database service operations.

The BDRMC's web application protection service identifies and updates application vulnerability patches, and provides features such as prevention of SQL injection and protection against application-side data leakage.

(7) Data Backup and Recovery

A data backup refers to the process of replicating the entire system or a subset of data to alternative storage media in order to prevent any compromise to data availability and integrity that may arise due to operational errors, system failures, or human factors. Data recovery involves restoring data to a specific point in time using valid backup data, as and when required.

With regard to data backup and recovery, the system provides local backup and recovery functions and real-time remote backup of critical data. Hot redundancy of important data processing systems, including boundary switches, boundary firewalls, core routers, application services, and database services, is also provided to ensure high system availability. Through backup and recovery services, backup centers are established, and important data is backed up in real time to backup sites using communication networks, achieving data backup and recovery. A full data backup of important application systems is carried out every day, and backup media is stored off-site, with designated backup and recovery strategies. For major network and security equipment policies, periodic backups are exported.

A data backup and recovery system should be deployed in the system management and operation domain to implement different security policies for critical data resources of core production systems and Demilitarized Zone (DMZ). A regular recovery testing plan should be developed to achieve periodic recovery testing and validate the integrity of backup data.

(8) Protection of residual information

Preventive measures must be taken to mitigate the potential risks associated with residual information, and it is crucial to ensure that the storage space containing residual information of users is thoroughly cleared before being released or reallocated to other users upon their departure.

When devices are replaced, a comprehensive data erasure process should be carried out, and clearance ought to be conducted on individual files, folders, and the residual disk space.

Table 10.1 Data classification

Data classification	Data level control requirements	Whether to collect, use, and store data of this category
Personal information data	Name, mobile, email, account, and password	Yes
Business data	System business data	Yes

10.3 Data Security Design

10.3.1 Data Security Protection Design

Data Security Classification and Grading Management.

The data carried by the business system should be managed in accordance with the requirements outlined in the "Data Classification and Grading Guidelines" (DB52_T 1123-2016) and "Data Classification and Grading Guidelines" (GT3-SP-PE-20120520), which specify the classification and grading of data.

Data Classification.

Data is categorized into five distinct groups based on its purpose, which includes business data and personal information data. The classification of data is presented in Table 10.1.

Data Grading.

The grading of data must take into account the significance of the data to national security, social stability, and citizen safety, as well as the presence of sensitive information, such as state secrets and user privacy. The classification of data should be determined by assessing the extent of harm to national security, public order, public interest, and legal rights and interests of individuals, legal entities, and other organizations (the targets of infringement) in the event of damage to data of varying degrees of sensitivity. Please refer to Table 10.2 for the data grading system.

10.3.2 Data Security Protection Design

In accordance with the Cybersecurity Law 2.0, data security protection comprises three primary elements: guaranteeing the confidentiality and integrity of data during transmission, guaranteeing the confidentiality and integrity of data during local storage, and replicating the entire system or a portion of the dataset to other storage media while using effective backup data to restore data at a specific point in time when data integrity is unexpectedly compromised.

Table 10.2 Data level

Data level	Data level control requirements	Whether to collect, use, and store data of this category
Open data	1. Unrestricted sharing between business 2. Can be completely accessible to the public	Yes
Internal data	1. As a general rule, business should share their internal data unconditionally. However, in cases where sensitive data may have an impact on the rights and interests of citizens, legal entities, and other organizations, the sharing of such data should be subject to certain conditions set by the business 2. The decision to open internal data should be made in compliance with national laws and regulations. Generally, data should be opened or de-identified as long as it is in accordance with the relevant national laws and regulations	Yes
Confidential data	Involved State secrets concerning national security	No

Data Communication Transmission Security Protection.

The measures to ensure data security during communication transmission are as follows:

(1) The integrity of data during communication should be ensured through the use of checksum or encryption technologies.
(2) The confidentiality of data during communication should be ensured through encryption technologies.

To comply with these requirements, we will implement a newly acquired data encryption system in this project. This system will perform data integrity checks and channel encryption, guaranteeing the integrity of communication and confidentiality of sensitive information fields or complete messages during communication.

Data Local Storage Security Protection.

The primary data that necessitates safeguarding in local storage are personal data retained in the emergency warning system and emergency warning data that pertains to emergency warning business.

The protective measures for securing personal information storage predominantly encompass two facets:

(1) Utilize encryption algorithms to conduct integrity verification and data encryption on personal information to guarantee the confidentiality and integrity of the data. Before storing usernames, passwords, and submitted data in the database, they must be encrypted to avoid access to crucial system information and prevent data leakage.

Apply data anonymization methods to obfuscate personal details, such as phone numbers and identification numbers, prior to preserving them in the database, as a measure to mitigate the risk of pilferage and misuse of sensitive personal information.

To safeguard emergency warning business data, encryption algorithms are primarily utilized to perform integrity checks and encrypt personal information, ensuring data integrity and confidentiality.

Besides utilizing technical measures to protect the security of the core data storage, the security of the data storage environment itself has been strengthened by conducting rigorous testing. In this regard, security vulnerability scanning equipment deployed in the business' network was employed to scan for weak passwords, vulnerabilities, feeble security configurations, and backdoors. Additionally, this equipment was utilized to verify and rectify any database vulnerabilities found during the scanning process.

Data Usage Security Protection.

Through the utilization of web application protection services provided by BDRMC, features such as application SQL injection prevention and XSS protection are implemented to monitor and safeguard business data activity. This ensures that business data adheres to normal business logic and prevents any data leakage on the application side.

Data Backup and Recovery.

Data backup refers to the process of replicating the entire system or a subset of data to other storage media to prevent data availability and integrity from being compromised due to operational errors, system failures, or human factors. Data recovery involves restoring data at a specified point in time using valid backup data as needed.

With regard to data backup and recovery, BDRMC provides local and remote backup services to back up and recover important data. The real-time backup function ensures system availability by using communication networks to back up essential data to backup sites. The center also offers data-level disaster recovery backup and application-level active-active services to ensure high system availability. Important application systems are fully backed up on a daily basis, and backup media is stored off-site with designated backup and recovery strategies. Backup export is periodically conducted for the main network and security devices.

A data backup and recovery system is implemented in the system management and operation domain to enforce different security policies for crucial data resources in the core production system and DMZ area. Regular recovery testing plans are established to validate the integrity of backup data.

10.4 Design of Security Management System

It is essential to establish a unified information security management system and implement various management systems. This "three part technology, seven part management" approach recognizes that technology and products serve as the foundation, but security management is the key. Establishing an excellent security management framework is crucial to implementing good security strategies repeatedly within the framework, achieving continuous security of the information system. System planning and design should consider aspects such as security management systems, security management institutions, personnel security management, system construction, and operation and maintenance management. A unified information security management system should be established, and various management systems should be implemented to ensure that the department's security management system has macroscopic design, clear responsibility and authority, and reasonable institutional requirements. Innovative technologies, such as security visualization and unified operation and maintenance management, should be applied to simplify security operation and maintenance management, reduce the burden of security operation and maintenance, and improve the efficiency of security operation and maintenance management. Ultimately, the aim should be to achieve overall defense and zone isolation, active protection, internal and external prevention, self-defense and active immunity, and depth defense with equal emphasis on technical management.

10.4.1 Security Management Policies and Systems

Based on the unique characteristics of the business' information security management work, it is essential to enhance the overall policy and security strategy of information security work, and provide clear guidance on the overall objectives, scope, principles, and security framework of security management work. It is crucial to establish security management systems for various management contents in security management activities and develop operating procedures for daily management operations executed by management personnel or operators. This will form a comprehensive information security management system consisting of security strategies, management systems, operating procedures, etc., and will guide and effectively regulate the information security management work of business at all levels. To ensure the effectiveness of the management system, it is necessary to formulate strict regulations on the process, methods, scope, and other aspects of the publication of the management system. The security management system should also be evaluated and revised periodically. Effective management systems can significantly reduce security issues caused by human factors. Additionally, when designing the management system for an information network system, it is important to consider the system's characteristics and the special environment in which the system is located. The design diagram of the security management strategy and system is presented in Fig. 10.2.

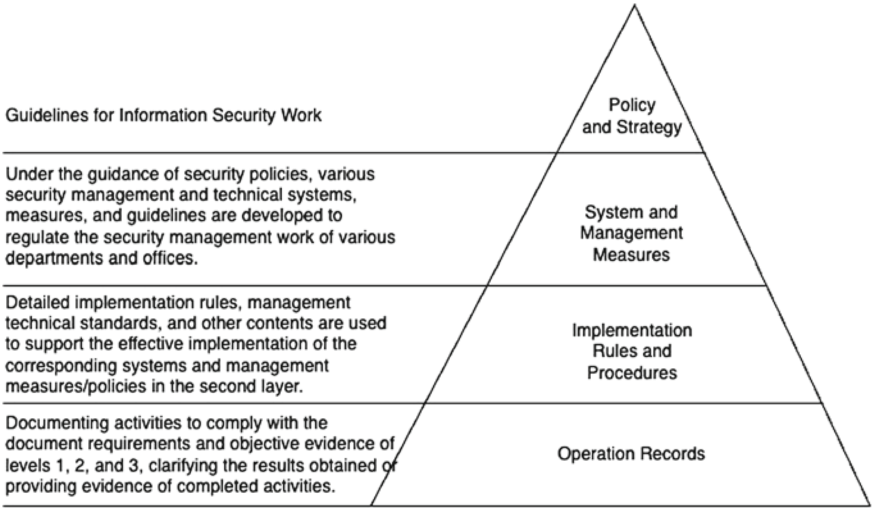


Fig. 10.2 Design diagram of security management strategy and system

10.4.2 Security Management Organization and Personnel

To align with the organizational structure and personnel characteristics of the business, it is vital to strengthen organizational structure of information security management. To enhance the efficiency and effectiveness information security management, it is essential to align the organizational structure with the business structure and personnel characteristics. A strong information security management team and other security management organizations should be established to clarify the form and mode of operation of the information security management organization. The roles and responsibilities of different positions should be clarify defined, such as system administrators, network administrators, security administrators, etc. It is also necessary to supervise and manage the allocation of personnel, authorization approval, communication and collaboration, audit and inspection, personnel recruitment, personnel departure, and security awareness education and training. The security management organization and personal structure is shown in Fig. 10.3. This structure allows for better coordination and communication between different departments and personnel to ensure the smooth operation of information security management.

10.4.3 Security Construction Management

From the viewpoint of information security management, it is necessary to enhance the information security-related content in the information system engineering construction management system. This includes the protection of information

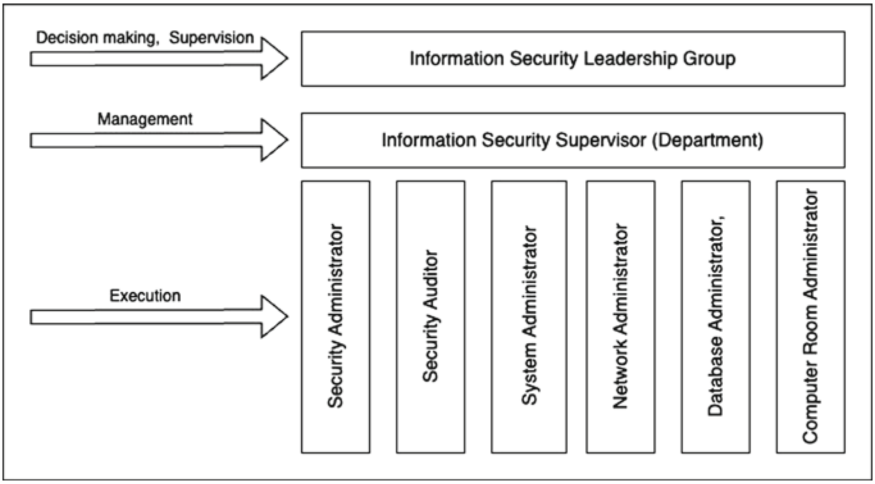


Fig. 10.3 Security management organization and stuff diagram

systems at different levels, such as classification, security design schemes, procurement and utilization of security products, self-developed software, outsourced software development, engineering execution, testing and acceptance, system delivery, and management of service providers. The security management strategy and content should be integrated throughout the entire engineering cycle, starting from the initial classification design phase to the acceptance and evaluation phase. Additionally, the security guidance and testing for the software development process and software delivery should be reinforced.

10.4.4 Security Operation and Maintenance Management

In accordance with the regulations governing information system security operations and maintenance within the framework of the business information security management system, it is imperative to utilize the physical environment, network system, information security protection, operation and maintenance management, monitoring, audit, and unified security management center to continuously enhance the measures and methods of system operation and maintenance security management, and strengthen the scientific norms of operation and maintenance security management. The objective is to ensure the secure and stable operation of the system. This includes the implementation of environmental management, asset management, medium management, equipment maintenance management, vulnerability and risk management, network and system security management, malicious code prevention and management, configuration management, password management, change management, backup and recovery management, security incident handling, emergency plan management, outsourced operation and maintenance management, etc.

In addition, it is crucial to establish and enhance daily scanning, detection, and reinforcement of network system security vulnerabilities, and to monitor and protect against malicious code. It is also important to implement mechanisms for security configuration changes, log auditing, analysis, data backup and recovery, and security event monitoring and notification, and to evaluate and verify the effectiveness of security policies and mechanisms.

10.5 Conclusions

The design of a network security protection system for cloud-based business systems is a critical component of digital transformation. This article proposes a multi-layered and all-encompassing security protection system design by analyzing the characteristics and potential security threats of cloud platforms. This design takes into account considerations related to physical, network, and data security, as well as security management strategies and systems, security management organizations and personnel, security construction management, and security operation and maintenance management. The security protection system adopts a range of security measures, including firewalls, VPNs, encryption, and periodic security audits and updates, to enable businesses to safeguard sensitive data and prevent cyber-attacks. Moreover, this article underscores the necessity of adequate training and education for system administrators and users. Through the implementation of these measures, cloud-based business systems can benefit from comprehensive protection in terms of network security, thereby providing reliable support for digital transformation.

References

1. Chen, X., Li, L.: Research on enterprise information security risk assessment model. Sci. Technol. Inf. **33**, 198–199 (2016)
2. Ding, Q.: Research on risk-based information security management. Sci. Technol. Innov. Appl. **5**, 86–88 (2015)
3. Wang, J.: Analysis of the application of information security technology in enterprises. Sci. Technol. Rev. **35**(11), 73–76 (2017)
4. Wu, Q., Yang, L.: Design of information security protection system based on strategy. Comput. Eng. **44**(7), 162–164 (2018)
5. Ma, Y., Li, J.: Research on enterprise information security risk management. Intell. Mag. **3**, 89–92 (2019)
6. Chen, X., Liu, Y., Chen, Q.: Information security protection scheme of business system based on multi-layer defense. Electron. Des. Eng. **29**(1), 1–3 (2021)
7. Fan, Y., Hong, B., Li, Y.: Design and implementation of information security protection for business systems. Commun. Technol. **54**(7), 27–30 (2021)
8. Chen, L., Fan, Y., Yang, B.: Design and implementation of information security protection for business systems. Comput. Digit. Eng. **46**(8), 1411–1416 (2018)

9. Wang, R., Zhang, Y., Li, Y., et al.: Information security protection method of business system based on artificial intelligence. Comput. Sci. **46**(9), 157–160 (2019)
10. Zhu, R., Zheng, Q., Chen, J.: Research on information security protection of business systems based on cloud computing technology. Comput. Digit. Eng. **5**, 46–47 (2019)
11. Li, P.: Research on information security protection scheme of business system based on blockchain technology. Inf. Technol. **37**(2), 1–3 (2020)
12. Lin, K., Wang, B., Lu, R.: Research on enterprise information security monitoring and early warning system based on Internet of Things technology. J. Comput. Sci. Explor. **13**(1), 63–71 (2019)
13. Christos, K., Kostas, T., Stefanos, G.: Security analysis of business process models: state of the art. Comput. Secur. **31**(3), 357–374 (2012)
14. Alsharari, N.K., Darshan, M.T., Muhammad, M.: Challenges and opportunities in business system security. J. Secur. Commun. Netw. **6**(1), 11 (2018)

Chapter 11
Industrial Area Power Load Forecasting Based on Seasonal Kalman Filter

Jinjin Li, Fanghua Mo, Qiuhua Chen, and Jun Chen

Abstract The stable and economic operation of the power system depends on accurate regional electricity load forecasting. This article proposes a seasonal Kalman filter prediction model based on Fourier transform spectrum maximization period analysis, which solves the problem of electricity load prediction in regions with different data distributions. The results indicate that the seasonal Kalman filter model has good predictive ability, low volatility, and stable error.

11.1 Introduction

Economic development has led to an increasing demand for electricity, and at the same time a rising demand for power quality. This leads to the rapid development of intelligent power systems. One of the decisive factors for the stable economic operation of the power system is the prediction of power demand [1]. Due to the influence of many random factors, such as society, society, economy, and natural conditions, the power load variation curve has a very complex nonlinear form. How to use historical data to make reasonable power demand forecasts has become one of the keys to intelligent power grid research. Zhou Xie made a detailed analysis of the main external factors affecting the load characteristic index of the power grid, studied the impact of climate factors, economic factors, holiday factors, and other factors on the power grid load, and sorted out the internal relationship of the load characteristic index. The research shows that the influence of various factors on power load has a

J. Li · J. Chen
Guangxi Power Grid Corporation Metering Center, Nanning Guangxi 530023, China

F. Mo (✉)
Guangxi Power Grid Corporation, Nanning Power Supply Bureau, Nanning Guangxi 530023, China
e-mail: 357030634@qq.com

Q. Chen
Guangxi Power Grid Corporation Wuzhou Power Supply Bureau, Wuzhou Guangxi 543000, China

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_11

certain randomness [2]. By studying the current situation at home and abroad, Wang Huizhong et al. briefly described the characteristics of short-term load forecasting and various factors affecting the forecasting accuracy, expounded the intelligent methods of short-term load forecasting for power systems, and analyzed and compared the advantages and disadvantages of various methods [3].

Due to the randomness of relying on external data for power load forecasting, and the power load of the same industry always being a certain periodicity, it is predictable. Therefore, many researchers tend to use the information of the power load curve directly to predict electricity consumption behavior. For example, Wang Jianjun et al., based on the historical data of China's electricity consumption from 1973 to 2004, fitted a similar exponential regression curve according to its trend chart, and then analyzed and identified its residual series using time series [4]. Kalman filter is a linear model based on minimized covariance estimation error, which has the advantages of simple calculation and a solid theoretical foundation. Yu Jingwen et al. briefly introduced the power quality problem and its analysis and detection methods. Then, the basic principles of three Kalman filters, namely conventional Kalman filter, extended Kalman filter, and untraced Kalman filter, are summarized systematically and their applications in power quality analysis are compared and analyzed [5]. Huang Z discusses the feasibility of applying Kalman filtering techniques, including estimating state models in dynamic variables. In the large and small interference tests of a multi-machine system, the Kalman filtering of the dynamic model is proposed, and the sensitivity analysis of the sampling rate and noise level shows that the dynamic state estimation performance is good [6]. Considering the characteristics of the power system itself, Ma Jingbo et al. built a load system model, observation model, and system parameter model with a time-varying coefficient based on the historical electricity consumption data at the same time on different dates. The time-varying noise statistical estimator is used to perform adaptive estimation of noise covariance, and the predictive equation is used to predict the load of the next day. The research shows that the predictive ability of the predictive equation of the adaptive time-varying noise estimator considering historical data is stronger than the general Kalman prediction model [7]. However, the study is based on the assumption that the electricity load is a stationary series at the same time every day. When generalized to the forecast of the total load of the whole day, the original assumption may not be valid, that is, the total load of the whole day does not necessarily constitute a stationary series with the load of the same period of the previous week or month. In the real production environment, the cycle rule is more complicated. Even for individual manufacturers in the same industry, the cycle of the electricity load is not the same.

Therefore, an adaptive Kalman filter prediction model (FFT-KF) combined with fast Fourier algorithm is proposed in this paper. Based on the fast Fourier model, the period of the power load time series curve is identified, the periodic shock factor is established, and the daily power load of various industries in each station area in the region is adaptively predicted, and then the daily power load in the whole region is obtained through linear combination. In order to analyze the adaptability of the FFT-KF model in theory and practice, this paper will compare the generalization ability of the FFT-KF model and classical Kalman filter in terms of prediction ability, in order to estimate the gap between the FFT-KF model and classical Kalman filter. Considering the influence of industry factors, this paper studies the generalization ability of the FFT-KF model under the influence of industry factors by analyzing the residual distribution and cumulative error of the FFT-KF model in power estimation of different industries.

11.2 Classic Kalman Filter Load Forecasting Model

The core idea of the classical Kalman filter is that according to the measured value at this moment and the predicted value and error at the previous moment, the weighted average is obtained at this moment, and the value of the next moment is predicted [8]. The errors are prediction errors and measurement errors, which exist independently and are not affected by measurement data. The predicted value of the previous time and the measured value of the current time are normal distributions, and the fusion of the two normal distributions can obtain the desired value more accurately. Kalman filtering first constructs a hidden Markov model for the system. Generally, the state of hidden Markov model cannot be directly observed, which needs to be divided into a state transition equation and observation equation, and then implemented by Bayes' theorem. First, estimate a prior probability based on previous experience, that is, Kalman's posterior estimate, and then add new information, that is, the measurement value, so that with the new information, the prediction of the value is more accurate [9, 10].

The Kalman filter model believes that the world is full of noise, and even the signal from the sensor will have various biases due to electromagnetic interference. In the power system, because the power load is composed of a certain sampling point, the sampling point itself is also affected by the accuracy of the power instrument. Based on the principle of minimum estimated mean square error, the Kalman filter model uses a recursive method to solve the linear filtering problem of discrete data [11].

The Kalman filter consists of a system state equation and a measurement equation, and follows the linear unbiased minimum mean square error estimation criterion to achieve recursive prediction using prediction and correction [12].

For a linear constant system state equation xk and measurement equation $y_{k,}$

$$x_k = \Phi_{k|k-1} x_{k-1} + \omega_{k-1} \tag{11.1}$$

$$y_k = H_k x_k + v_k \tag{11.2}$$

Among them, x_k and x_{k-1} are the electricity load values at time k and $k-1$, $\Phi_{k|k-1}$ is the correlation coefficient between k-1 and the electricity load at time k, and for stationary loads $\Phi_{k|k-1}$ is the identity matrix I, ω_{k-1} is the normal distribution process noise with the mean value of 0; y_k is the measured value of the electricity load at time k, and H_k is the measurement matrix, ν_{k-1} is the measurement noise of normal distribution with mean value of 0.

Because ω_{k-1}, ν_{k-1} is the normal distribution white Gaussian noise with the mean of 0, there is

$$\hat{x}_k = \Phi_{k|k-1}\hat{x}_{k-1} \tag{11.3}$$

$$\hat{y}_k = H_k\hat{x}_k \tag{11.4}$$

$\hat{x}_k$, $\hat{x}_{k-1}$, $\hat{y}_k$, are the estimated values of x_k, x_{k-1} and y_k, respectively.

By using the actual measurement value y_k at time k to correct the estimated electricity load $\hat{x}_k$, there is

$$\begin{aligned} \hat{x}_k &= \Phi_{k|k-1}\hat{x}_{k-1} + K_k(y_k - \hat{y}_k) \\ &= \Phi_{k|k-1}\hat{x}_{k-1} + K_k(y_k - H_k\Phi_{k|k-1}\hat{x}_{k-1}) \end{aligned} \tag{11.5}$$

where K_k is the gain matrix that minimizes the mean square deviation P_k of $(x_k - \hat{x}_k)$, i.e.

$$\begin{aligned} P_k &= E[(x_k - \hat{x})(x_k - \hat{x})^T] \\ &= \Phi_{k|k-1}P_{k-1}\Phi_{k|k-1}^T + Q_{k-1} \end{aligned} \tag{11.6}$$

$$K_k = P_k H_k^T (H_k P_k H_k^T + R_k)^{-1} \tag{11.7}$$

Among them, R_k is the measurement noise variance, and Q_{k-1} is the process noise variance.

Equations (11.3)–(11.7) constitute the recursive equation system of the Kalman filter prediction model. The classic Kalman filter load prediction model is aimed at predicting the electricity load in a stationary time series, but has poor performance for periodic fluctuations in load [13].

11.3 Seasonal Kalman Filter Electricity Load Forecasting Model

The classic Kalman filter load forecasting method is based on a stationary time series of daily load values at the same time, and the predicted duration is usually daily or weekly [14]. However, for non-stationary data, this algorithm is not suitable. In fact, in the medium to long term, the load has cyclical fluctuations, especially in the industry's electricity load, which fluctuates greatly with seasonal changes. For individual manufacturers in the same industry, the cycle of their electricity load is different. It is necessary to identify the period of historical electricity load data and add a period factor to the Kalman filter prediction [15].

The data collection of power load in the production process of power enterprises is carried out, and the power load curve in the industrial manufacturing process of power enterprises is collected and recorded, and the collected power load curve is self-adaptive cycle identification. Perform a fast Fourier transform on the power load identification data to obtain the spectrum sequence s_j. Take the subscript of the sequence with the largest spectrum as the period T, and T is equal to j corresponding to the maximum value of the spectrum sequence s_j. Then

$$s_j = \sum_{k=0}^{n-1} e^{-\frac{2\pi}{n} j^k} h_k \tag{11.8}$$

Among them, $h_k \in (h_0, h_2, h_3, \ldots, h_{n-1})$ is the true value sequence of power load, $k = 1, 2, 3, \ldots, n$, $j = 0, 1, 2, \ldots, n-1$.

In order to incorporate periodic factors and introduce a prediction model with two parameters, the state equation and measurement equation are as follows:

$$x_k = \Phi_{k|k-1} x_{k-1} + \Psi_{k|k-1} u_k + \omega_{k-1} \tag{11.9}$$

$$y_k = H x_k + v_k \tag{11.10}$$

where $u_k = x_{k-T}$, that is, the power load data at T before k is taken as the cycle factor, and Ψ_k is the correlation coefficient of the load at k and $k-T$. Since the cycle is T, Ψ_k is taken as the identity matrix I.

Let $A_{k|k-1} = [\Phi_{k|k-1}\ \Psi_{k|k-1}]$, $X_{k-1} = [x_{k-1}\ u_k]^T$, then there is

$$x_k = A_{k|k-1} X_{k-1} + \omega_{k-1} \tag{11.11}$$

Similarly, a recursive formula can be obtained by minimizing the minimum mean square error.

Prediction estimation equation:

$$\hat{x}_k = A_{k|k-1} X_{k-1} + K_k \left[y_k - H_k A_{k|k-1} X_{k-1} \right] \tag{11.12}$$

Prediction gain equation:

$$K_k = P_k H_k^T (H_k P_k H_k^T + R_k)^{-1} \tag{11.13}$$

Prediction covariance equation:

$$P_k = A_{k|k-1} P_{k-1} A_{k|k-1}^T + Q_{k-1} \tag{11.14}$$

In practice, it is difficult to accurately grasp the initial states x_0 and P_0. However, due to the continuous use of new information in the recursive process of Kalman filtering to correct the state, when the filtering time is sufficiently long, the influence of the initial state value x_0 on x_k will decay to zero, and the influence of the initial covariance P_0 on the filtering estimation covariance matrix P_k will also decay to nearly zero. Therefore, the initial conditions for filtering can be approximately determined. In this model, both the state value and the measured value are power load values, so the measurement matrix H_k is taken as the identity matrix I [9, 10].

The Kalman filtering model quantifies the common periodic phenomena in real production by adding periodic factors, thereby enhancing the predictive ability of classical Kalman filtering methods.

11.4 Case Study

This article selects the load data of representative electricity consumption units in the non-ferrous metal smelting and rolling processing industry, chemical raw material and chemical product manufacturing industry, and cement manufacturing industry in Guangxi region from 2014 to 2015 as examples. The classic Kalman filter and seasonal Kalman filter are used to estimate the electricity load day by day, and the advantages and disadvantages of the Kalman model with seasonal factors are analyzed, and the degree to which industry factors affect seasonal Kalman filtering.

Taking the non-ferrous metal smelting and rolling processing industry as an example, taking the data of the industry in 2014 as training data, and bringing it into Eq. (11.8), the spectral period of the industry's data in 2014 is 106 days, equivalent to about 3 and a half months.

H is initialized to [0.5, 0.5]; initialize process noise variance Q_1 to 4, measure noise variance R_1 to 4, and calculate the predicted value for 2015.

Figure 11.1 is a comparative diagram of using classical Kalman filtering and seasonal Kalman filtering to predict the smelting and rolling industries of non-ferrous metals, respectively.

The randomness of the classical Kalman filter in Fig. 11.1 is strong. For example, at points with a number of days equal to 52,100,110,190, there is a significant difference between the predicted results of the classical Kalman filter and the actual measured values. This indicates that the classical Kalman filter is susceptible to noise interference, resulting in significant errors in the predicted results. The Kalman filter incorporating seasonal factors avoids the situation of excessive prediction error, allowing the prediction error to converge to a smaller range. Figure 11.2 shows the cumulative probability distribution of classical Kalman filtering and seasonal Kalman filtering.

The image on the left of Fig. 11.2 represents the cumulative error distribution of classical Kalman filtering, while the image on the right represents the cumulative distribution of seasonal Kalman filtering. From the graph, it can be observed that

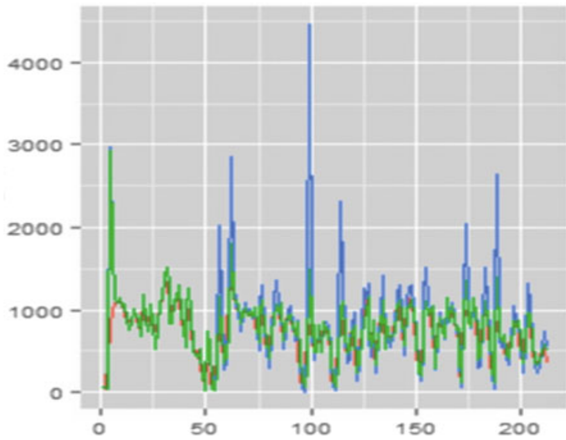


Fig. 11.1 Comparison of classical Kalman filter and seasonal Kalman filter predictions

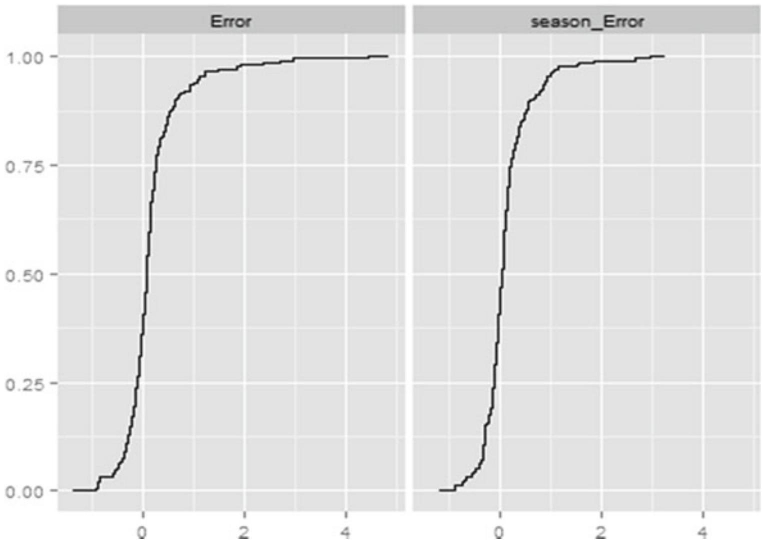


Fig. 11.2 Cumulative probability distribution of classical Kalman filter and seasonal Kalman filter

compared to classical Kalman filtering, seasonal Kalman filtering has a steeper cumulative distribution curve and converges at a lower upper bound. This fully demonstrates that seasonal Kalman filtering has stronger predictive ability than classical Kalman filtering.

The magnitude and patterns of electricity consumption vary among different industries. Therefore, it is necessary to consider whether the algorithm can maintain good predictive ability after incorporating industry factors. Using the non-ferrous metal smelting and rolling processing industry, chemical raw material and chemical

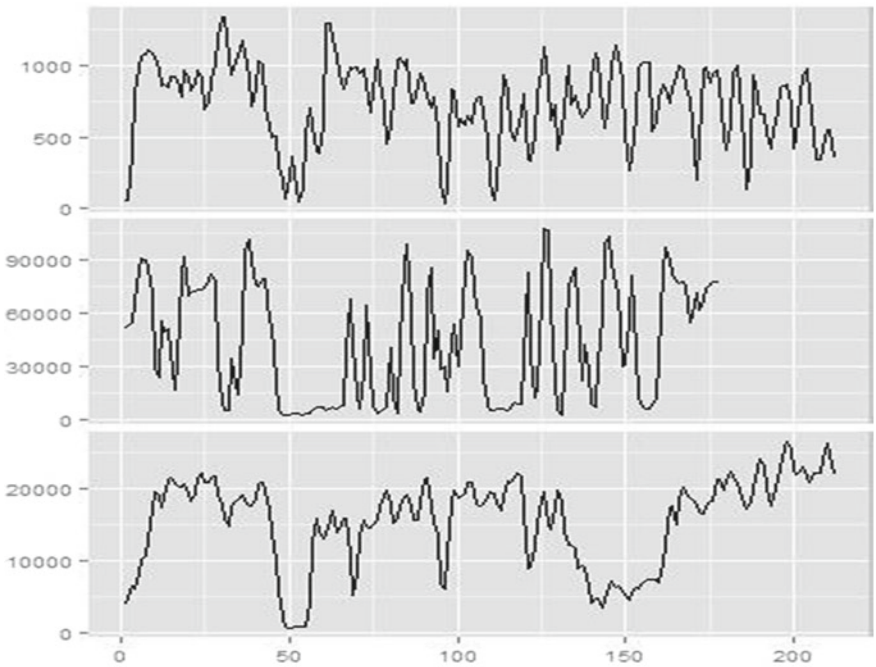


Fig. 11.3 Time series diagram of industry's electricity load to be predicted

product manufacturing industry, and cement manufacturing industry as control variables, analyze the generalization ability of seasonal Kalman filter models. Figure 11.3 is a time series comparison chart of the predicted electricity load for three industries.

From Fig. 11.3, it can be observed that the fluctuation of the power load curve shows weak periodicity, and the data fluctuation range is relatively large. The power load range of different industries is completely different, with a value of cement > non-ferrous metals > chemical engineering, and a variance of cement > non-ferrous metals > chemical engineering. Figure 11.4 shows the frequency distribution of generalization error of seasonal Kalman model for three industries.

From Fig. 11.4, it can be observed that the probability distribution of errors is basically similar, with a nearly bell-shaped distribution centered around 0. From the graph, it can be analyzed that the Kalman filter prediction model still has good predictive ability. However, for different industries, the generalization ability of seasonal Kalman filter prediction models varies significantly. The degree of concentration serves as the dividing line for distinguishing the generalization ability of the model, and its generalization ability is ranked as Nonferrous Metals > Chemical Industry > Cement Manufacturing. Research has shown that seasonal Kalman has good generalization ability for predicting electricity load, but industry factors have a certain influence on the stability of the model.

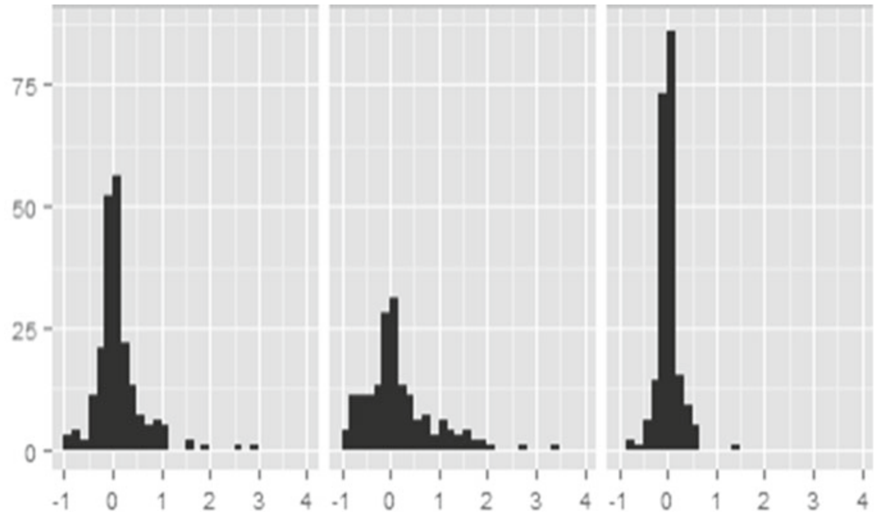


Fig. 11.4 Multiple industry prediction error frequency pairs

11.5 Conclusion

Regional electricity load forecasting is a key factor in determining the stable and economic operation of the power system. Based on the periodic characteristics of industry electricity load data, this article improves the common time-varying Kalman filter by incorporating seasonal factors. Taking the non-ferrous metal smelting and rolling processing industry, chemical raw material and chemical product manufacturing industry, and cement manufacturing industry as examples, the generalization ability of classic time-varying Kalman filtering and seasonal Kalman filtering was compared and analyzed. Research shows.

(1) In terms of model stability, seasonal Kalman filtering is superior to classical Kalman filtering models, and its upper bound of error is narrower than classical Kalman filtering.
(2) The error distribution of seasonal Kalman filtering is a bell-shaped curve centered around 0, and industry factors have a certain impact on the predictive ability of seasonal Kalman filtering.

The seasonal Kalman filtering method studied in this paper can use the recurrence method to continuously revise the state estimation according to the new information. When the time series is long enough, the influence of the state value of the initial state and the covariance matrix on the estimation will decay to zero. Therefore, the Kalman filter model can update the state information continuously and obtain more accurate estimates. This method can be used not only for short-term load forecasting, but also for ultra-short-term load forecasting.

References

1. Dan, W.: Overview of load forecasting methods for power systems. Technol. Inf. **34**, 079 (2014)
2. Xie, Z.: Analysis of power load characteristic indicators and their internal correlation. Changsha Univ. Technol. (2013)
3. Huizhong, W., Jia, Z., Ke, L.: Summary of research on short-term load forecasting methods in power systems. Electr. Autom. **37**(1), 1–3 (2015)
4. Jianjun, W., Zongyi, H.: The application of ARMA model in power demand forecasting in China. Econ. Math. **23**(1), 64–68 (2006)
5. Jingwen, Y., Hui, X., Boying, W.: Overview of power quality analysis methods based on Kalman filtering. Grid Technol. **34**(2), 97–103 (2010)
6. Huang, Z., Schneider, K., Nieplocha, J.: Feasibility studies of applying Kalman filter techniques to power system dynamic state estimation. In: 2007 International Power Engineering Conference (IPEC 2007), pp. 376–382. IEEE (2007)
7. Jingbo, M., Honggeng, Y.: The application of adaptive Kalman filtering in short-term load forecasting of power systems. Grid Technol. **29**(1), 75–79 (2005)
8. Mastorocostas, P.A., Theocharis, J.B., Bakirtzis, A.G.: Fuzzy modeling for short term load forecasting using the orthogonal least squares method. IEEE Trans. Power Syst. **14**(1), 29–36 (1999)
9. Charytoniuk, W., Chen, M.S., Van Olinda, P.: Nonparametric regression based short term load forecasting. IEEE Trans. Power Syst. **13**(3), 725–730 (1998)
10. Lu, J.H., Zhan, Y., Lu, J.A.: The non linear chaotic improved model of electric power system short term load forecasting. Proceed. CSEE **20**(12), 80–83 (2000)
11. Bakirtzis, A.G., Theocharis, J.B., Kiartzis, S.J., et al.: Short term load forecasting using fuzzy neural networks: IEEE trans on power. System **10**(3), 1518–1524 (1995)
12. Rahama, S., Bhatnagar, R.: An expert system based algorithm for short term load forecast. IEEE Trans. Power Syst. **3**(2), 392–399 (1998)
13. Minggan, L., Jianli, S., Pei, L.: Short term load forecasting of power systems based on Kalman filtering. Relay **32**(4), 9–11 (2004)
14. Ke, W., Lihua, C.: Short term load forecasting for holidays based on Kalman filtering. Electr. Technol. **1**, 3–6 (2014)
15. Min, Z., Haiyan, B., Jinping, L.C., Jianguang, D.: Research on short-term load forecasting method based on Kalman filtering. Grid Technol. **27**(10), 39–42

Chapter 12
Application of Lidar Technology in Power Engineering Surveying and Mapping

Guozhu Yang, Maojie Tian, Chuntian Ma, Junlei Li, and Bangbo Zhao

Abstract Airborne lidar technology provides high-precision and real-time data support for power grid planning, survey, design, construction and operation management with its digital, high-precision and 3D characteristics, and comprehensively improves the efficiency of power grid construction and management. This paper analyzes the composition and principle of airborne lidar technology as a fundamental breakthrough point, and discusses the application of airborne lidar technology in power engineering surveying and mapping. 30 inspection points were collected evenly in this survey area, and the results showed that the median error in the plane was 0.15 m, the maximum error was 0.7 m, the median error in elevation was 0.16 m and the maximum error was 0.12 m, which met the requirements of making 1:2000 Digital Elevation Model (DEM) and Digital Orthophoto Map (DOM) from point cloud data. After the route selection, the height of crossing the point line is analyzed by using the point cloud data, which can be used by the tower arrangement of line electrical specialty. Compared with the previous projects that need to go to the site to measure the line height, it saves time and improves the operation efficiency.

12.1 Introduction

In recent years, with the rapid development of the economy and society, the increasing demand for electricity has brought a severe test to the State Grid. How to build power lines quickly and efficiently, manage the completed lines scientifically, ensure the normal operation of the power grid and ensure the safe transmission of power are particularly important. Power grid engineering has the characteristics of a wide coverage area and a complex geographical environment, so it is necessary to use a measurement method that is not limited by geographical conditions and can efficiently and accurately obtain large-scale 3D spatial coordinate information of the

G. Yang (✉) · M. Tian · C. Ma · J. Li · B. Zhao
State Grid Electric Space Technology Company Limited, Beijing 102209, China
e-mail: gzyang3912@163.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_12

surface to realize fast and accurate survey and design of transmission lines [1]. Airborne lidar technology is one of the revolutionary achievements in photogrammetry and remote sensing in recent years, and it is the most advanced 3D aerial remote sensing technology at present. At present, airborne lidar measurement technology is mainly used to quickly obtain large-area 3D topographic data and quickly generate digital products such as DEM, especially the real topographic maps of forest-covered areas and mountainous areas, including strip-shaped target topographic maps. As capillaries of electric systems, transmission lines play an irreplaceable role in transmitting electric energy [2, 3]. With its unique characteristics of penetrating trees, bamboos and vegetation, 3D lidar can obtain information such as fine topography, landforms and the height of trees and bamboos in vegetation areas, which plays an irreplaceable role in other data acquisition methods and brings a brand-new technical trend to transmission line survey [4]. Thanks to the emergence and development of 3D lidar technology, the problems of spatial positioning and measurement accuracy can be well solved, and the problems of route selection and daily operation and maintenance can also be effectively solved. Airborne lidar technology provides high-precision and real-time data support for power grid planning, survey, design, construction and operation management with its digital, high-precision and 3D characteristics, and comprehensively improves the efficiency of power grid construction and management [5, 6]. Compared with the traditional photogrammetry technology, using airborne lidar to design transmission lines is not restricted by weather, light, vegetation and other conditions, and can quickly, cheaply and accurately obtain 3D topography, aerial digital images and other spatial geographic information data. This paper analyzes the composition and principle of airborne lidar technology as a fundamental breakthrough point, and discusses the application of airborne lidar technology in power engineering surveying and mapping.

12.2 Airborne Lidar Technology

12.2.1 Composition of Airborne Lidar System

There are many kinds and structures of lidar, and the form, volume and weight of the whole machine are also very different. All lidars are composed of three main parts: transmitting, receiving and signal processing. According to its structure, lidar can be divided into monostable and bistable. In the bistable system, the transmitting part and the receiving part are placed in different places to improve the spatial resolution. The laser power supply is used to drive the laser to generate a pulsed laser, and the function of the emission optical system is to collimate the laser output by the laser. The receiver is mainly composed of a receiving optical system, photodetector and receiver circuit. The function of the photodetector is to convert the laser light collected by the receiving optical system into a current signal, which is amplified, filtered, amplitude detected and frequency detected by the receiver circuit.

The airborne lidar measurement system takes the plane as the observation platform and the laser scanning ranging system as the sensor, and obtains the 3D spatial information of the earth's surface in real time [7]. The airborne lidar measurement system mainly consists of five parts: lidar ranging system, dynamic differential GPS, INS attitude measurement system, digital camera and computer, wherein the flight platform can be replaced by an airplane; the working principle of the laser scanner is basically similar to that of ordinary microwave radar, but the wavelength of the laser scanner is smaller than that of ordinary microwave radar. The point cloud data obtained by airborne lidar systems can penetrate the leaf canopy of vegetation. Compared with the traditional photogrammetry technology, data acquisition is faster and the absolute accuracy is higher.

12.2.2 Working Principle of Airborne Lidar System

The interaction between the laser beam emitted by lidar and the detected object to produce an optical echo is the basis of lidar detection [8, 9]. In fact, the working process of an airborne lidar system is a process of repeatedly transmitting and receiving laser points. This working process expresses the results of sampling the surface of the target object by the system through a point cloud diagram composed of spatial points with a certain resolution (see Fig. 12.1).

The laser ranging unit includes a laser transmitter and a receiver. Laser scanning is an active working mode. The laser transmitter generates a laser beam, and the

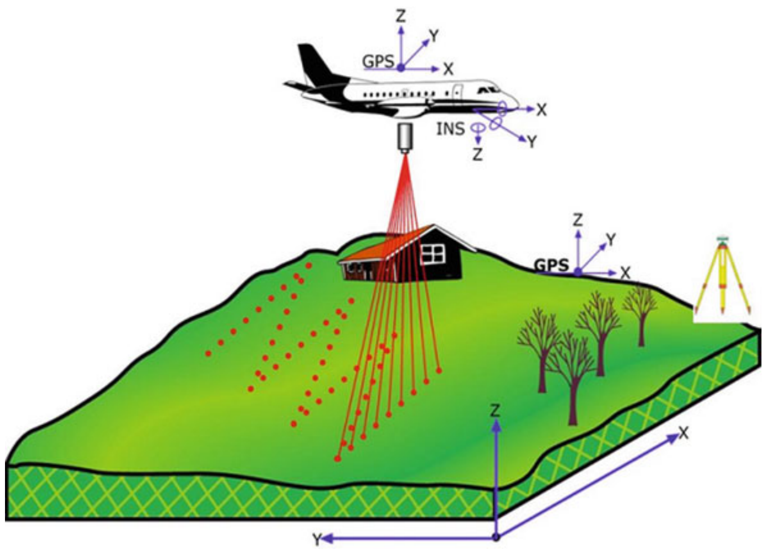


Fig. 12.1 Working principle of airborne lidar system

scanning device controls the direction of the laser beam emitted. After the receiver receives the reflected laser beam, it is recorded by the recording unit. At the same time, the Inertial Navigation System (INS) attitude measurement system can accurately measure the instantaneous space attitude parameters of aircraft: roll angle, tilt angle and heading angle.

After the airborne radar laser system receives the echo, the detected optical signal can be recognized by the computer system and processed accordingly. Finally, various data about electric power engineering surveying and mapping are generated, stored in the surveying and mapping system for operation and use, and can also provide corresponding data reference for future daily management and operation.

12.3 Application of Lidar Technology in Power Engineering Surveying and Mapping

12.3.1 Data Processing

With the rapid development of science and technology and the pursuit of airborne lidar technology, airborne lidar technology has been widely used in computer games, mold manufacturing, cultural relics excavation, tunnel construction and other fields; especially in the last decade, with the continuous innovation and development of various professional software and computer hardware environment, its application fields have become more extensive, and it has gradually entered every field of people's daily life from the laboratory [10]. After processing and analyzing the data by using the professional software of a 3D Geographic Information System (GIS), the 3D real scene along the field line can be moved indoors, and the computer display of a 2D map, 3D map and plane cross-section map can be realized. In this way, designers can fully consider the influencing factors such as landform, social environment and ecological environment, and then combine their own experience and technology to optimize the design of the line and output the results that meet the standards of the power industry.

Make a flight plan, including band division, heading overlap, lateral overlap, laser wavelength, scanning angle, altitude, camera shooting interval, etc. The laser scanning measurement data, Global Positioning System (GPS) measurement data, flight attitude data and aerial photography data are jointly processed; then these data are classified, denoised and filtered to generate orthographic images and 3D models [11, 12].

Data preprocessing refers to the process of positioning, orientation, calibration and coordinate transformation of the original laser data, and it is also necessary to determine the external orientation elements of the digital image. The positioning and orientation part uses GPS data, IMU data and various parameters provided by the system. In order to correct the errors of the three attitude angles of the lidar, it is calibrated according to specific ground objects and airborne lidar data samples.

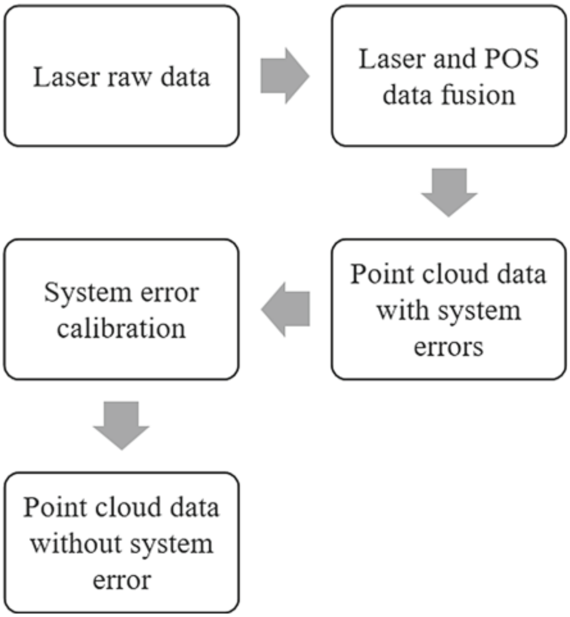


Fig. 12.2 Flowchart of point cloud preprocessing

The preprocessing process is that the original point cloud data mainly contains the angle and time information of laser emission, and the data with distance and angle is processed by software, and the corrected data is fused with the integrated navigation data to generate point cloud data with errors. The operation flow is shown in Fig. 12.2:

In order to correct the roll value, the height difference (d_1, d_2) and the distance between two points on two opposite flight routes are measured. Calculate the correction value Δr of the roll angle according to the following formula:

$$\Delta r = \frac{d}{\text{swath width}} \tag{12.1}$$

By using the calibration parameters, the cross sections of the point cloud data of two opposite routes in the scanning direction are finally adjusted to overlap each other.

Pitch angle of system integration is mainly reflected in the deviation of laser point cloud in the vertical direction of laser scanning line (flight direction of aircraft). The comparison method of pitch angle is that we use two flight belts to fly back and forth, which are perpendicular to the roof line of the house and fly over the spire house at the same time. Pitch angle is calculated by trigonometric geometry:

$$\Delta P = \frac{D}{2H} \tag{12.2}$$

where D represents the distance between planes flying back and forth to obtain the same position information. H stands for flying altitude.

The error of heading direction will change the position of the center of the scanned object and deform the object, so it is relatively difficult to check the heading angle error. The length of the calibration baseline (the distance from the nadir of the route to the center of the target object) is short, and the object offset is small, which makes the calibration error large.

According to the error influence law of `heading`, the formula for calculating heading error can be obtained as

$$\Delta p = \arctan \frac{\sqrt{D}}{S} \tag{12.3}$$

where S is the plane distance from the target center to the intersection of the routes.

12.3.2 Selection of Key Parameters of the System at Different Flight Altitudes

With the support of airborne LIDAR technology, power enterprises can quickly scan and recover the accurate 3D spatial positions of terrain and objects in line corridors by using laser lidar measurement technology, and can record the information of environmental variables during operation in real time. The geographical environment and other conditions around lines and equipment are displayed perfectly and intuitively by means of stereoscopic simulated images, thus providing clearer and more accurate information reference for the management and decision-making of power enterprises.

In the application of airborne lidar to ground mapping, according to different mission objectives, the flying height of the aircraft is usually between 1000 and 6000 m. At the flying height of 3000 m, the radar system should work in the near-saturation imaging mode, in which the root mean square error of ranging first decreases and then increases with the fringe width, and there is an optimal fringe width to minimize the ranging error.

In the airborne surveying and mapping system, the number of sampling units N in each sampling channel is set to 1200. In order to ensure that the distance gate width is enough to cover the depth of field of the measured object, the time slot width $t_{bin} = 1.26\,\text{ns}$ of the system is the corresponding distance gate width Gate $= 172\,\text{m}$ at this time.

In the mapping experiment of fixed targets, the position of fringe image in the sampling channel can usually be changed by adjusting the delay between detector and laser. That is to say, if the depth of field of the measured object is not particularly large, we can ignore the influence of fringe bending on ranging accuracy. However, in the application of airborne mapping to the ground, we usually can't know the

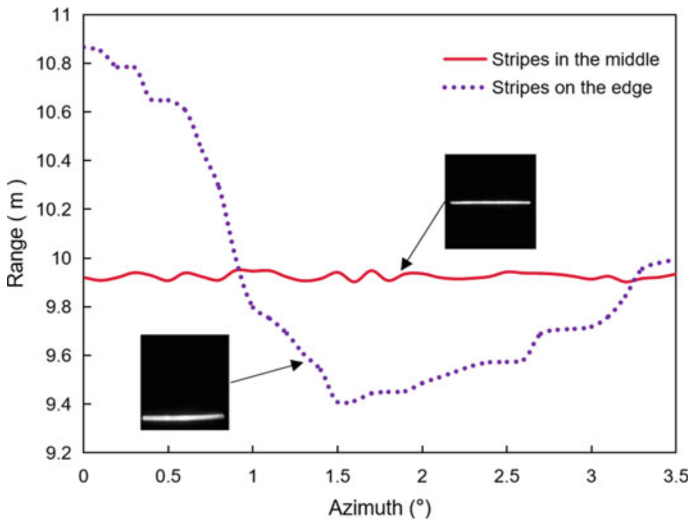


Fig. 12.3 Bending of fringe images at different positions

elevation information of the measured object in advance, and the fringe image may appear at any position of the sampling channel, so the distance deviation caused by fringe bending must be considered at this time.

From Fig. 12.3, we can see that when the fringe image is at the edge of the sampling channel, the distance extraction result is obviously curved; when the fringe image is in the center of the sampling channel, the distance extraction results are basically the same.

According to the calculation result of the signal centroid, the distance value corresponding to the centroid position can be queried in the calibration matrix, so that more accurate distance extraction results can be obtained. By changing the time delay between the detector and the laser, the fringe images move in the whole time-resolved channel, and the fringe images at different positions are stored at certain intervals, interpolating the distance values at all centroid coordinates to obtain the distance calibration matrix of all charge coupled device (CCD) pixels.

12.3.3 Accuracy Analysis

Strong wind, snow and ice weather may cause conductor dancing, change the distance between conductors and split conductors, and easily cause short-circuit discharge. Lidar can accurately measure the distance between each split sub-conductor under the condition that the line is live and not in contact with the line, and provide guidance for making spacer bars.

Based on the 3D scene processed by lidar measurement technology, the real situation in the field is displayed on the computer, and the whole line can be understood from the global to the microscopic, as shown in Fig. 12.4.

See Fig. 12.5 for the accuracy requirements of lidar data processing:

The accuracy of classified point clouds is checked by stereo image pairs, that is, the obvious corner points of ground objects are collected on stereo image pairs, and the same corner points are collected on the point cloud model, and then the plane and elevation are compared to check their accuracy. 30 inspection points were collected

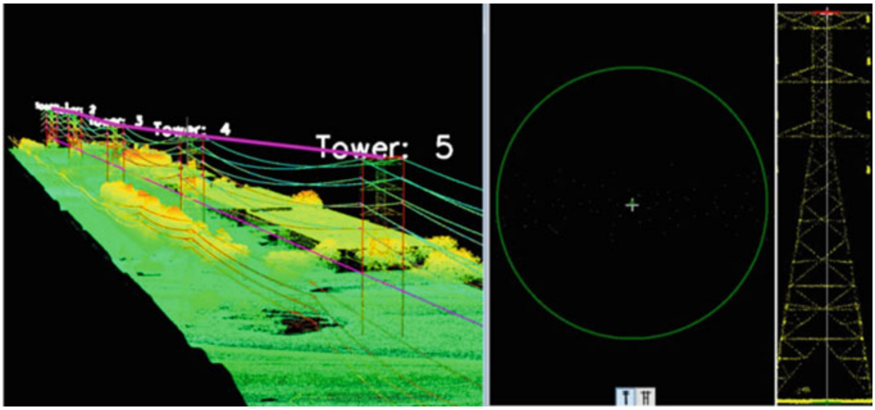


Fig. 12.4 Line 3D laser point cloud obtained by lidar measurement technology

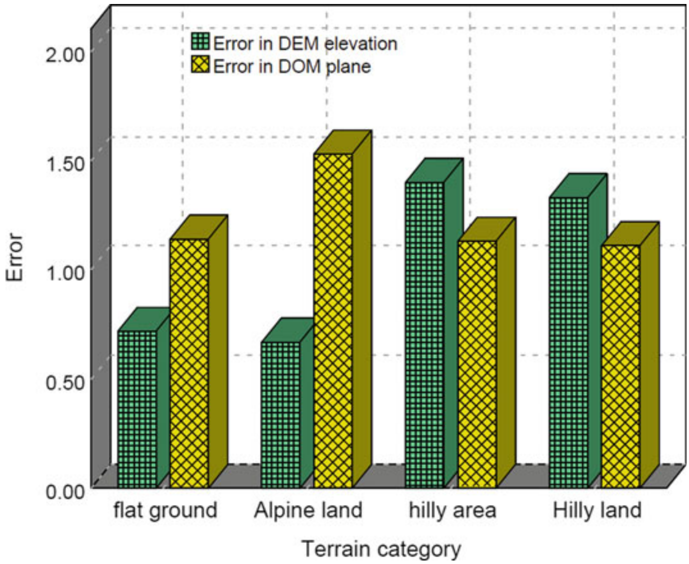


Fig. 12.5 Required precision

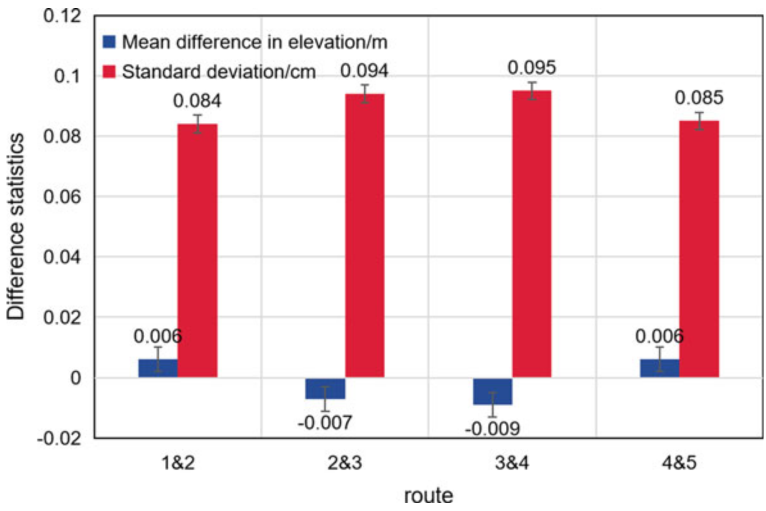


Fig. 12.6 Statistics of elevation difference in overlapping areas

evenly in this survey area, and the results showed that the median error in the plane was 0.15 m, the maximum error was 0.7 m, the median error in elevation was 0.16 m and the maximum error was 0.12 m, which met the requirements of making 1:2000 DEM and DOM from point cloud data.

After band adjustment, the internal accuracy (relative consistency) of point cloud data must be evaluated. It is usually realized by checking the consistency of lidar points in overlapping zones. In addition to zonal mapping, airborne lidar data usually have overlapping areas between different routes. Figure 12.6 shows the statistical results between five parallel routes in a small area.

The first line indicates that the difference of most points in the overlapping area of route 2 and route 3 is within the range of -0.007 ± 0.094. For areas with large slopes, this index of elevation difference also reflects the accuracy of the plane, because the error of the plane will bring the error of elevation. It is difficult to distinguish the plane difference of laser point cloud data because of its discontinuity. The plane difference can be realized by comparing the plane differences of the same name points of the feature objects directly extracted from the original data.

The better point cloud data of the laser reflection surface of the conductor can fit the outgoing line type. For many years, the reflection surface of the wire has been poor, so it is difficult to return to the pulsed laser and obtain continuous point clouds. For power lines of 10 kV and below, the conductor section of communication lines on the ground is small, so it is impossible to return a pulsed laser, and it cannot be identified in the aerial survey industry. The identifiable power line elevations are shown in Table 12.1.

After calculation, the mean error between the fitting elevation of power line point cloud and the measured elevation in this project is 0.17 m, which can meet the requirements of power line selection accuracy. After the route selection, the height

Table 12.1 Line height measurement

Circuit	Approximate difference/m
1	−0.042
2	−0.157
3	−0.141
4	−0.144
5	−0.161

of crossing the point line is analyzed by using the point cloud data, which can be used by the tower arrangement of line electrical specialty. Compared with the previous projects that need to go to the site to measure the line height, it saves time and improves the operation efficiency.

12.4 Conclusion

At present, airborne lidar measurement technology is mainly used to quickly obtain large-area 3D topographic data and quickly generate digital products such as DEM, especially the real topographic maps of forest-covered areas and mountainous areas, including strip-shaped target topographic maps. As capillaries of electric systems, transmission lines play an irreplaceable role in transmitting electric energy. Using an airborne lidar system to design power lines, only a few ground control points and field mapping work are needed to obtain high-precision 3D coordinates of ground points. Using airborne lidar technology to obtain terrain data has high accuracy and high speed, which can greatly reduce or even avoid the field survey work of surveyors and greatly improve work efficiency. Laser can penetrate the canopy of vegetation to reach the surface, and at the same time obtain information on the ground and vegetation, which is more abundant than that obtained by traditional aerial photogrammetry, and can also reduce the layout of ground control points. Based on the above characteristics, airborne lidar technology will play an increasingly important role in the construction and maintenance of the national power grid.

References

1. Meng, Q., Guo, H., Zhao, X., et al.: Loop-closure detection with a multiresolution point cloud histogram mode in lidar odometry and mapping for intelligent vehicles. IEEE/ASME Trans. Mechatron. **26**(3), 1307–1317 (2021)
2. Thatcher, C., Lim, S., Palaseanu-Lovejoy, M., et al.: Lidar-based mapping of flood control levees in South Louisiana. Int. J. Remote Sens. **37**(24), 5708–5725 (2016)
3. Douglas, T.A., et al.: Degrading permafrost mapped with electrical resistivity tomography, airborne imagery and LiDAR, and seasonal thaw measurements ERT measurements of discontinuous permafrost. Geophysics **81**(1), 71–85 (2015)

4. Gorsevski, P.V., Brown, M.K., Panter, K., et al.: Landslide detection and susceptibility mapping using LiDAR and an artificial neural network approach: a case study in the Cuyahoga Valley National Park. Ohio. Landslides **13**(3), 467–484 (2016)
5. Ouyang, Z.H., Cui, S.A., Zhang, P.F., et al.: Iterative Closest Point (ICP) performance comparison using different types of lidar for indoor localization and mapping. Lasers Eng. **46**, 47 (2020)
6. Famili, A., Sarasua, W.A., Shams, A., et al.: Application of mobile terrestrial LiDAR scanning systems for identification of potential pavement rutting locations. Transp. Res. Rec. **2675**(9), 1063–1075 (2020)
7. Brown, R., Hartzell, P., Glennie, C.: Evaluation of SPL100 single photon lidar data. Remote Sens. **12**(4), 722 (2020)
8. Du, S., Li, Y., Li, X., et al.: LiDAR odometry and mapping based on semantic information for outdoor environment. Remote Sens. **13**(15), 2864 (2021)
9. Zhang, R., Yang, B., Xiao, W., et al.: Automatic extraction of high-voltage power transmission objects from UAV lidar point clouds. Remote Sens. **11**(22), 2600 (2019)
10. Gentil, C.L., Vidal-Calleja, T., Huang, S.: IN2LAAMA: inertial lidar localization autocalibration and mapping. IEEE Trans. Rob. **99**, 1–16 (2020)
11. Li, T.J., Zhang, J.G., Zhao, Z.: Characteristics of light and small laser radar system for electric power inspection and some matters needing attention. Bull. Surv. Map. **5**, 4 (2018)
12. Duan, M.Y.: Research on 3D reconstruction method of airborne lidar point cloud power line. J. Surv. Map. **12**, 1 (2016)

Chapter 13
Research on Smart Car Safety Based on YOLOv3 Target Detection Algorithm

Yanxia Liu

Abstract In recent years, smart cars have achieved rapid development. Environmental awareness technology is the basis of behavior prediction, decision-making and control of moving targets around smart cars, which largely determines the intelligence level of cars. In the complicated and changeable road traffic environment, the negligence and inattention of vehicle drivers may cause irreversible losses. Real-time classification and location of surrounding objects is a key link to ensure road traffic safety during vehicle driving. Among them, the lane line detection and lane departure warning system based on vision is a hot topic in the research of smart cars at present, and it is also one of the key technologies to realize the safe driving assistance function of automobiles. Therefore, this article combines the improved YOLOv3 algorithm with target object detection and distance measurement. The program outputs 6 categories of target object classification and location information and 4 categories of longitudinal distance between the target object and the camera. The video detection speed reaches 29.8 frames per second, meeting the real-time requirements. This method can provide a reference for vehicle-assisted driving in natural road traffic scenarios.

13.1 Introduction

With the rapid development of China's economy, the living standards of Chinese residents have gradually improved. Compared with trains and airplanes, automobiles, as a more flexible and convenient means of transportation, have become an important part of social life and production, and their number has increased rapidly [1]. With the huge number of cars and the accelerating pace of life in modern society, it is not uncommon that drivers' behaviors such as running red lights, speeding and overloading caused by fatigue driving, and information such as vehicles, pedestrians

Y. Liu (✉)
Nanchang Institute of Technology, Nanchang 330044 Jiangxi,, China
e-mail: 269106172@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_13

and traffic signs on both sides or above the road in front of them are not found in time [2].

In China, according to statistical data from the Traffic Management Bureau of the Ministry of Public Security, there were 3,906,164 road traffic accidents nationwide in 2010, a year-on-year increase of 35.9%. 65,223 people were killed and 251,103 injured, resulting in a direct property loss of 930 million yuan [3]. According to the analysis of accidents, more than 90% of traffic accidents are caused by drivers' untimely response and operational errors when encountering dangerous situations. These factors objectively reflect the following aspects: (1) drivers' distraction when driving; (2) driver's long-term fatigue driving; (3) drivers fail to handle the vehicle in a timely and correct manner at dangerous times due to overspeed and overload; (4) drivers misjudged the road conditions and mismanipulate the vehicle [4]. However, due to the variety of targets to be detected in driving scenes and the complexity of the environmental background, target detection algorithms for autonomous driving scenes still have great challenges, and some efforts still need to be made in terms of accuracy and real-time performance. Therefore, how to improve vehicle safety performance and the accuracy and speed of traffic target category detection is the key to achieving autonomous or unmanned driving [5].

In recent years, with the progress of on-board chip technology, target recognition, sensor technology, artificial intelligence, computer communication technology and decision algorithms of self-driving cars, self-driving cars have achieved great development. Self-driving cars integrate environmental awareness, dynamic decision planning, intelligent control and execution, so that they can collect surrounding information in real time through sensors, make analysis and decisions according to surrounding road conditions, plan driving routes, control the acceleration and deceleration of cars, and turn left and right [6]. In order to improve the increasingly serious traffic problems, alleviate the increasingly congested traffic environment and better configure traffic resources, the system of intelligent transportation systems has also been established. With the advent of the 5G era, intelligent transportation will play an increasingly important role in the future of intelligent life. It is a close connection between advanced science and technology and transportation manufacturing, transportation, and services, combining users, vehicles, and roads, A data-driven, accurate, safe, energy-saving, and comprehensive modern transportation system has been formed [7]. However, if only a fixed number of sliding window templates are applied, unsatisfactory areas may result. In order to identify different objects, we need to extract visual features. Feature extraction is an important step in the target detection process, and the quality of feature extraction directly affects the effectiveness of target detection [8].

This article mainly studies object detection algorithms from two aspects: theoretical innovation and practical application value. Finally, a target detection system is implemented based on the improved YOLOv3 algorithm. In order to increase the robustness of the system, the model is trained by data set expansion. Users can upload videos or pictures, complete the labeling of target detection, and return the results to users. Tests and experiments show that the system is robust.

13.2 Overview of Research on Target Detection of Smart Car Safety

13.2.1 Vision-Based Target Detection

Visual sensor is one of the most important sensors in intelligent vehicle systems, with advantages such as low cost, high frame rate, rich information, and long observation distance. However, it is susceptible to environmental impact and lacks in-depth information. In the field of intelligent driving assistance, the advantage of cameras lies in their ability to effectively detect and classify obstacles in front of them [9]. The detection algorithm based on deep learning needs to be trained for large-scale data sets, so it has a certain generalization ability. The more similar the characteristics of the detected images and the targets in the training set, the higher the accuracy of the model. Therefore, in order to be able to achieve detection targets in different weather conditions such as sunny days, cloudy days, rainy days, and different scenes such as urban streets, gas stations, parking lots, and highways, the image data collected from a real driving platform is required for deep neural network training [10].

First of all, YOLOv3 uses Darknet-53 as the basic network structure, which was originally designed to detect and identify 80 kinds of general targets, but this paper only needs to detect 4 kinds of automobile targets, so there are many redundant structures in the basic network that can be optimized. However, because only four types of vehicles need to be detected in this paper, the depth and width of the Darknet-53 basic network are somewhat redundant. The whole data set production process is shown in Fig. 13.1.

The whole process is mainly divided into five parts, which are determining the acquisition object, shooting the video on the spot, framing and cleaning the video, and manually expanding the data set samples and labeling.

The purpose of vehicle and pedestrian target detection algorithms is to detect the location and type confidence information of vehicles and pedestrians in images. They are generally divided into single-stage and dual-stage-based algorithms. Compared to a single-stage detector, a dual-stage detector separates location and classification, requiring the generation of a region suggestion box for location, which consumes a lot of time. The input layer of a conventional neural network is the input quantity. Each node in the hidden layer has a weight and deviation, which are calculated and transferred to the next node. Finally, the calculation result is output through the output layer. In conventional neural networks, every node between layers is fully connected, and different nodes in the same layer are completely independent. For ramp signals and step signals, the input and output signals can be guaranteed to be the same.

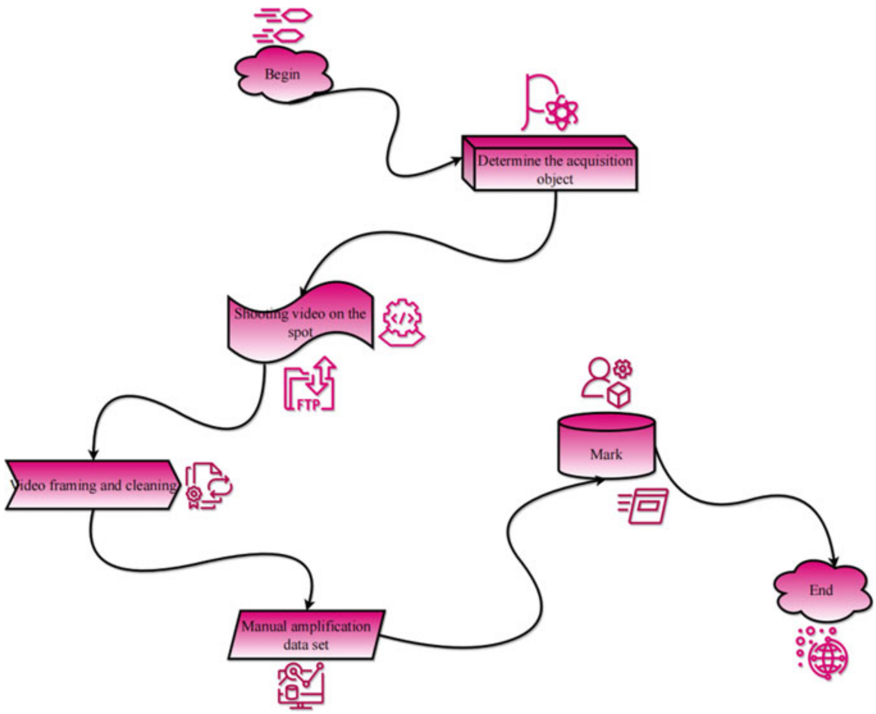


Fig. 13.1 Flowchart of data set making

13.2.2 Lane Line Detection and Feature Extraction

The image captured by the camera is a two-dimensional image, while the actual scene is a three-dimensional space. Therefore, it is necessary to calibrate the camera and determine the geometric relationship between the three-dimensional scene and the two-dimensional image based on the imaging model of the camera. An ideal camera model is a pinhole model composed of the optical center, imaging surface, and optical axis of the camera. Objects in space are projected onto the imaging plane through the optical axis. The actual camera imaging model needs to consider projection distortion. In fact, when the camera is shooting images, the ideal installation position of the camera is perpendicular to the shooting plane, which can ensure that the image maintains its original scale without distortion. Compared to traditional target detection algorithms, deep learning target detection algorithms not only have a breakthrough in detection accuracy, but also have improved detection speed. However, in autonomous driving scenarios, target detection based on deep learning has significant shortcomings in terms of real-time detection and detection accuracy. Even if a large number of feature factors and feature weights can be extracted, the description of the obstacles in front is still limited, resulting in the robustness of the detection model not being high. Moreover, in the face of complex scenes, high-order images are even more stretched. Therefore, the difficulty of traditional target detection is limited, not only the detection accuracy can not be greatly improved, but also the detection speed can not achieve real-time performance.

Target detection algorithms have gradually developed from traditional algorithms to deep learning algorithms. Traditional table detection algorithms have low detection accuracy, cannot meet real-time requirements for detection speed, and are less robust to complex scenarios. This makes it more difficult to apply such target detection in the engineering field. In practical projects, it is difficult to find enough image data to complete training tasks. In order to ensure the completion of the project, in addition to searching for more data, we also need to conduct data augmentation operations on the collected data through relevant means. This is to make up for the shortcomings of insufficient data collected by ourselves, and to simulate the characteristics presented by targets in different environments. Therefore, the development of deep learning technology simplifies the setting of parameters in traditional target detection, reduces the use of experience by people, and converts human experience into data form. At the same time, deep learning exhibits learning ability for a large number of sample sets through a non-linear network structure relationship, which can achieve good results in automatic driving target detection.

13.3 Research on Intelligent Vehicle Safety Based on YOLOv3 Target Detection Algorithm

13.3.1 YOLOv3 Target Detection Algorithm Model

Because in practical engineering applications, technicians pay great attention to the reasoning performance of the model; the less computational power the model needs and the higher its accuracy, the more it can play a role in engineering. The smaller the parameters of the model, the smaller the storage space, and the smaller the running space required for reasoning. Compared with the original model, it can also provide a higher ability to process simultaneous input data, increase the number of pictures for simultaneous reasoning, and can be used for multi-channel reasoning. The methods of model compression are generally divided into model pruning, model quantization, and neural architecture search. In practical application, model pruning and quantization are post-processing methods based on trained models. Neural architecture search is a model preprocessing method, which requires high time consumption.

In order to be able to detect fine-grained features, it further uses three feature maps with different scales for object detection based on YOLOv2's passthrough structure; YOLOv3 still uses the YOLOv2 method to obtain anchors using K-means clustering, increasing the number to 9, and assigning 3 different anchors to each scale feature graph, which greatly improves the ability of the network to detect small-size objects. In addition, it also changes the softmax function of the prediction layer to a logistic

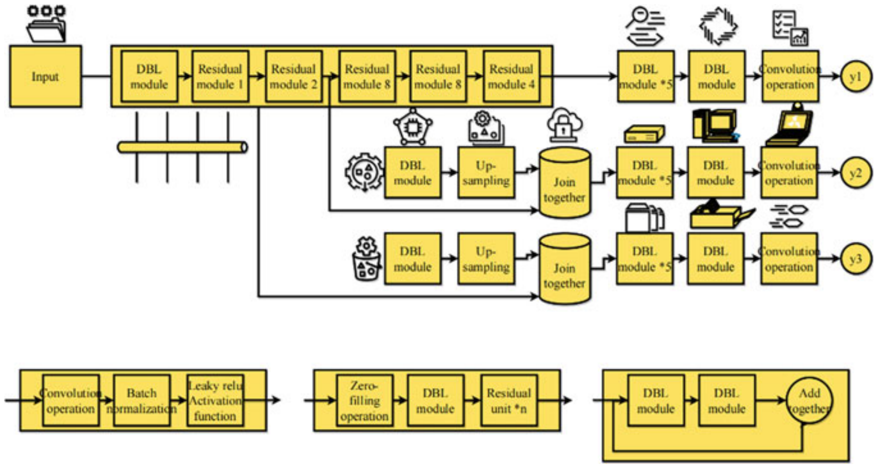


Fig. 13.2 YOLOv3 network structure

function, which can support the output of multiple label objects. The overall network structure of YOLOv3 is shown in Fig. 13.2.

In recent years, the vision-based object detection method using a convolutional neural network (CNN) has been very successful, such as YOLOv3 in this paper. Darknet-mini changed the feature map from 416*416*3 to 26*26*128 after four times of down-sampling, and obtained abundant feature information. In order to extract more local features, the feature map filter was reduced to 64 through three DBL convolution modules before input. Definition $size_{pool} \times size_{pool}$ indicates the size of the sliding window, and $size_{fmap} \times size_{fmap}$ indicates the size of the feature map, which can be obtained as follows:

$$size_{pool} = \left[size_{fmap} / n_i \right] \tag{13.1}$$

Due to the absence of a fully connected layer in YOLOv3, it is not necessary to adjust the size of the feature map to a fixed-size feature vector. In order to obtain more feature information and improve accuracy, three feature maps with a size of 26 * 26 * 64 are concatenated with the input feature maps of the SPP module to finally obtain a 26 * 26 * 256 large channel feature map. These feature maps extract and merge features from multi-scale local regions as output, helping to eliminate the impact of differences in extracted feature information caused by different target scales, enabling the network to learn target features more comprehensively.

The classification loss function of YOLOv3 is a cross-entropy loss function, and the cross-entropy loss function does not treat all samples differently, that is to say, the contribution of each sample to the error loss is only related to its own error loss, which will cause many car types to contribute more to the error loss, thus leading the gradient adjustment process, while a small number of other types of vehicle samples contribute less to the error, making it difficult for the network to fully learn the effective characteristics of other types of vehicle samples during training. Although $\lambda = 0.5$ basically balances the contribution of both background and vehicle samples in the balanced cross-entropy (CE) loss function, it does not distinguish between difficult samples between vehicles, as shown in Eq. (13.2). Therefore, we use the focus loss function to reduce the training weight of easy car samples, and we can dig deeper into the characteristics of other types of vehicles:

$$CE(p_t) = -\lambda \log(p_t) \tag{13.2}$$

The focus loss function (FL) mainly introduces an adjustment factor $(1 - p_t)^{\gamma}$ based on the cross-entropy loss, with adjustable parameters $\gamma \geq 0$. The formula for the focus loss function is shown in Eq. (13.3):

$$FL(p_t) = -(10p_t)^{\gamma} \log(p_t) \tag{13.3}$$

When $\lambda = 0$, the focus loss function is equivalent to the cross-entropy loss function. If it continues to increase, the loss difference between insoluble and insoluble samples becomes larger and larger. In other words, the loss of large samples is decreasing, and the loss of small samples is increasing.

13.3.2 Analysis of Experimental Results

Because the training data set of the deep learning model has a large training sample, and the picture of the training set has a high pixel value of 1920×1080, which occupies a large amount of memory, the ordinary computer platform is not competent for the training task of YOLOv3, so it needs to be completed on the platform with GPU graphics card to accelerate the calculation. In order to comprehensively measure the detection performance of YOLOv3 and improved YOLOv3 models, this experiment combines the cross-union ratio and loss value in the training process, and evaluates the applicability of the model to the detection task through the indicators of recall rate, accuracy rate, average accuracy, and mean average accuracy. The commonly used evaluation criteria in target detection are average accuracy mean mAP. The calculation of mAP involves the calculation of Precision, Recall, and AP. The definitions of precision and recall involve confusion matrix, which is shown in Table 13.1.

When using migration training for YOLOv3-pro, the loss value decreases rapidly. When not using migration training, the decline is slower. This is because when using

Table 13.1 Evaluation model prediction confusion matrix

Positive: It's a car		Model prediction	
Negative: Not a car		Positive	Negative
Real label	Positive	True positive	False negative
	Negative	False positive	True negative

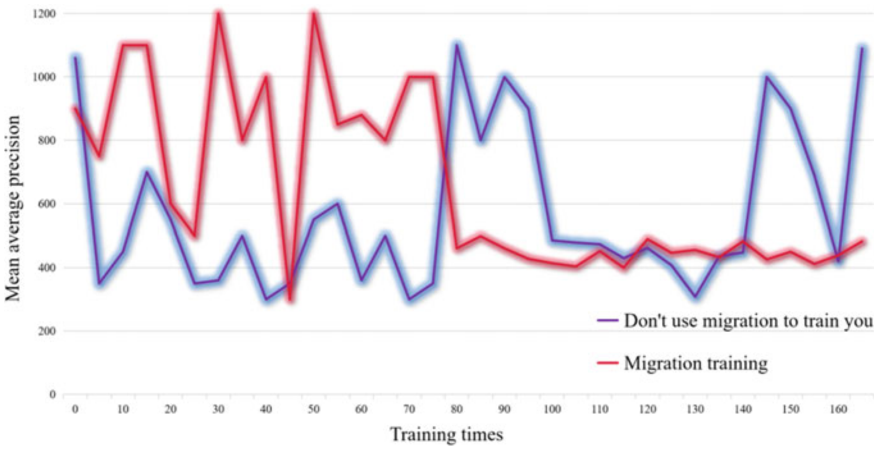


Fig. 13.3 Effect of training times on mAP

pre-training weights for migration training, the underlying weight parameters can be used as initial values for training, and these initial values fuse some target features. Therefore, the training method based on transfer learning can find and locate the target faster, and the loss value decreases faster. When migration training is not used, the model needs to learn target features from 0, which slows down the speed of identifying and locating the target, resulting in a slower decrease in the loss value. The changes of mAP in the two training methods under different training times are shown in Fig. 13.3.

As can be seen from the figure, the value of the map of the two training methods is basically around 0 when the training times are very small. This is because there are few pictures sent to the network in the early stage, and the model does not fully learn the characteristics of the target, resulting in low detection accuracy. With the increase in training times, more and more pictures are sent to the network, and the model gradually grasps the characteristics of the target, so the map values of the two models slowly rise. The verification data set is an MSCOCO data set. Training process of Pytorch platform: loading model, initializing data, defining optimizer, training, and forecasting. Explain that the L_{HOU} loss function can solve the shortcomings of the L_{HOU} loss function to some extent. When training the model, all the values of the loss function are saved in the log. In order to compare the changes of several loss functions conveniently, the comparison of several loss functions is shown in Fig. 13.4.

From the image, we can see that several loss functions tend to stabilize after 400,000 iterations, and the L_{HOU} loss function converges faster during the initial iteration.

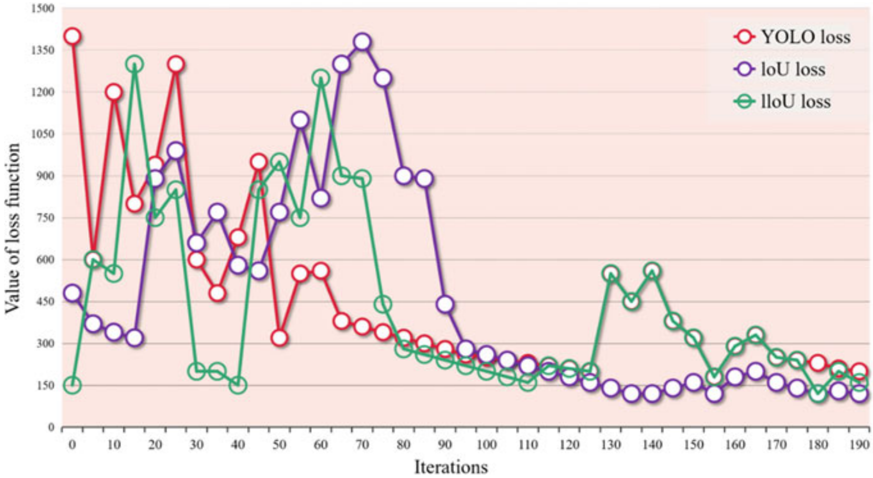


Fig. 13.4 Comparison of several loss functions

13.4 Conclusion

Object recognition and positioning in road traffic environments play an important role in vehicle dynamic path planning and avoiding traffic accidents. With the development of cities and the continuous improvement of people's safety awareness, the research on assisted driving technologies such as road object detection has become the focus of attention of scholars and research institutions at home and abroad. In the research of vehicle safety-assisted driving technology, lane line detection and lane departure warning system have become two important research directions. Lane line detection and feature parameter extraction are key technologies to achieve lane departure warning systems. Based on the background of road traffic, by studying the characteristics of traffic targets, this paper determines that the objects of this algorithm are four kinds of dynamic targets: people, cars, bicycles, and battery cars. In order to meet the real-time and reliability requirements of autonomous driving, based on the YOLOv3 target detection algorithm, this paper improves the YOLOv3 model by optimizing the data set, adjusting the structural model, optimizing the training parameters, and using migration training, so as to improve the recognition effect in autonomous driving scenes.

Acknowledgements Project Name: Science and Technology Research Project of Jiangxi Provincial Department of Education (Fund number: GJJ2202708).

References

1. Xu, H., Guo, M., Nedjah, N., Zhang, J.D., Li, P.: Vehicle and pedestrian detection algorithm based on lightweight YOLOv3-Promote and Semi-Precision acceleration. IEEE Trans. Intell. Transp. Syst. **23**(10), 19760–19771 (2022)
2. Sun, D.H., Lei, Z.J., Shao, H., Wang, Y.Z.: Ship Target Detection Algorithm Based on Improved YOLOv3 for Maritime Image. J. Adv. Transp. 2021, Article ID 9440212 (2021)
3. Chen, B., Pei, X., Chen, Z.: Research on target detection based on distributed track fusion for intelligent vehicles. Sens.S (Basel, Switz.) **20**(1), 56 (2020)
4. Liu, P., Song, C.L., Li, J.M., Yang, Simon X., Chen, X.Y., Liu, C.F., Fu, Q.: Detection of transmission line against external force damage based on improved YOLOv3. Int. J. Robot. & Autom. **35**(6), 460–468 (2020)
5. Wang, H., Yu, Y., Cai, Y., Chen, X.B., Chen, L., Liu, Q.C.: A Comparative study of State-of-the-Art deep learning algorithms for vehicle detection. IEEE Intell. Transp. Syst. Mag. **11**(2), 82–95 (2019)
6. Yang, B., Wang, J.: An improved helmet detection algorithm based on YOLO V4. Int. J. Found. Comput. Sci. **33**(6), 887–902 (2022)
7. Yang, J., He, W.Y., Zhang, T.L., Zhang, C.L., Zeng, L., Nan, B.F.: Research on subway pedestrian detection algorithms based on SSD model. IET Intel. Transport Syst. **14**(11), 1491–1496 (2020)
8. Zhang, L., Xu, R., Miao, H., et al.: Research on infrared dim and small target detection algorithm based on frequency domain difference. Int. J. Remote Sens. **43**(8), 11 (2022)
9. Wu, D., Wu, Q., Yin, X., Jiang, B., Wang, H., He, D.J., Song, H.B.: Lameness detection of dairy cows based on the YOLOv3 deep learning algorithm and a relative step size characteristic vector. Biosys. Eng. **189**, 150–163 (2020)
10. Huang, Z.J., Sui, B.W., Wen, J.Y., Jiang, G.H.: An intelligent ship image/video detection and classification method with improved regressive deep convolutional neural network. Complex. 2020, Article ID 1520872 (2020)

Chapter 14
Research and Construction of the Integrated Management System for the Grid-Connected Operation of the System Terminal

Jian Zhang, Bo Li, Ying Zeng, Xingnan Li, and Zhan Shi

Abstract With the rapid development of the electric power industry, it is particularly important to improve the management efficiency and reliability of the electric power system, and it is urgent to ensure the safe and stable operation of the system. Therefore, this paper studies and builds the integrated management system of system terminal grid connection operation. Using advanced computer technology and modern communication, to realize the comprehensive monitoring and control of all aspects of the power system, the system includes data collection, data analysis and management, information visualization and other functional modules. The experiment proves that the system has a high degree of reliability and security, and can realize adaptive and optimal control under changing environmental conditions. The integrated management system of grid connection operation of the system terminal can provide technical support and guarantee the reliable operation of the power system.

14.1 Introduction

The power system is one of the indispensable infrastructures in modern society, and it is the necessary energy support [1] to ensure the social and economic operation and people's lives. With the continuous development of China's power industry and the continuous opening of the power market, the scale of the power system is expanding, the operating environment is becoming more and more complex and the requirements of power system management are getting higher and higher. How to improve the management efficiency and reliability of the power system to ensure its the safe and stable operation is an urgent problem to be solved [2]. With the continuous

J. Zhang (✉) · B. Li · Y. Zeng · X. Li · Z. Shi
Electric Power Dispatching and Control Center of Guangdong Power Grid Co., Ltd., ,
Guangzhou 510000, China
e-mail: minwell@126.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_14

development of computer technology and communication technology, the technical level of terminal equipment and monitoring system is constantly improved, so that the comprehensive monitoring and control of [3] each link of the power system can be realized. This paper aims to study and construct a system to improve the management efficiency and reliability of power systems. This paper adopts advanced computer technology and modern communication to realize the comprehensive monitoring and control of all aspects of the power system. The system includes data collection, data analysis and management, information visualization and other functional modules. Through the research and construction of the integrated management system, the efficiency and reliability of the power system management are improved, and the technical support and support for the safe and stable operation of the power system [4].

14.2 General Architecture Design of the Integrated Management System for the Grid-Connected Operation of Electric Power System Terminals

Power system terminal grid operation integrated management system is data collection, processing, analysis and management in the integration of the integrated system; it is composed of multiple components, and the function of the whole system covers the power system terminal equipment monitoring, data acquisition and processing, running status evaluation, abnormal early warning and response, etc.; the overall architecture is shown in Fig. 14.1.

The functional requirement of the system terminal grid-connected operation integrated management system is to ensure the normal operation of the photovoltaic power generation system and to provide comprehensive monitoring, management and analysis functions. The following is an expansion of the functional requirements:

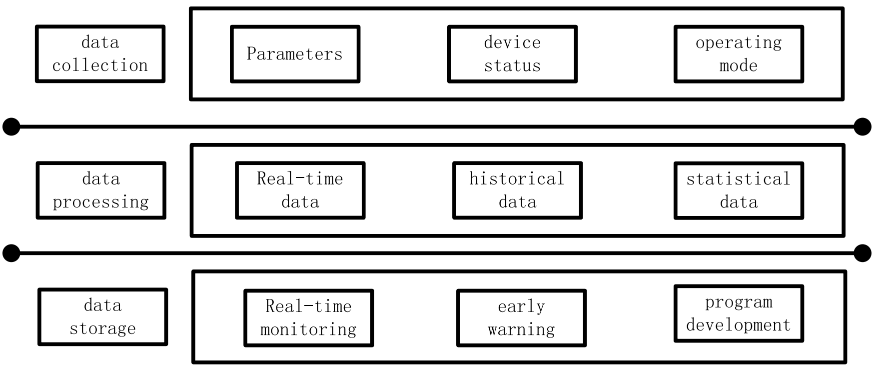


Fig. 14.1 Overall architecture design of the integrated management system

Real-time monitoring of terminal equipment parameters: the system needs to be able to monitor the operating state of the power system terminal equipment in real time, including temperature, voltage, current and other key parameters. Through sensors and monitoring equipment, collect equipment data and ensure the accuracy and timeliness of the data.

Data collection, storage and management: The system shall have the capability of data collection, and shall store and manage the data collected from the terminal equipment. Data storage can be performed using a database or cloud platform for subsequent analysis and queries.

Data processing and analysis: The system needs to clean, normalize and merge the collected data to ensure the quality and consistency of the data. Through data analysis and mining technology, the system can extract the trends and rules of the power system operation, to provide support for the system operation management.

Operating status assessment: Based on the collected data, the system shall be able to evaluate the operating status of the power system terminal equipment. Through the set rules and algorithms, the performance and health status of the equipment are analyzed and determined, and the operation quality evaluation index of the equipment is provided.

Abnormal monitoring and early warning: the system shall have the function of abnormal monitoring and early warning, and can timely detect the abnormal condition of the equipment and issue early warning signals. When the equipment has a failure, temperature abnormality or other abnormal conditions, the system can automatically send an alarm to notify the relevant personnel, in order to take timely measures for repair or maintenance.

Fault diagnosis and maintenance support: The system shall provide fault diagnosis and maintenance support functions to help the operation and maintenance personnel accurately locate the equipment faults, and provide corresponding maintenance guidance and suggestions.

Visualization interface and report generation: The system needs to provide an intuitive visual interface to display the monitoring data in the form of charts, graphics or maps, so that users can quickly understand the operating status of the power system. In addition, the system shall have the function of generating reports for recording and summarizing the operation and performance indicators of the system.

Through the realization of the above functional requirements, the system terminal grid operation integrated management system can provide comprehensive, accurate and real-time power system operation status information, help realize the efficient operation and management of photovoltaic power generation system and provide timely early warning and maintenance support, to ensure the stable operation of the power system, improve the efficiency and reliability of the system and provide decision support and troubleshooting information.

14.3 Hardware Design

The hardware design of the integrated management system is an important part of the system, including computer, network equipment, sensors and other hardware equipment. This system chooses the Advantech brand industrial grade computer; the model can choose AdvantechIPC-610H. Network equipment is a bridge connecting the various equipment and sensors of the system, which needs to meet the stability and real-time performance of the system data transmission. The system adopts Huawei S5720-36C-EI-A, with high speed, low delay, multi port, high reliability and other industrial grade switches. The sensor is the sensing equipment of the system, which needs to collect various parameter data on the site, including temperature, humidity, pressure, flow, etc. The model of the sensor is Honeywell PX2AG1XX025PSCHX and the quantity shall be configured according to the actual requirements of the system. In addition to the above hardware equipment, the system will also need to be equipped with the corresponding accessories, such as data cable, power cord and adapter. These need to be characterized by high quality, high reliability and low failure rate to ensure the stable operation of the system. The hardware design of the integrated management system should consider the factors of functional requirements, data processing capacity, communication performance and stability to select and configure appropriate equipment and accessories. At the same time, the corresponding standards and specifications should be followed to ensure that the system can operate stably and reliably.

14.4 Software Design

14.4.1 Data Acquisition Module

The data acquisition module is used to collect the real-time operation data of the power system terminal equipment, including the voltage, current, power, frequency and other parameters, as well as the equipment status, operation mode and other information [5]. Data collection can be uploaded and collected through intelligent terminals through field sensors and monitoring devices, but the collected data generally needs to be processed and cleaned to ensure the accuracy and reliability of the collected data [6]. Data cleaning technology is mainly used to process the original data; in order to remove noise, errors and outliers, we first need to deal with the missing data. The usual method is to fill the missing values with the mean, median, crowd, etc., and to remove repeated data, and the data with missing values can be filled by formula (14.1):

$$x_j = 1/N * sum_{i=1}^{N} x_i \tag{14.1}$$

Shown as formula (14.1), N is the number of data present, and i is the fill value of missing values. Data distribution analysis can include statistical indicators such as the central trend, degree of dispersion and deviation of the observed data. By calculating the indicators of the mean, the median, the standard deviation and the skewness, we have an intuitive understanding of the overall distribution of the data. Moreover, visualization tools such as drawing histograms, boxplots and scatter plots can also help to reveal the distribution patterns and anomalies of the data. After analyzing the data distribution, data points that do not match the data distribution can be excluded. These data points that do not match the data distribution may be outliers or erroneous records, and they may have a negative impact on subsequent data processing and analysis. By identifying and excluding these data points, we can maintain the accuracy and consistency of the data.

Next, to facilitate subsequent data processing, the original data will need to be converted into a target format. This may involve operations such as data type transformation, unit conversion, normalization or standardization to facilitate uniform data processing and comparison: for example, converting date and time data into a unified format, encoding text data or participle processing. In addition, in order to obtain a more comprehensive data perspective, it is necessary to gather different data sources together to form a whole data set and discover the laws and relationships between the data. By integrating data from different data sources, a more comprehensive analysis can uncover potential trends, patterns or correlations. Data pooling may involve operations such as data merging, connection, association and data integration. Ensuring data consistency, accuracy and completeness when integrating data is needed to avoid introducing errors or bias.

To sum up, nalysing data distribution, excluding abnormal data points, transforming data formats and integrating different data sources can provide a more accurate, complete and reliable data basis for subsequent data processing and analysis, help reveal the laws and relationships of data and obtain valuable insight and decision support.

14.4.2 Data Analysis and Management Module

The data analysis management module is used to store and manage the collected data, including real-time data, historical data, statistical data, etc. Data storage adopts Apache Cassandra distributed database, which has the characteristics of high scalability, high-speed reading and writing ability, and large capacity storage capacity, and can meet the data management requirements of the integrated management system of grid-connected operation of power system terminals. In addition, the collected data also needs to be analyzed and processed to extract useful information and indicators, such as power grid load prediction, power quality analysis, fault diagnosis and others [7]. Data analysis can use the traditional statistical analysis method; for the data with outliers, formula (14.2) can be used to process:

$$z = (\mathrm{x} - \mathrm{mean})/\mathrm{std} \tag{14.2}$$

where x is the raw data, mean is the mean of the data, std is the standard deviation of the data and z is the normalized data. If the z-value is greater than a certain threshold, the data point can be treated as an outlier. The data analysis and management module detects and handles the outlier values of the collected data by applying various outlier value detection methods. Among them, the standardized method (z-score) is a common method of outlier detection. This method transforms the data into a standard normal distribution by standardizing the raw data, where each data point is subtracted from the mean of the data and divided by the standard deviation of the data. Then, a threshold can be set to treat the normalized data point as an outlier when they exceed it. In addition to the standardized methods, the data analysis management module can also use cluster analysis-based and box-plot-based methods for outlier detection. Through effective outlier detection and processing, the credibility of data and the accuracy of analysis can be improved, so as to provide a more accurate basis for subsequent data analysis and decision-making.

14.4.3 Visualization Module

The visualization module is used to provide decision support for the power system operation and management personnel, including real-time monitoring, early warning, scheme formulation, instruction issuing, etc. Decision support can adopt a rule-based method, combined with artificial intelligence technology, to realize automatic decision-making, and collect, store and analyze decision data and information in the form of charts, curves, maps and other visualization, so as to facilitate users to intuitively understand the operating status and change trend of the power system [8]. Visual display can adopt modern interactive data visualization technology, such as large-screen displays, mobile terminal applications, virtual reality and other forms, to improve the efficiency of information communication and sharing [9, 10]. Artificial intelligence technology can automatically analyze and judge the operation state and change the trend of the power system through data mining, machine learning and other means, so as to provide more accurate decision support for the power system operation management personnel [11]. Commonly used visualization methods in the visualization module include histogram, line map, scatter map, heat map, map, etc. Among them, the heat map is a visualization method that can intuitively display the data density, and its mathematical model can be expressed as Eq. (14.3):

$$f(x, y) = \frac{1}{2\pi\sigma^2}\mathrm{e} - \frac{(x - x_0)^2 + (y - y_0)^2}{2\sigma^2} \tag{14.3}$$

where x_0 is the expected early warning value along the X-axis in the visual heat map, and y_0 is the expected early warning value along the Y-axis in the visual heat map. In the integrated management system of the power system, the heat map can be used to show the operation status of each equipment in the power system, so as to provide timely monitoring and early warning information for the operation management personnel of the power system and help them take timely measures [12, 13]. In addition, virtual reality technology can also be used in the integrated management system of the power system [14]. By establishing a three-dimensional model of the power system, data display and interactive operation in the virtual environment can improve the efficiency of information communication and sharing.

14.5 Test the Experiment

14.5.1 Experimental Preparation

In order to test the performance of the integrated management system of grid-connected terminal operation, the experiment used two photovoltaic power generation systems using different models of terminal equipment. This can evaluate the performance difference of the different terminal devices when operating the integrated management system. The first photovoltaic system uses model A1 terminals, while the second photovoltaic system uses model A2 terminals. In order to ensure the normal operation of the system. To carry out the test, the ThinkPad X1 Carbon Intel Core i5 flagship version was selected as our computer model. The laptop is equipped with a powerful processor and memory to provide enough computing power to support the operation of the integrated management system. In addition, the corresponding monitoring equipment and terminal equipment are also needed to realize the monitoring and control of the photovoltaic power generation system. To meet this need, JZ-KQ 1 monitoring equipment and NPort 6400/6600 series terminal equipment were selected for the experiment. The JZ-KQ 1 monitoring device is able to collect various data on the photovoltaic system, such as power, voltage, current and temperature information. The Nport 6400/6600 series terminal equipment provides a stable and reliable connection between the computer and the terminal equipment of the photovoltaic power generation system.

Before conducting the experiment, ensure that the computer and monitoring equipment are correctly connected to the terminals of the two photovoltaic systems. In this way, the remote monitoring and control of the two photovoltaic power generation systems can be realized through the system terminal grid connection operation of the integrated management system. Once the system equipment is connected, the integrated management system can begin to remotely monitor and control the photovoltaic system. Through the integrated management system, it can communicate with the terminal equipment of the photovoltaic power generation system, obtain real-time data and send control commands. In the experiment, the output power, voltage, current and other parameters of the photovoltaic power generation system can be monitored, and the corresponding control operation can be carried out according to the need. The collected data should be fully documented during testing. These data are very important for the subsequent performance analysis. Experiments can record the performance of photovoltaic power generation systems in different time periods, and compare the differences between the A1 model and the A2 model terminal equipment. In addition, the experiment should record the operation and control commands for the system in order to associate them with the changes in the performance.

14.5.2 Experimental Result

The experimental test results include the data of the total power generation, average power generation, maximum power generation and minimum power generation of the photovoltaic power generation system. Table 14.1 shows the data obtained from the remote monitoring and control of the two photovoltaic power generation systems through the system terminals when using different types of terminal equipment.

Data analysis for terminal device type A1: Values of generating capacity 1 (kW/h) are between 56, 58 and 57. The generating capacity of 2 (kW/h) is between 54, 57 and 55. The total generating capacity (kW/h) is the sum of capacity 1 and capacity 2, 110, 115 and 112, respectively. The average power generation (kW/h) is the average of the total power generation and is calculated as 55, 57.5 and 56. The maximum power generation (kW/h) is the maximum of generating capacity 1 and generating capacity 2, at 56 and 58, respectively. Minimum power generation (kW/h) is the smaller value of generating capacity 1 and generating capacity 2, at 54 and 57, respectively.

Data analysis for terminal device type A2: The value of generating capacity 1 (kW/h) is between 52, 53 and 51. The generating capacity of 2 (kW/h) is between 53, 51 and 50. The total generating capacity (kW/h) is the sum of capacity 1 and capacity 2, 105, 104 and 101, respectively. The average power generation (kW/h) is the average of the total power generation, and the calculated results are 52.5, 52 and 50.5. The maximum power generation (kW/h) is the maximum of generating capacity 1 and generating capacity 2, at 53 and 51, respectively. Minimum power generation (kW/h) is the smaller value of generating capacity 1 and generating capacity 2, at 52 and 50, respectively.

Table 14.1 Experimental result

Serial number	Terminal equipment type	Generating capacity1 (kW/h)	Generating capacity2 (kW/h)	Gross generation(kW/h)	Average power generation (kW/h)	Maximum power generation (kW/h)	Minimum power generation (kW/h)
1	A1	56	54	110	55	56	54
2	A1	58	57	115	57.5	58	57
3	A1	57	55	112	56	57	55
4	A2	52	53	105	52.5	53	52
5	A2	53	51	104	52	53	51
6	A2	51	50	101	50.5	51	50

By nalysing these data, the average power generation of terminal device type A1 is slightly higher than terminal device type A2, 56 and 52.33 kW/h, respectively. The maximum generating capacity of the terminal equipment type A1 is 58 kW/h, and the minimum generating capacity is 54 kW/h. The maximum capacity of terminal equipment type A2 is 53 kW/h and the minimum capacity is 50 kW/h.

It can be seen from the experimental results of the system that the use of the system terminal grid-connected operation of the integrated management system can improve the average and total generating capacity of the photovoltaic power generation system, effectively monitor and control the operation status of the photovoltaic power generation system and improve the efficiency of the system. Through the remote monitoring and control system, the faults in the photovoltaic power generation system can be found and solved in time, so as to improve the reliability and stability of the system [15]. The system can realize the real-time monitoring and remote control of the photovoltaic power generation system, and improve the flexibility and operability of the system, so that the system can better adapt to various environments and operating states.

14.6 Conclusion

The integrated management system plays a very important role in the operation of the power system, and the system terminal grid connection operation integrated management system is an important means to apply the integrated management system to the terminal equipment and realize remote monitoring and control. This paper adopts advanced computer technology and modern communication to realize the comprehensive monitoring and control of all aspects of the power system. The system includes data collection, data analysis and management, information visualization and other functional modules. The integrated management system can improve the efficiency and reliability of the power system management, reduce the operation cost and provide an important guarantee and support for the safe and stable operation of the power system. In the future research and practice, we can improve and optimize the shortcomings of the integrated management system of system terminal grid operation, strengthen the compatibility of equipment and user interface design, and enhance the scalability and maintainability of the system.

Acknowledgements This work is supported by the Key Science and Technology Project of China Southern Power Grid Co., Ltd. (036000KK52220038).

References

1. Zhou, J., Mao, Z.B., Shao, Z.Q., et al.: Research on the comprehensive benefit evaluation of low-carbon energy big data management system—The application scenarios and business model of "green energy code." Price Theory Pract. **1–5**, 03–09 (2023)
2. Liang, G.S., Zhang, W.: Development and application of FPSO integrated energy efficiency management system. Chem. Des. Commun. **48**(12), 24–26 (2022)
3. Zhao, Z.Q., Li S.X.: Design and implementation of an integrated smart energy management system. Autom. Appl., (12): 172–174 + 181 (2022)
4. Gui, Y., Zhou, L., Liu, F., et al.: Key technology for thermal management of high-power density propulsion system. Intern. Combust. Engine Accessories **22**, 4–6 (2022)
5. Xiong, D.H., Li, S.F., Liang, Y.K.: Research on river and lake information management system based on GIS and internet of Things. Internet Things Technol. **12**(11), 121–123 (2022)
6. Chen, S.H.: Design and application of intelligent property integrated operation and maintenance management system. Green Constr. Smart Build. **11**, 59–62 (2022)
7. Tang, X.: Functional design of an integrated intelligent energy management system. Energy-saving **41**(10), 60–62 (2022)
8. Xing, H.F., Che, H., Miao, B.Z., Yang, B.: Design and implementation of intelligent safety supervision integrated information management system. J. Tonghua Norm. Univ. **43**(10), 73–80 (2022)
9. Fu, Y., Li, W., Xiong, N., Fang, Y., He, X.M.,Wang, Z.F.: Influence of off-grid/grid-connected operation on stability of large-scale photovoltaic system. Energy Rep., **9**(S7) (2023)
10. Liu, Y., Guo, W.C., Ding, J.H.: Parameter uncertainty and sensitivity of pumped storage system with surge tank under grid-connected operating condition. J. Energy Storage, 62 (2023)
11. Elsir Mohamed, Al-Sumaiti Ameena Saad, El Moursi Mohamed Shawky, Al-Awami Ali Taleb.: Coordinating the day-ahead operation scheduling for demand response and water desalination plants in smart grid. Appl. Energy, 335 (2023)
12. Tong, Z.J., Zhang, C., Wu, X.T., Gao, P.C., Wu, S., Li, H.Y.: Economic optimization control method of Grid-Connected microgrid based on improved pinning consensus. Energs. 16(3) (2023)
13. Hamdan, I., Youssef Marwa, M. M., Noureldeen Omar.: Influence of interval type-2 fuzzy control approach for a grid-interconnected doubly-fed induction generator driven by wind energy turbines in variable-speed system. SN Appl. Sci., **5**(1) (2022)
14. Liu, Z.F., Zhao, H.Y.: Analysis of grid-connected operation characteristics and limit access capacity of offshore wind farms. Energy Rep., **8**(S13) (2022)
15. Moji L.P., Hohne P.A., Kusakana K., Numbi B.P.: Optimal operation of a hybrid multisource energy system considering grid load shedding. Energy Rep., **8**(S15) (2022)

Chapter 15
Accuracy Analysis and Verification of Demand Superposition Calculation Method Based on Power Curve

Jinjin Li, Fanghua Mo, Xiuqing Lin, and Yupeng Qin

Abstract At present, some enterprises use the method of multi-circuit demand superposition to obtain the actual maximum demand for the two-part electricity price settlement. There are two main ways to realize the multi-circuit demand superposition. One is to use the communication means to superimpose the power information of the electric energy metering devices of each circuit in real time to obtain the actual maximum demand. Another method is to collect and superimpose the power curve data of the metering device remotely through the information system. This paper evaluates the uncertainty of the demand superposition calculation method based on different power curves, analyzes and evaluates the accuracy of its calculation, and compares and verifies it with actual case data. The results show that the calculation method of dispatching the power curve combining the second-level curve and the minute-level curve has both economy and accuracy, and has great practical application value.

15.1 Introduction

According to the Interim Measures for the Management of Sales Electricity Prices, large industrial users with a transformer capacity of 315 kVA or above must implement a two-part electricity price. The two-part electricity price consists of two parts: the electricity degree price and the basic electricity price [1]. The electricity degree price is charged based on the user's electricity consumption, while the basic electricity price is charged based on the user's transformer capacity or maximum demand [2].

J. Li · X. Lin · Y. Qin
Guangxi Power Grid Corporation Metering Center, Nanning Guangxi 530023, China

F. Mo (✉)
Guangxi Power Grid Corporation, Nanning Power Supply Bureau, Nanning Guangxi 530023, China
e-mail: 357030634@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_15

The maximum demand refers to the measurement of the average power consumption of customers during a certain period of time during the agreed settlement period, with its maximum indicated value retained as the maximum demand for the settlement period. For users with multi-circuit power supply, the real-time power of each circuit may not necessarily reach its maximum value at the same time, and the actual maximum demand value is often less than the sum of the maximum demand values of each circuit. With the approval of both the power grid company and the customer, the actual maximum demand can be calculated using the method of demand superposition, which is used to settle the basic electricity bill [2–5].

15.2 Current Situation of Demand Superposition

There are currently two main ways to achieve demand superposition: one is to perform superposition calculation locally on site through demand superposition devices, and the other is to perform superposition calculation on power curve data collected remotely through information systems.

15.2.1 Local Stacking

The demand superposition device is based on the on-site electric energy metering device, adding a demand superposition module to the master meter [6]. Through communication lines, the electrical energy data of the sub-meter is superimposed on the master meter, and the superimposed demand value is obtained through calculation.

The demand superposition device directly uses data from on-site devices for real-time stacking, and communicates between sub-tables and parent tables in real time, with high accuracy. At the same time, the maximum demand value can be directly read out at the table end, which is also in line with the customary mode of obtaining measurement values through physical measuring instruments [7].

However, this method also has shortcomings. Firstly, there are currently no relevant verification regulations to verify and test the demand superposition function; secondly, it is only applicable to the demand superposition of parent and sub-tables in the same place, and does not support scenarios where measuring instruments for each circuit of remote meters are installed in different places; thirdly, the demand superposition function is expensive, with a set of demand superposition devices costing nearly 100,000 yuan. Currently, only some imported meter manufacturers have developed the demand superposition function for on-site devices.

15.2.2 Remote Stacking

With the construction of measurement automation systems, a large amount of data line information from electricity meters has been remotely collected. The collected data items include voltage, current, power, and other curve data. By overlaying the curve data of multiple circuits at the same time point and taking the maximum value of slip, the actual maximum demand value can be obtained [8].

The characteristic of remote stacking is that it does not require additional modification equipment on site and is more convenient to implement. The disadvantage is that this method requires high data collection requirements, including collection frequency, collection success rate, and clock accuracy. The quality of data collection directly affects the accuracy of calculations.

Whether it is local or remote stacking, it is necessary to evaluate the uncertainty of the acquired demand for stacking, so as to judge the accuracy of the stacking demand [9]. The demand obtained by the local stack and remote stack has not been verified, has no legal effect, and can only be used as a reference, with the consent of both the power company and the customer.

15.3 Uncertainty Analysis of Power Curve Superposition

15.3.1 Measurement Model

The power superposition formula for multi-circuit power supply based on a power curve is

$$P_{\max} = \mathrm{Max}\left(\frac{1}{n}\sum_{i}^{n+i-1} p_i\right) \tag{15.1}$$

p_i—Multi-loop power and value at acquisition point i

n—Collection points within the demand cycle.

In the power system, the demand period for maximum demand is 15 min, and the demand is calculated based on a slip period of 1 min. Therefore, the collection interval of the power curve should be less than or equal to 1 min [10].

At present, the minimum collection interval for load curves supported by electricity meters is 1 min; the minimum collection interval for the load curve of the scheduling system can reach 1 s.

15.3.2 Analysis of Sources of Uncertainty

There are two main sources of uncertainty in analyzing mathematical models:

(1) The uncertainty introduced by the acquisition accuracy;
(2) The uncertainty introduced by the collection frequency.

Uncertainty Introduced by Acquisition Accuracy

The uncertainty brought by different collection devices varies. For metering devices, which usually include energy meters and transformers, assuming the accuracy level of the energy meter is *a* and the accuracy level of the measuring transformer is *b*, in the case of uniform distribution, according to the uncertainty propagation theorem, the uncertainty u_1 introduced by the acquisition accuracy level is

$$u_1 = \sqrt{\left(\frac{a}{\sqrt{3}}\right)^2 + \left(\frac{b}{\sqrt{3}}\right)^2} \tag{15.2}$$

Uncertainty Introduced by Acquisition Frequency

The maximum demand is to collect power data at a certain frequency, and then compare the collected power data to take the maximum value as the maximum demand. Therefore, the uncertainty introduced by the acquisition frequency includes two components: one is the uncertainty component caused by the maximum deviation of the sampling value itself, and the other is the uncertainty component generated by comparing different acquisition point sizes [11].

(1) uncertainty introduced by maximum deviation.

Within a collection interval, it can be considered that there are countless instantaneous values of the actual load, and the instantaneous value of the load curve is only a randomly obtained value among them, and this value is the maximum uncertainty depending on the extreme deviation. The maximum deviation is the average deviation degree between a set of data and its maximum value, and the calculation formula is as shown in (4):

$$s_{\max} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - x_{\max})^2}{n-1}} \tag{15.3}$$

x_i is the instantaneous value collection point, x_{max} is the maximum value of the instantaneous value collection point, and n is the number of collection points. S_{max} is the extreme value deviation of the collection point.

Therefore, the uncertainty u_2 introduced by the maximum deviation is

$$u_2 = \frac{s_{\max}}{x_{\max}} \tag{15.4}$$

(2) uncertainty introduced by comparison.

The maximum demand is obtained by comparing the values of each collection point one by one. Due to the different deviations in the maximum values of different collection points, comparing only the values of the collection points themselves can generate uncertainty.

X_n is any point in the collection point except x_{max}, and the maximum deviation of x_n is s_n. If $x_{max} + s_{max} > x_n + s_n$ is true for any point, then there is no misjudgment in $x_{max} > x_n$, and no comparative uncertainty is introduced.

If $x_{max} + s_{max} < x_n + s_n$ exists, it is judged that $x_{max} > x_n$ will have uncertainty, and the introduced comparison uncertainty is the maximum $(x_n + s_n)$ value minus $(x_{max} + s_{max})$, so the introduced uncertainty u_3 by comparison is

$$u_3 = \frac{x_n + s_n - (x_{\max} + s_{\max})}{x_{\max}} \tag{15.5}$$

By synthesizing the uncertainty introduced by the maximum deviation of collection point and the comparison of collection point values, the uncertainty introduced by collection frequency u4 is

$$u_4 = \sqrt{u_2^2 + u_3^2} \tag{15.6}$$

Relationship Between Uncertainty of Different Acquisition Frequencies

If a certain collection interval is further subdivided for collection, the maximum power value within the original collection interval will fall into one of the subdivided intervals, and its maximum value will remain unchanged. However, as the collection frequency increases, more values will be collected, and there will inevitably be points larger than the original collection values [12]. Therefore, as the collection frequency increases, the uncertainty introduced will decrease.

15.3.3 Uncertainty Synthesis

Synthesize uncertainty components based on uncertainty propagation rate

$$u_c = \sqrt{u_1^2 + u_4^2} \tag{15.7}$$

Take the inclusion factor $k = 2$, then the extended uncertainty is

$$U = 2u_c \tag{15.8}$$

15.4 Comparison of Uncertainty Between Different Data Sources

At present, the main systems for collecting user power data include measurement automation systems and scheduling systems. The maximum collection interval for the load curve of the energy meter itself in the measurement automation system is 1 min per point. The scheduling system will also collect power curves of large users through measurement and control devices, with a minimum collection interval of up to 1 s per point [13].

15.4.1 Uncertainty Introduced by Acquisition Accuracy

The uncertainty brought by different collection devices varies. For metering devices, they usually include energy meters and transformers. The accuracy level of energy meters used for large users is generally level 0.5, while the accuracy level of transformers used for metering is level 0.2, which is uniformly distributed. According to the propagation rate of uncertainty [14]:

$$u_{\mathrm{j}1} = \sqrt{\left(\frac{0.5\%}{\sqrt{3}}\right)^2 + \left(\frac{0.2\%}{\sqrt{3}}\right)^2} = 0.3109\,\% \tag{15.9}$$

For the measurement and control device of the dispatch system, the accuracy level of the measuring transformer is level 0.5, and the accuracy level of the measurement and control device is level 0.5. According to the uniform distribution and uncertainty propagation rate, that is

$$u_{d1} = \sqrt{\left(\frac{0.5\%}{\sqrt{3}}\right)^2 + \left(\frac{0.2\%}{\sqrt{3}}\right)^2} = 0.4082\,\% \tag{15.10}$$

15.4.2 Uncertainty Introduced by Acquisition Frequency

The measured curve data comes from the load records at the electricity meter end, and the minimum collection interval for the load records supported by the electricity meter is 1 min per point [15]. The scheduling data comes from the measurement and control device, and the minimum interval of its power curve can reach 1 s per point.

(1) Uncertainty introduced by maximum deviation.

During the same period of time, the maximum value of the design quantity curve data is $x_{j\text{-}max}$, and the maximum deviation of the maximum value is $s_{j\text{-}max}$;

the maximum value of the scheduling curve data is $x_{d\text{-}max}$, and the maximum deviation of the maximum value is $s_{d\text{-}max}$. Due to the small collection interval of scheduling data, which is 1 point per second, by observing a large amount of scheduling second-level curve data, its value remains unchanged within multiple collection intervals, and there is no fluctuation in the load value, that is, the maximum deviation is $s_{d\text{-}max} = 0$.

Therefore, for the scheduling load curve, the uncertainty introduced by the collection frequency is 0, so the uncertainty introduced by the maximum deviation of the collection point u_{d2} is 0.

Due to the fact that the actual load during the same period is consistent with the maximum value of the load, that is

$$x_{j-\max} + s_{j-\max} = x_{d-\max} + s_{d-\max} \tag{15.11}$$

Due to $s_{d\text{-}\max} = 0$, $x_{j\text{-}\max} + s_{j\text{-}\max} = x_{d\text{-}\max}$.

Then that is

$$s_{j-\max} = x_{j-\max} - x_{d-\max} \tag{15.12}$$

Therefore, according to Eq. (15.4), that is

$$u_{j2} = \frac{s_{j-\max}}{x_{j-\max}} = \frac{x_{d-\max} - x_{j-\max}}{x_{j-\max}} \tag{15.13}$$

(2) Uncertainty introduced by comparison.

Due to the absence of maximum deviation in scheduling second-level data [16], there is no isjudgement in its comparison, and its comparative uncertainty is 0, $u_{d3} = 0$.

When the measurement data is $x_{j\text{-}max} + s_{j\text{-}max} > x_{jn} + s_{jn}$, $u_{j22} = 0$. In the case of $x_{j\text{-}max} + s_{j\text{-}max} < x_{jn} + s_{jn}$, the uncertainty introduced by the comparison of measurement data is as follows:

$$u_{j3} = \frac{x_{jn} + s_{jn} - (x_{j-\max} + s_{j-\max})}{x_{j-\max}} \tag{15.14}$$

According to Eq. (15.11), that is

$$u_{j3} = \frac{x_{dn} - x_{d-\max}}{x_{j-\max}} \tag{15.15}$$

(3) Uncertainty introduced by synthetic acquisition frequency.

According to the uncertainty propagation theorem, in the measurement curve data, the uncertainty u_{j2} introduced by the maximum deviation and the uncertainty u_{j3} introduced by the comparison were synthesized, and the uncertainty u_{j4} introduced by the acquisition frequency of the measurement curve data was obtained as follows:

$$u_{j4} = \frac{x_{dn} - x_{j-\max}}{x_{j-\max}} \tag{15.16}$$

Similarly, in the scheduling curve data, the uncertainty u_{d2} introduced by the maximum deviation is combined with the uncertainty u_{d3} introduced by the value to obtain the uncertainty u_{d4} introduced by the acquisition frequency of the scheduling curve. Since $u_{d2} = 0$ and $u_{d3} = 0$, the synthesized uncertainty $u_{d4} = 0$.

15.4.3 Comparison of Uncertain Sizes

According to Eq. (15.7), the scheduling curve data calculates the uncertainty u_d introduced by demand, and the measurement curve data calculates the uncertainty u_j introduced by demand, as follows:

$$u_d = \sqrt{u_{d1}^2 + u_{d4}^2} = u_{d1} = 0.04082\% \tag{15.17}$$

$$u_j = \sqrt{u_{j1}^2 + u_{j4}^2} = \sqrt{0.3109\%^{\ 2} + u_{j4}^2} \tag{15.18}$$

When u_{j4} is greater than 0.2645%, the uncertainty of u_{j} is greater than u_d.When u_{j4} is greater than 0.2645%, there is

$$u_{j4} = \frac{x_{dn} - x_{j-\max}}{x_{j-\max}} = \frac{x_{dn}}{x_{j-\max}} - 1 \tag{15.19}$$

$$\frac{x_{dn}}{x_{j-\max}} > 1.0026 \tag{15.20}$$

That is to say, when the maximum value of the second-level curve is greater than 1.0026 times the maximum value of the minute-level curve, the uncertainty of the metering curve calculation will be greater than the uncertainty of the scheduling data. Therefore, the accuracy of calculating the maximum demand using the scheduling curve should be higher [17].

15.5 Instance Validation

Select a power supply user for a certain circuit, which is a 220 kV steel plant with a total of 3 circuits. The main wiring diagram is shown in Fig. 15.1.

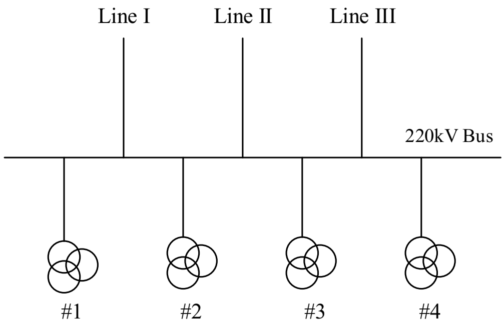


Fig. 15.1 Main wiring diagram

Among them, the user has a total of 3 circuits, namely I line, II line, and III line. The metering point is on the incoming side of the line, and the user has installed a demand superposition device on site. The maximum demand is calculated by using on-site demand superposition devices, minute curve data of energy meters, and scheduling second-level curve data, respectively.

15.5.1 Demand Data of the Superposition Devices

The maximum demand value obtained by the user through the demand stacking device in August 2020 is 193.05 MW, and the occurrence time is 22:37 on August 1, 2020.

15.5.2 Minute Level Curve Data of Measurement Automation System

Super-position the minute and curve data of each line to obtain the superimposed power and curve. According to the calculation method of maximum demand, the maximum demand value is 193.88 MW.

15.5.3 Using Scheduling Data on a Second-Level Curve

Super-position the scheduling second-level curve data of each line to obtain the superimposed power and curve. According to the calculation method of maximum demand, the maximum demand value is 193.40 MW.

15.5.4 Comparison

The difference between the minute curve of the metering automation system and the demand superposition device is 0.43%, and the difference between the calculated data of the scheduling curve and the demand superposition device is 0.18%.

It can be seen that after using a large amount of curve data for calculation, the maximum demand is very close to the demand superposition device, and it also indicates that the smaller the collection interval of curve data, the higher the accuracy of its calculation.

In the case where the collection interval of curve data in the measurement automation system cannot be further reduced, good calculation accuracy can also be achieved by using second-level curve data from scheduling data.

15.6 Conclusion

This article provides a method for calculating the maximum demand by using load curves for superposition, while allowing for demand superposition. The uncertainty of this method is within a certain range and has a certain rationality. The degree of uncertainty depends to a certain extent on the fluctuation of the load. If the fluctuation rate is relatively large, the uncertainty is large and it is not suitable to use this method for calculation. The calculated value can be used as a reference. When using this method, it is necessary to calculate its measurement uncertainty to determine the reliability of its measurement results. This method is a feasible reference method that is acceptable to both parties when other measurement conditions are not available.

With the increase of power supply circuits for power customers, the cost of on-site demand superposition device transformation will also increase further. The application of the method of calculating superposition demand based on curve data introduced in this paper will be further emphasized and studied, and the collection frequency of metering curve data can be further improved in the future, so as to reduce the uncertainty caused by the calculation of demand superposition of metering curve data.

References

1. Chen Qing.: Selection of the agreed maximum demand in two-part electricity pricing: Power demand side management **14**(4), 53–54 (2012)
2. Ying, M., Hui, W.: Analysis of the reasons affecting the accuracy of maximum demand: Jiangxi Electric. Power **33**(6), 52–54 (2009)
3. Su Zhen: Problems in the current maximum demand measurement and charging regulations: power demand side management, **6**(6), 25–26 (2004)
4. Zuming, F.: Research on achieving maximum demand superposition function based on Langier energy meters. Commun. World **11**, 163–164 (2015)

5. Yahua, C., Ningyu, H.: Research on the accuracy of measurement based on double power circuit user maximum demand superimposition. Guizhou Electr. Power Technol. **20**(1), 57–58 (2017)
6. Ye Rongxuan.: Research on electric energy demand model and control strategy. Nanjing: Nanjing Univ. Technol. (2007)
7. Su Hongbang, Qi Yanshou, Yang Yongqian, et al. Analysis of abnormal maximum demand for a 4-circuit power customer of State Grid Qinghai Electric Power: Qinghai Electric Power, 1, 65–68 (2021)
8. Huang Yafen.: Exploration of the application of electricity information collection system (load control terminal) in smart Grid: Hangzhou, Zhejiang University (2012)
9. Song Yuexin.: Research on key technologies for accurate management of industrial load electricity demand: Jinan, Jinan University (2020)
10. Baogeng, Y.: Research and application of maximum demand regulation technology for large enterprise power grids: automation technology and application **37**(3), 14–106 (2018)
11. Li Shaodong, Sun Minghui, Huang Aili.: Research and application of multi circuit and different transformer ratio demand superposition technology in electricity metering: Hongshuihe, **138**(4), 58-61(2019)
12. Zhihan, X., Liu, O.: Analysis of abnormal situation in meter reading and settlement for demand settlement users: Power demand side management **17**(2), 46–48 (2015)
13. Lei, Z.: Research on strategies for improving China's two part electricity pricing: China electric power. Education **8**, 162–164 (2013)
14. Mingcheng, J., Fuguo, J., Jiannan, Z.: A brief discussion on the two part electricity price: Northeast electric power technology **8**, 30–32 (2016)
15. Wang Feijun.: Research on design and application of marketing management system for county level power supply enterprises: Beijing, North China electric power university (2016)
16. Daibo.: Research and application of customer demand superposition technology for multi loop power supply based on measurement automation systems. Electromechanical Inf., 27, 128–130 (2016)
17. Zhihong, H.: Application of demand superposition technology in electricity metering. Yunnan Electr. Power Technol. **2**, 48–50 (2008)

Chapter 16
Construction of Security Control and Protection of Internal and External Networks of Communication Terminals Under the Heterogeneous Network Environment

Jian Zhang, Ying Zeng, Bo Li, Xingnan Li, and Zhan Shi

Abstract The construction of network security control and protection system of internal and external communication terminals under the heterogeneous network environment can promote the research and development of power network security technology, promote the information construction and intelligent upgrading of the power grid, and improve the overall operation level and competitiveness of the power grid. Based on this, this paper proposes the construction of security control and protection of communication terminals under the heterogeneous network environment of the power grid. By analyzing the condition of grid heterogeneous network environment, the communication terminal network security control system to build design, build the network isolation, certification and authorization, traffic monitoring, and security reinforcement, in order to improve the network heterogeneous network environment communication terminal network security control efforts, comprehensive guard against all kinds of network attacks or malicious software intrusion, improve the overall security performance of power grid system.

16.1 Introduction

Heterogeneous network refers to the network environment composed of multiple different types of networks, including the traditional wired network and wireless network, as well as all kinds of new networks such as the Internet of Things and 5G. Surrounded by the multiple network environment, the security control and protection of the internal and external networks of communication terminals is particularly important [1]. The security protection system designed in the internal and external

J. Zhang (✉) · Y. Zeng · B. Li · X. Li · Z. Shi
Electric Power Dispatching and Control Center of Guangdong Power Grid Co., Ltd., Guangzhou 510000, China
e-mail: minwell@126.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_16

network of the communication terminal can effectively carry out real-time monitoring and security protection of the power grid, which is conducive to ensuring the overall safe and stable operation of the power grid, and can improve the reliability and security of the power grid to a certain extent [2]. Based on this, according to the current needs of the power grid, the overall design of the security control and protection system under the heterogeneous network environment. By constructing the full aspects of the system, enhance the security of data transmission and user authentication from the authentication and authorization design. By monitoring the data flow of the power grid, realize the functional requirements of the system to automatically identify and handle the power grid security events. If a problem occurs, the security audit system can promptly issue alarms and address the corresponding issues, thereby enhancing the efficiency of power grid security management and emergency response. It helps to safeguard against various types of cyber attacks and malicious software intrusions, protecting network and information security.

16.2 Overall Design of the Security Control and Protection System

Under the heterogeneous network environment of the power network, the network environment involved includes internal network, public network, Internet of things, etc., the characteristics of the network environment are different, with the deepening of the power grid information construction, the power grid is facing increasing network security threats, the power grid contains a large number of key information, Such as: energy production, transmission and distribution, equipment operation and other data, once such data leakage or damage will cause serious impact and loss on the operation of the power grid [3]. To strengthen the power grid communication terminal network's internal and external security, to prevent all kinds of network attacks or malicious software intrusion, through the communication terminal network's internal and external security protection system construction, the control system can realize from the application layer to the network layer comprehensive security protection, improve the use of the overall network security and stability, protect the information security of communication terminal, reduce manual intervention, improve the management efficiency, shorten the troubleshooting time [4]. The general design of the security control system of internal and external networks of communication terminals under heterogeneous network environment is shown in Fig. 16.1.

As can be seen from Fig. 16.1, the security control system of the communication terminal is constructed by the application layer, system layer, data layer and network layer respectively, the application layer contains various applications and security management control platforms, mainly responsible for implementing various security functions such as authentication, access control, security audit, etc.; the system layer is various operating systems and system management tools, It is mainly responsible for providing basic system resource management and security control functions,

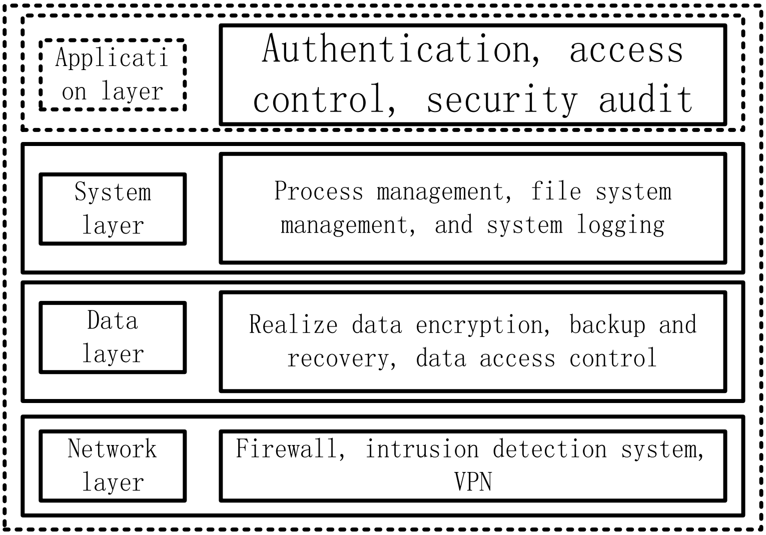


Fig. 16.1 General design diagram of the internal and external network security control system of the communication terminal

such as process management, file system management, system logging, etc. [5]. The data layer includes various data storage and management tools, mainly responsible for realizing data encryption, backup and recovery, data access control and other functions to ensure the security and reliability of data; the network layer is various network equipment and security protocols, mainly responsible for realizing the security and reliability of network communication, such as firewall, intrusion detection system, VPN, etc. [6].

16.3 Software Design

16.3.1 Internal and External Network Isolation

Network isolation and security authentication are the key steps in guarding against cyber attacks and data leakage. Filter and control of network traffic through network devices, to realize the physical isolation and logical isolation between different network areas, its purpose is to prevent attackers from spreading viruses, Trojan horses and other malicious code through the network, so as to avoid damage to terminal devices [7]. By deploying the Virtual Private Network (VPN) server, the data of the internal network is encrypted, the VPN client is installed on the terminal device of the internal network, and configured to establish a secure communication tunnel with the VPN server. When the internal network terminal devices request the

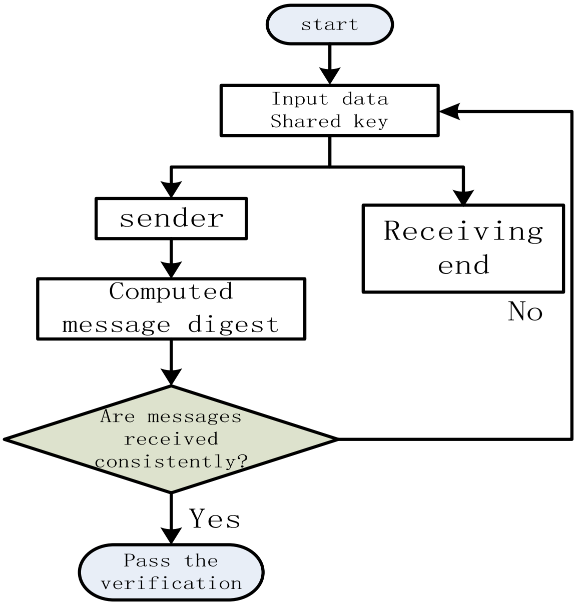


Fig. 16.2 SHA-2 authentication process

VPN server through the VPN client to establish the secure communication tunnel. After establishing the VPN connection, the data transmission between the internal network terminal device and the external network will be encrypted to ensure the confidentiality and integrity of the data transmission [6]. The SHA-2 encryption hash algorithm in the communication side can be verified by the configuration in the VPN device. Its implementation process is shown in Fig. 16.2.

According to Fig. 16.2, the process of SHA-2 authentication is realized by sharing the key in the communication. During setting up the key, note that the key should be sufficiently long and complex. To ensure safety, during the communication process, the sender takes as input the data to be sent and the shared keys and calculates a message summary using the SHA-2 algorithm. The sender sends the calculated summary of the message to the receiver [8]. After the receiving party receives the message, also takes the data to be received and shared keys as input, calculates a message summary using the SHA-2 algorithm, the receiver compares the calculated message summary with the message summary sent by the sender. If the same occurs, the authentication is considered to have passed. Otherwise, the authentication is considered to fail [9].

16.3.2 Communication Terminal Authentication and Authorization

The secure communication between the communication terminal and the authentication server can be realized using the TLS/SSL protocol. The communication terminal applies for a certificate from the authentication server [10]. After verifying the identity of the communication terminal, the authentication server generates the certificate and returns it to the communication terminal. The certificate contains the public key of the communication terminal and the public key of the authentication server, and the authentication and authorization structure are shown in Fig. 16.3.

During the TLS/SSL handshake process, the communication terminal and the authentication server exchange information such as the certificate and encryption algorithm, and conduct authentication and key negotiation [11]. The communication terminal and the authentication server use the symmetrical key generated during the handshake process to encrypt and decrypt the communication data to ensure the confidentiality and integrity of the communication. The communication terminal and the authentication server need to configure the TLS/SSL protocol to ensure the security of the protocol. The configuration items include the password suite, key length, certificate expiration time, TLS version, etc. Before transmitting the data, the communication terminal needs to encrypt the data, and the authentication server needs to decrypt the data. After receiving the data, the authentication server needs to verify the data to ensure the integrity of the data and the legitimacy of the identity. During the communication process, abnormal situations of TLS/SSL handshake failure, certificate verification failure, and key negotiation failure may occur [12]. The communication terminal and the authentication server need to handle such abnormal situations and take corresponding security measures, such as connection interruption, recording and logging, etc.

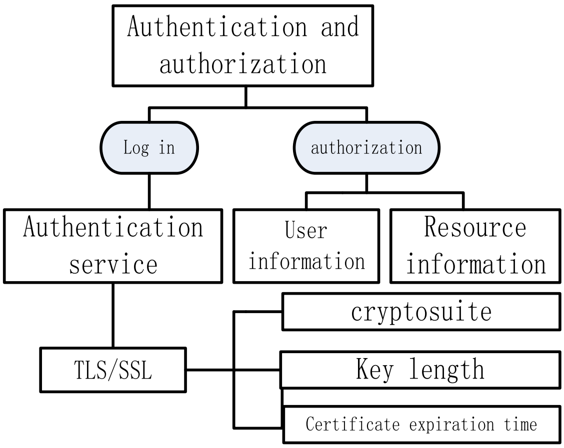


Fig. 16.3 Communication terminal authentication and authorization structure

16.3.3 Data Traffic Monitoring

Using tcpdump can monitor the security control of internal and external networks of communication terminals in the heterogeneous network environment. In Linux system, use sudo apt-get install tcpdump command and use ifconfig command to view the network card, IP address, MAC address and other information network interfaces. According to the network interface required to be monitored, select the corresponding network card for monitoring [13]. With the tcpdump command, you can set the corresponding filtering rules according to the protocol, source IP address, target IP address, port and other information to be monitored, and then start the network traffic monitoring and record the corresponding data. The Wireshark network protocol analysis tool is used to analyze the captured data, find abnormal traffic, determine network performance problems, find security vulnerabilities, etc., according to the monitoring data analysis results, and realize the system alarm and defense measures [14]. Its test method is shown in Fig. 16.4.

Using Wireshark analysis and detection, we can effectively detect and defend against network security problems. After opening the Wireshark software, select the network interface from which you want to capture packets and start capturing.

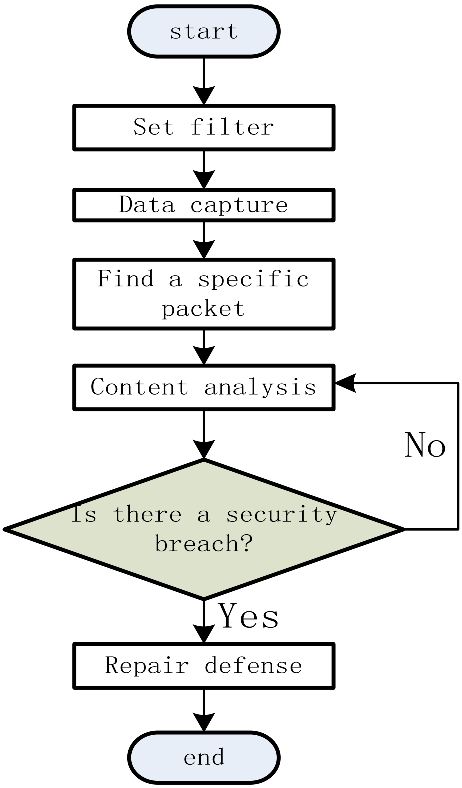


Fig. 16.4 Wireshark Flow chart of network protocol analysis

Wireshark will listen to the selected network interface and display the captured packets [15]. By using Wireshark's filtering capabilities, you can selectively display the packets of interest. The filtering expressions allow you to limit the displayed packets based on criteria such as source/destination IP addresses, port numbers, and protocol types. For example, filtering packets with the source IP address of 192.168.0.1: ip.src = = 192.168.0.1; filtering packets with source or destination port numbers of 80 or 443: tcp.port = = 80 or tcp.port = = 443; filtering packets with the TCP protocol: tcp. By using Wireshark's analysis features, you can inspect the captured packets and identify potential security vulnerabilities through packet analysis. Based on the analysis results, you can then set up alert rules to implement the system's alerting functionality.

16.3.4 Safety Reinforcement

To set alarm rules, you must first determine the monitoring and alarm targets. These include abnormal traffic, abnormal behavior of specific protocols, and security vulnerabilities. Second, the packets captured by Wireshark are analyzed and noted for potential security issues or patterns of unusual behavior. For example, frequent connection attempts, a large number of error packets, abnormal packet sizes, and so on. Based on the analysis results, you can define alarm rules to detect and trigger alarms. Alarm rules can be based on specific packet characteristics, traffic modes, and protocol behaviors. For example, "If an unauthorized access attempt (such as an unknown IP address) is detected, an alarm is triggered; an alarm is triggered if the abnormal behavior of a specific protocol, such as an abnormal HTTP request or DNS query, is detected. This section describes how to configure alarm rules in the system to implement the alarm function. Then, determine the severity of the alarm and the notification method. Alarm levels can be classified into different severity levels to facilitate alarm classification and handling [16]. Alarms can be notified by email, short message, or instant message to ensure that related personnel receive alarm notifications in a timely manner. Finally, periodically evaluate the validity and adaptability of the alarm rules. This section describes how to update and optimize alarm rules based on the actual situation and experience to improve alarm accuracy and reduce the false positive rate.

16.4 Conclusion

In the heterogeneous network environment of the power grid, the internal and external network is of positive significance for the long-term construction of power grid. Through adopting modern security management technologies, such as VPN, tcpdump and AES, it can effectively collect, process, visualize and analyze various types of log data to help power grid administrators quickly discover security threats and abnormal events. At the same time, with the help of advanced network isolation, certification and authorization and other security technologies, the comprehensive security protection and control of the internal and external networks of communication terminals can be realized to ensure the stable operation of the power grid in the heterogeneous network environment of the power grid. In the future, further exploration and application of emerging security technologies such as blockchain, AI, and machine learning can be conducted to continuously enhance the level of power grid network security and ensure the reliability, security, and stability of the power grid.

Acknowledgements This work is supported by The Key Science and Technology Project of China Southern Power Grid Co., Ltd. (036000KK52220038).

References

1. Zhao, H.S., Sun, J.J., Peng, Y.H., et al.: Radio power terminal identification method based on multiscale windows and regional attention residual network. J. Electr. Technol. **38**(01), 107–116 (2023)
2. Huang, J.T., Wang, Z.Y.: Design of electric power safety production monitoring system for real-time data analysis. Energy Environ. Prot. **44**(12), 256–261 (2022)
3. Li, Y.C., Yang, H.F., Cui, J.B., et al.: D2D communication security and reliability analysis of power 5G network based on the nearest location relay selection strategy. Comput. Appl. **42**(S2), 168–174 (2022)
4. Kang, C.Q., Chen, Q.X., Su, J., et al.: Scientific problems and research framework of large-scale flexible resource virtual power plant. Autom. Electr. Power Syst. **46**(18), 3–14 (2022)
5. Qian, W., Liu, S.S., Sun, J.Y., et al.: Design of power data monitoring and management system based on component security situation awareness. Big Power Data **25**(08), 76–83 (2022)
6. Wu, Q.: Thinking on security problems in embedded communication system design based on Zigbee—Review of zero trust network: building security system in untrusted network. Chin. J. Saf. Sci. **31**(09):199 (2021)
7. Li, Y.F., Liu, R., Cao, L.J., et al.: Research and practice of power mobile application and terminal security traceability management and Control technology. Inf. Secur. Commun. ConfidIty. **08**, 94–100 (2022)
8. Lan, Z., Wu, F.R., Yu, X.P., He, D., Tu, C.M., Xiao, F.: Transient instability analysis and stability strategy research of isolated island microgrids containing heterogeneous micro sources. Power Grid Technol. 1–16 (06–26) (2023)
9. Ma, X.M., Dong, C., Mao, X.Y., Jiao, Y.X., Li, H.: A method for compressing and storing massive multi heterogeneous smart grid data based on state estimation. Mot. Control. Appl. **50**(02): 67–72+81(2023)
10. Lan, Z., Diao, W.Y., Zeng, J.H., He, D., Tu, C.M., Jiang, F.: Pre synchronization control strategy for virtual synchronous generators in heterogeneous micro source isolated island microgrids. Power Syst. Autom. **46**(19), 154–161 (2022)
11. Chen, X., Xin, Y.Z., Zhao, W.Y., Sun, B.Y., Talipeng Nurbaheti.: Multi source heterogeneous data fusion model for smart early stage platform of power grid construction. J. Electr. Power, **37**(01): 76–83 (2022)
12. Qiu, H.J., Lian, G.X.: Design of power grid security perception system based on heterogeneous dual active data model continuity multilevel indicators. Microcomput. Appl. **37**(08), 197–200 (2021)

13. Peng, Y.L., Huang, W., Shuai, Z.K.: Research on instantaneous active power distribution of isolated island micro grid with heterogeneous micro sources. J. Electr. Eng. China **41**(15), 5167–5179 (2021)
14. Leng, O.Y., Hui, Q., Song, Y.P., Sun, P.: Optimal control model for power output of transmission end grid considering uncertainty and data heterogeneity of renewable energy. Renew. Energy **38**(08), 1116–1121 (2020)
15. Xu, D.M., Li, Z., Cui, G.Z., Hao, W.J.: Distributed finite time secondary control of heterogeneous battery energy storage systems in isolated microgrids. Control. Decis. **36**(08), 2034–2041 (2021)
16. Yan, M.Q., Leng, D., Liu, W.: Research on load balancing algorithm for heterogeneous wireless networks in smart grid based on matter element analysis. Commun. Power Technol. **36**(09), 114–115 (2019)

Part III
Approaches Based on Multimedia and Big Data Analysis

Chapter 17
Strategies for Improving Collaborative Filtering Library Circulation Services Based on AR Technology Under the Perspective of Five Education Initiatives

Zhenwei Wang

Abstract In order to solve the problem of difficult internal guidance in university libraries, a library lending navigation system based on collaborative filtering and Augmented Reality (AR) technology is designed. A joint collaborative filtering model based on attention mechanism is also proposed to solve the data sparsity problem of library recommendation system. In addition, the study introduces AR technology for book location marking and circulation navigation in order to provide better circulation service experience in library scenarios. The AR plane detection and ray collision detection technologies enable users to locate the exact position of books through their smartphones. Experiments on Movielens and BookCrossing datasets show that the model exhibits excellent recommendation in both top-K Hit Ratio (HR) and Normalized Discounted Cumulative Gain (NDCG) metrics. The results show that the accuracy localization error is higher than that of the normalized discounted cumulative gain (NDCG), with the improvement ranging from 2.6% to 26.4%, respectively. The experimental results show that the maximum accuracy positioning error is only 0.62 m, and in most cases, the positioning accuracy of the fine positioning is higher than that of the coarse positioning. In summary, the joint collaborative filtering model based on attention mechanism and AR navigation technology can effectively improve the efficiency of library lending service and user satisfaction.

Z. Wang (✉)
Guangdong University of Science & Technology, Dongguan 523083, China
e-mail: 276704647@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_17

17.1 Introduction

With the continuous development of digitalization and networking, scenario service has become a way to proactively recommend information resources and services suitable for users' needs. In the library scenario, checking out books is one of the most basic requirements. However, with the increasing number of print and electronic books, it is difficult for readers to find books that meet their needs, which has led to the problem of "information overload" [1–3]. It is an urgent problem for libraries to find out readers' preferences for information resources and to provide them with personalized resources. In addition, this project proposes an augmented reality-based reader browsing service, which can help readers shorten the time to find books and increase their interest in reading [4–6]. However, there is also an urgent need to solve the problem of how to access a large number of books in a library while reducing the cost of locating the user's position in an indoor environment. Therefore, combining machine learning models for personalized book recommendations and low-cost navigation services using computer vision and augmented reality technologies can meet readers' needs for more personalized and intelligent borrowing, which is very important for library information services [7–9].

17.2 Research on Collaborative Filtering AR Lending Navigation Technology

17.2.1 Collaborative Filtering Lending Recommendation Model Design

A universal recommendation system in the library scenario can recommend suitable books based on users' preferences and improve the efficiency and experience of lending services. However, due to the large amount of data, the recommendation system faces the challenge of data sparsity in the initial stage of startup [10, 11]. Therefore, this chapter proposes a joint collaborative filtering model based on the attention mechanism to solve the data sparsity problem. The overall model is shown in Fig. 17.1.

The model in Fig. 17.1 consists of an input layer, an embedding layer, an attention layer, a fusion layer, a convolutional layer, a fully connected layer, and an output layer. The fusion layer mainly consists of a coupling information matrix, while the embedding layer mainly consists of information vectors for users and projects. The recommendation model takes into account the information of display attributes and implicit behavioral information between users and books. In the recommendation model, the information and attributes related to the user and the item, such as the user's occupation or interest, the type of the item, etc., often directly reflect the relevance and characteristics of the user and the item. At the same time, the associated

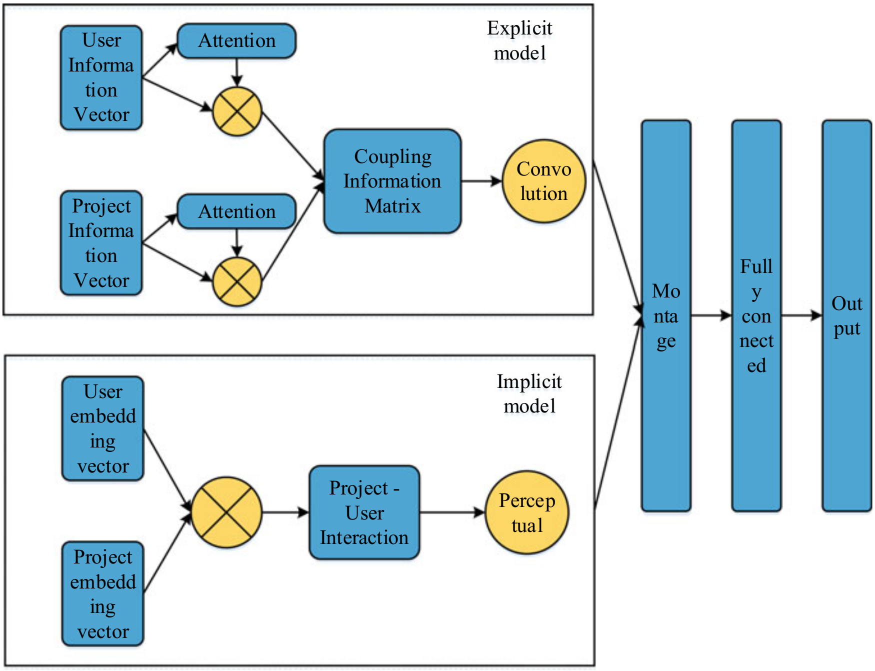


Fig. 17.1 Collaborative filtering borrowing recommendation model

behaviors between users and items, such as users' favorites or shares of items, can also reflect the connection and characteristics between users and items. In the correlation attributes between users and projects, projects with similar attributes can be recommended to users based on the information of users and projects with similar attributes, which is also the basic idea adopted in user-based collaborative filtering and project-based collaborative filtering. The coupled collaborative model further explores the correlation between users and items by calculating the coupled information matrix between user's feature information and item's feature information, so as to accurately recommend items that meet users' needs. For the display attributes of users and books, similar books can be recommended by calculating the similarity. In order to extract the influence of each attribute on user and book preferences, this model introduces an attention mechanism. Specifically, this model employs a network of explicit joint attention mechanisms to mine the explicit attribute features of users and books. The attention extraction formula is shown in Eq. (17.1).

$$\begin{cases} A_{uout} = A_u \cdot v_u \\ A_{iout} = A_i \cdot v_i \end{cases} \tag{17.1}$$

In Eq. (17.1), v_u and v_i represent the information vectors of the user and the project respectively, A_u and A_i represent the corresponding attention weights respectively,

A_{uout} represents the output of the user information vector and A_{iout} represents the output of the project information vector. The coupling matrix is constructed based on the relational information between the user and the project, and the elements of the matrix are calculated using Eq. (17.2).

$$g_{ij} = \frac{1}{\sqrt{A_{uouti} \cdot A_{uoutj} + 1}} \tag{17.2}$$

In Eq. (17.2), i and j represent the vectors of users and items, respectively, and A_{uouti} and A_{uoutj} represent the elements of users and items, respectively. After that, the coupling information matrix is used to extract the coupling information features as shown in Eq. (17.3).

$$h_{conv} = Conv(W_{conv} R_c + b_{conv}) \tag{17.3}$$

In Eq. (17.3), W_{conv} represents the neuron weights, b_{conv} represents the bias value, $Conv$ represents the convolution operation, and h_{conv} the output of the convolution layer. To further optimize the feature learning effect, this model uses a two-layer neural network structure for learning the low-dimensional representations of users and books. Specifically, the attention score w is multiplied by the attribute features and fed into the two-layer neural network to finally obtain the low-dimensional embedding vectors of users and books. In the neural network, the ReLU function is used to improve the fitting and generalization ability of the model. During the learning process, the model uses root mean square error (RMSE) as a loss function to measure the error between predicted ratings and actual ratings. In the prediction process, the predicted rating value is obtained by calculating the dot product between the user's embedding vector and the book's embedding vector. By combining the attention mechanism and the joint collaborative filtering model, the model can improve the recommendation effect in the case of sparse data, thus optimizing the lending recommendation service, satisfying the users' needs, and promoting the motivation of book lending by considering the display attributes and the implicit behavioral information and mining the association features between users and books comprehensively. The study uses word2vec technique to embed user rating information in the implicit Deep CF recommendation model designed to solve the sparse problem generated by the general one-hot coding method. This model constructs a new feature space and achieves more accurate recommendations through deep learning algorithms. Compared with traditional machine learning algorithms, the Deep CF model can improve recommendation accuracy and user experience, and personalize recommendations by learning user behavior patterns. The practical application of this model will help improve the efficiency and performance of recommendation systems. Deep CF uses element-by-element multiplication to generate a linear correlation vector in the operation, as in Eq. (17.4).

$$r = (r_{u1} r_{i1}, r_{u2} r_{i2}, \ldots r_{uk} r_{ik}) \tag{17.4}$$

In Eq. (17.4) the elements of the different embedding vectors are represented by r_{uk}. The user-item relationship obtained after introducing the nonlinear factors is shown in Eq. (17.5).

$$a_L = RELU\left(W_L^T r + b_L\right) \tag{17.5}$$

In Eq. (17.5), W_n and b_n represent the weight and bias values, respectively, and $RELU$ represent the activation function.

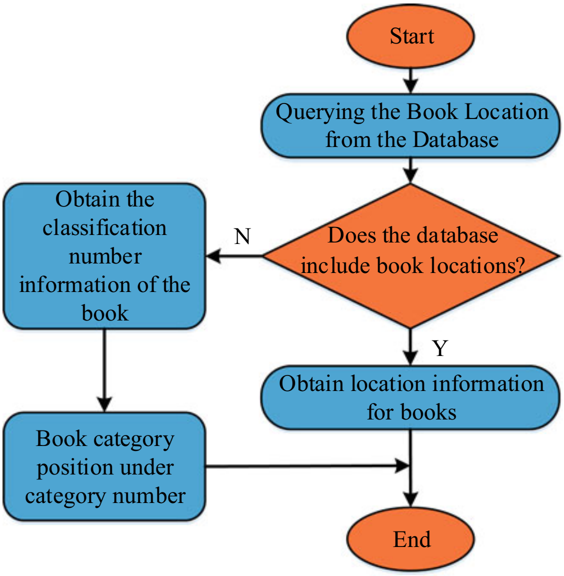
17.2.2 Library AR Circulation Navigation Study

To improve the quality of library services, accurate book location information needs to be collected. The research can use indoor location services and Augmented Reality (AR) technology to obtain the user's location and mark the books to obtain accurate book location information. When the books do not have tagged location information, the study can pre-tag the books with their classification numbers by humans and mark their approximate grid coordinates in the image map. In this way, at the early stage of system startup, users can also be directed to the vicinity of the shelf of the book of that category based on the grid coordinate markers, which makes it easy for them to find it. In addition, when the user uses the book markers, it can facilitate quick location the next time the user wants to look up the book again. Using these virtual rewards of positive feedback, users can be encouraged to collect book location information and share it with everyone to achieve the construction of a service for sharing purposes. Through the above approach, the study can effectively collect and maintain book location information to improve the efficiency of library services and user satisfaction. The location query process is shown in Fig. 17.2.

Using AR Plane Detection and Ray Collision Detection technologies, AR book marking service can be provided in the library scene. First, the AR Plane Manager component is introduced to detect horizontal and vertical planes in the Unity3D engine for visual display on the phone. Then, with the AR Raycast Manager component, rays (with AR Plane Collision Detection) are emitted to obtain the coordinates of a point on a specific plane, thus enabling the AR book marking service. This service enables quick search and finding of related books, making it easier to find the books users need. In addition, users can record the marker location in the form of a virtual bookmark, so that they can quickly locate the marker location when reading next time, thus improving reading efficiency. Therefore, the service can add real-time checkout and return function to AR book marking service to realize digital library self-service. This technology application can significantly improve the quality of library services and user experience, and promote the digitalization process.

The AR-based book navigation method designed in the study can achieve precise location of book marker position and navigation position by positioning the user's smartphone. Before performing the precise location, image feature point matching is required. Generally, matching can be performed using the BF or FLANN algorithms,

Fig. 17.2 Location query process



but both algorithms result in incorrectly matched point pairs. In order to ensure the correctness of the bit pose estimation, it is usually necessary to perform the rejection of the incorrectly matched point pairs. This is where the RANSAC algorithm can be used for image feature point matching. The algorithm calculates certain mathematical model parameters for a series of samples containing erroneous data, and selects the correct samples from them to achieve the rejection of matching pairs. Through the above steps, a more accurate position estimation and displacement vector of the user's cell phone with respect to the grid can be obtained, as well as a precise location of the book marker position and navigation position. Thus, the AR-based book navigation service is realized. First, the library scene needs to be scanned and modeled to get the grid layout and coordinate information of each grid location. Secondly, each book in the library needs to be photographed and scanned with high accuracy to obtain the feature information of the book and store it in the database. When using the AR smartphone to locate the book position, a camera calibration operation and image processing process are required to obtain an estimate of the user's location and orientation. Next, by using this estimated value and the pre-established grid layout information, the current grid location of the user can be determined. With the book feature information stored in the database, the books in the currently located grid can be compared and matched with the user query information. Thus, AR-based book navigation service is realized and the efficiency of library services is improved. In this system, the RANSAC algorithm is used to pair feature points in adjacent images to obtain visual transformations. Through coarse matching and filtering, the study can identify the same feature points in adjacent images. During the screening process, the study uses the Real Sample Consensus (RANSAC) algorithm to further exclude false matches, resulting in the correct camera shift and rotation matrix. The algorithm iteration threshold can be calculated using Eq. (17.6).

$$l_m = \frac{\ln(1 - p)}{\ln(1 - w^a)} \tag{17.6}$$

In Eq. (17.6), w represents the ratio of the number of points in the bureau to the number of data sets, a represents the number of points needed for the calculation, and p represents the probability of getting the correct model. To recover the actual scale, the study uses gridded maps to solve the image scale problem. In each grid, the study has a reference point that can be used as scale information using its grid edge length. After a pairwise geometric calculation, the study can calculate the exact position of the user on the map based on the position and distance of adjacent grids.

17.3 Research on Collaborative Filtered AR Lending Navigation Technology

17.3.1 Analysis of the Effect of Collaborative Filtering Lending Recommendation Model

When analyzing the effect of the Collaborative filtering loan recommendation model, the study uses two main data sets, one is the Movielens data set, and the other is the BookCrossing data set. All experiments were conducted on a computer equipped with an Intel Core i7 processor and 16GB RAM, using software such as Python 3.6 and TensorFlow 1.14.

In order to verify the improvement of image matching speed of the orientation sensor-based image database segmentation method, image matching was performed on 20 test case images, where the longest time, shortest time and average elapsed time required for successful test cases were used as detection metrics. The results are shown in Fig. 17.3.

It can be seen that the time fluctuation of image feature matching from both matching maximum time and matching minimum time in the map image feature library is not significant. When the image feature library is divided by using a directed sensor, the matching time decreases with the increase in the number of divisions. However, as the number of divisions increases, the increase in image feature matching speed becomes smaller at this time. This is because when matching using the feature database, the database read share increases as the matching time decreases, so the improvement is not as significant when the number of directional divisions is 8 compared to the number of divisions is 6. In order to verify the improvement of the enhanced image similarity scoring rule on the image matching accuracy, a comparison was made under the same experimental conditions, using the traditional method with the improved method for image matching, and the number of test images that

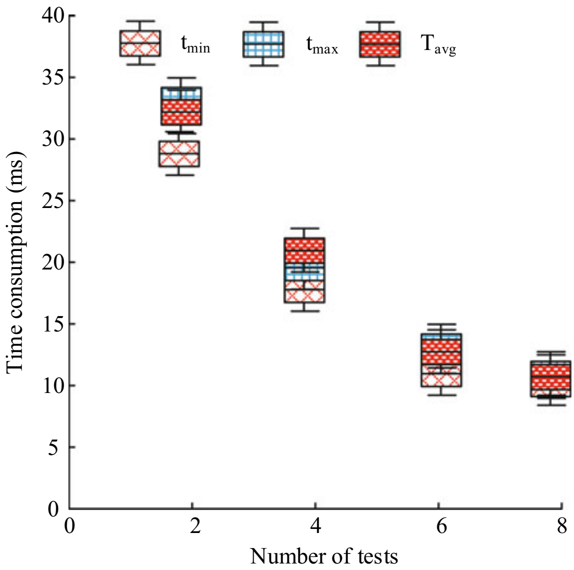


Fig. 17.3 Image matching elapsed time test results

correctly matched the recorded grid coordinates as a percentage of the total number of test images was counted, and the results are shown in Fig. 17.4.

It can be found that the improved method is 6.35% higher in the number of correct matches compared to the original method, which is due to the optimization of the improved algorithm in terms of matching test images in similar scenarios and

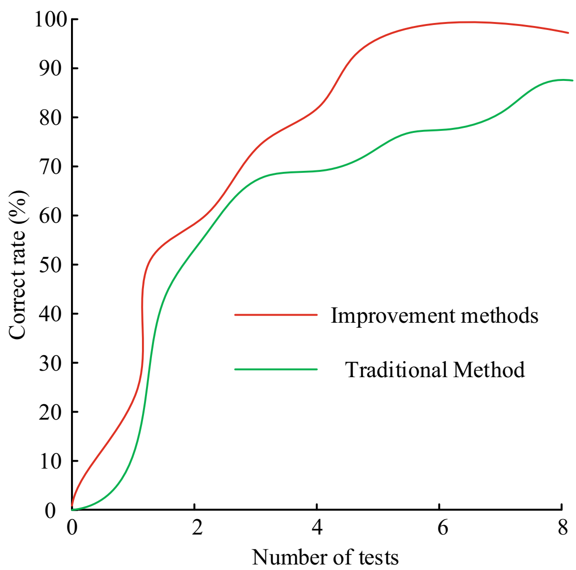


Fig. 17.4 Image matching accuracy test results

achieving certain results to make correct matches. In order to verify the performance of the research algorithm, the study uses two main datasets separately in the analysis of the effect of collaborative filtering lending recommendation model, one is Movielens dataset, which is a standard dataset often used in the field of recommendation systems and contains movie ratings, user information and movie information, etc. The other is the BookCrossing dataset, an online book exchange platform where users can share their unused books with other users for borrowing, which collects many users' rating data and is one of the popular datasets in the field of recommendation systems. The study uses two metrics, top-K Hit Ratio (HR) and Normalized Discounted Cumulative Gain (NDCG), for analysis. Among them, the HR metric indicates the ratio of the number of items actually interacted by users to the total number of items interacted when the algorithm recommends the top-K items to users, which reflects the recall rate of the algorithm. And the NDCG index is a recommendation quality index considering the position ranking, which has a good ability to evaluate the effect of ranking order. Therefore, both metrics have the ability to measure the algorithm's recommendation effect very well. The effect comparison chart is shown in Fig. 17.5.

Contrastive algorithm performance analysis is an important method to evaluate the performance of recommendation algorithms. In the study's evaluation of the Movielens and BookCrossing datasets, the study selected several common recommendation algorithms for comparison, including ItemKNN, BPR, MLP, DeepICF + a, DeepUCF + a, NeuMF, and CoupledCF. The study analyzes and compares HR and NDCG metrics for each algorithm on two datasets. On the Movielens dataset, the HR metric of the studied algorithm was 0.8509 and the NDCG metric was 0.5852 on the Movielens dataset, indicating that the algorithm outperformed all the previous algorithms in terms of effectiveness and significantly outperformed the other algorithms in terms of HR and NDCG metrics. Compared to the closest algorithm, CoupledCF,

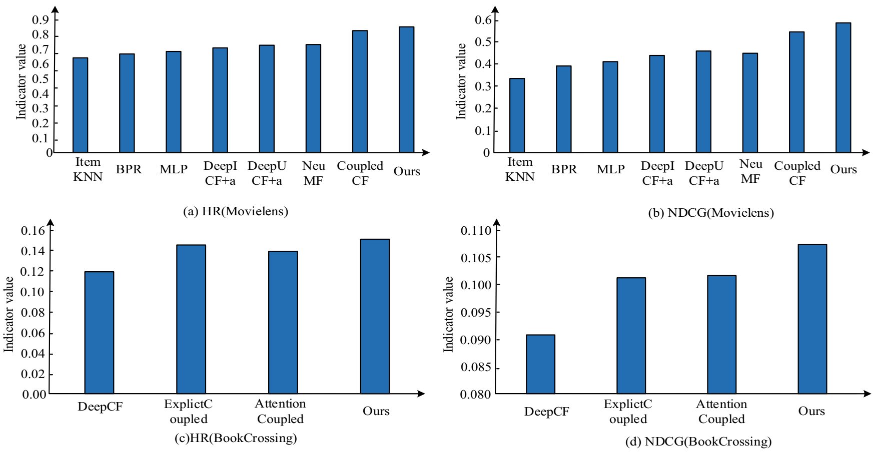


Fig. 17.5 Comparison of effects

the studied algorithm improved the HR metric by 2.6% and the NDCG metric by 8.6%. On the BookCrossing dataset, the studied algorithm also showed excellent recommendation results, outperforming all three other algorithms, improving HR metrics by 26.4% and NDCG metrics by 17.3%. These data suggest that the studied algorithms are better able to predict users' behaviors and preferences, thus providing better recommendations. In addition, better performance is shown on both datasets, indicating that the studied algorithm has good generalization and adaptability. On the BookCrossing dataset, the study also compares the HR and NDCG metrics of different algorithms on this dataset, and it can be seen that the algorithm proposed in the study has better performance on HR and NDCG metrics relative to the other three algorithms, with an improvement of 26.4% and 17.3%, respectively, which indicates that the algorithm of the study can better predict users' preferences and provide them with a better recommendation experience for users. Meanwhile, the performance of the explicitCoupled and attentionCoupled algorithms is also slightly better than that of DeepCF, indicating that both algorithms have some effectiveness. The proposed algorithm performs better than the other three algorithms in HR and NDCG metrics, with an improvement of 26.4% and 17.3%, respectively, which indicates that the studied algorithm can better predict users' preferences and provide better recommendation experience for users. From the results, it can be seen that the algorithms proposed in the study have relatively good performance and can provide better recommendations for the recommender system.

17.3.2 Analysis of the Effect of Lending Navigation

In this section, all experiments were conducted on a computer equipped with an Intel Core i7 processor and 16GB RAM, using software such as Python 3.6 and TensorFlow 1.14. In conducting the analysis part of the lending navigation effect, the study used two perspectives from coarse and fine positioning, and also analyzed the change status of the number of iterations of the model. The specific analysis results are shown in Fig. 17.6.

It can be seen that fine positioning and coarse positioning have somewhat opposite performance in terms of average positioning error. The average error of fine positioning is 0.27 m, which is relatively high, but the maximum error is only 0.62 m, which means that most of the positioning results are relatively accurate, but there are a few cases of large errors. The average error of coarse positioning is 0.31 m, which is relatively low, but the maximum error is 0.54 m, which means that most of the positioning results are relatively accurate, but there are also some cases with large errors. Comparing the CDF plots of precision positioning and coarse positioning, it can be found that the error of precision positioning is smaller in most cases. For example, within the error range of 0.2 m or less, the error probability of precision positioning is higher than that of coarse positioning, indicating that precision positioning is superior in this respect. At the same time, it should be noted that the error CDF curves of fine and coarse positioning are not completely smooth, and sometimes

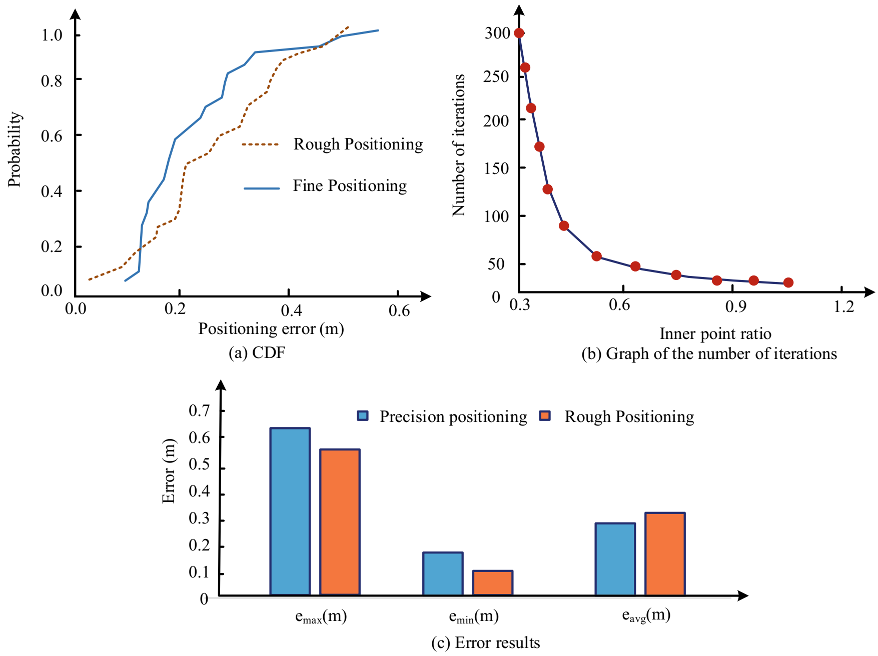


Fig. 17.6 Positioning rendering

there are obvious "sudden change points". This may be due to the existence of some special circumstances or outliers during the experiment, which causes some error measurements to deviate from the normal distribution and thus produce such abrupt changes. Therefore, according to the analysis of the error CDF diagram, the positioning accuracy of fine positioning is higher than that of coarse positioning in most cases. However, it should be noted that there exists a certain degree of instability in the error distribution, which needs to be reasonably selected and carefully weighed according to the actual scenarios and needs.

17.4 Conclusions

A book navigation system that can precisely locate the user's location and book location by coarse and fine localization techniques for more efficient book finding is investigated by combining a joint collaborative filtering model based on attention mechanism and a book navigation service system based on AR technology from the field of library lending service, while using AR technology. According to the results of Movielens and BookCrossing datasets, the collaborative filtering model designed in the study improved by 2.6% and 8.6% in HR and NDCG metrics, respectively, and 26.4% and 17.3% in BookCrossing dataset, respectively, indicating that this model can predict user behavior and preferences more accurately. In addition, the average error of fine positioning of the model is 0.27 m, which is relatively high, but the maximum error is only 0.62 m, which means that most of the positioning results are more accurate, but there are a few cases with large errors. The average error of coarse positioning is 0.31 m, which is relatively low, but the maximum error is 0.54 m, which means that most of the positioning results are relatively accurate. This shows that the model designed by the study has practicality. The model designed in this study is practical and has important value in accurately locating user and book locations to improve book search efficiency. However, further optimization of the model, such as how to reduce positioning errors and how to further improve recommendation accuracy, still requires in-depth exploration in future research. Meanwhile, with the continuous development and popularization of AR technology, how to better integrate AR technology with library service systems is also an important direction for future research.

Acknowledgements This work was financially supported by Guangdong University of Science & Technology (GKY-2021KYWY-38): "Research on the Service of University Library under the Model of Five Education Simultaneously Developing".

References

1. Khaleel, I., Wimmer, B.C., Peterson, G.M., et al.: Health information overload among health consumers: a scoping review. Patient Educ. Couns. **103**(1), 15–32 (2020). https://doi.org/10.1016/j.pec.2019.08.008
2. Phillips-Wren, G., Adya, M.: Decision making under stress: the role of information overload, time pressure, complexity, and uncertainty. J. Decis. Syst. **29**(sup.1), 213–225 (2020). https://doi.org/10.1080/12460125.2020.1768680
3. Al-Kumaim, N.H., Hassan, S.H., Shabbir, M.S., et al.: Exploring the inescapable suffering among postgraduate researchers: information overload perceptions and implications for future research. Int. J. Inf. Commun. Technol. Educ. (IJICTE) **17**(1), 19–41 (2021). https://doi.org/10.4018/IJICTE.2021010102
4. Gao, L., Shukla, S.K.: Book retrieval navigation system based on internet of things technology. In: Cyber Security Intelligence and Analytics: The 4th International Conference on Cyber Security Intelligence and Analytics, vol. 2, pp. 577–584. Springer International Publishing, Cham (2022). https://doi.org/10.1007/978-3-030-97874-7_75
5. Starrett, M.J., McAvan, A.S., Huffman, D.J., et al.: Landmarks: A solution for spatial navigation and memory experiments in virtual reality. Behav. Res. Methods **53**, 1046–1059 (2021). https://doi.org/10.3758/s13428-020-01481-6
6. Li, G., Dong, J., Wang, J., et al.: The clinical application value of mixed-reality-assisted surgical navigation for laparoscopic nephrectomy. Cancer Med. **9**(15), 5480–5489 (2020). https://doi.org/10.1002/cam4.3189
7. Dong, W., Wu, Y., Qin, T., et al.: What is the difference between augmented reality and 2D navigation electronic maps in pedestrian wayfinding? Cartogr. Geogr. Inf. Sci. **48**(3), 225–240 (2021). https://doi.org/10.1080/15230406.2021.1871646
8. Li, G., Dong, J., Wang, J., et al.: The clinical application value of mixed-reality-assisted surgical navigation for laparoscopic nephrectomy. Cancer Med. **9**(15), 5480–5489 (2020). https://doi.org/10.1002/cam4.3189

9. Linxweiler, M., Pillong, L., Kopanja, D., et al.: Augmented reality-enhanced navigation in endoscopic sinus surgery: a prospective, randomized, controlled clinical tria. Laryngoscope Investig. Otolaryngol. **5**(4), 621–629 (2020). https://doi.org/10.1002/lio2.436
10. Abdollahpouri, H., Adomavicius, G., Burke, R., et al.: Multistakeholder recommendation: survey and research directions. User Model. User-Adap. Inter. **30**, 127–158 (2020)
11. Yang, M.: Research on vehicle automatic driving target perception technology based on improved MSRPN algorithm. J. Comput. Cogn. Eng. **1**(3), 147–151 (2022)

Chapter 18
Optimal Scheduling of Urban Logistics Service Supply Chain Based on Artificial Intelligence Algorithm

Lisha Liu

Abstract With the comprehensive popularization of AI, the traditional supply chain logistics field has been unable to meet the social needs, leading to the reduction of industrial operation efficiency. In this article, artificial intelligence (AI) algorithm is considered to realize the intelligent growth of the supply chain logistics industry, the optimal scheduling problem of urban logistics service supply chain is studied and discussed, and a new optimal scheduling algorithm of logistics service supply chain is proposed. In the implementation of the algorithm, the merging position of merging points is deeply studied, and according to the characteristics of merging points and different constraints, they are processed separately, and good optimization results are obtained. The research shows that the algorithm in this article converges quickly, its accuracy rate can reach 95.14%, and its recall rate can reach 94.97%. Moreover, the simulation running time of the algorithm is short and its performance is excellent. It is hoped that this research can provide some theoretical basis and technical support for the optimal scheduling of urban logistics service supply chain, and provide a reliable basis for staff to carry out their work, so as to promote the further growth of urban logistics service supply chain.

18.1 Introduction

The advantages of AI can simulate human thinking. AI itself has the ability of human judgment and reaction, and the AI system is maintained by computer programs and related hardware devices [1]. As the application of AI is becoming more and more popular, people's daily lives and work are more dependent on the choice and decision of data information algorithms and intelligent systems [2]. For the strong AI or even super AI being developed, they are beyond human wisdom [3]. In the material transportation planning and logistics supply chain, people have high hopes for AI.

L. Liu (✉)
Shanghai Institute of Technology, Shanghai, China
e-mail: 78249691@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_18

The advantage of AI in the field of logistics is that it can always accurately predict the repeated events with long or short intervals [4]. It also includes the importance of the strike of air traffic controllers and dockers on air transport and shipping. Supply chain is an important link that closely connects the supply and demand sides, and the rationality of supply chain structure affects the efficiency and cost of supply chain operation [5, 6]. It is of great significance for promoting the growth of modern logistics industry to dig deep into the core value of AI and then improve the intelligent level of modern supply chain logistics.

In order to promote the transformation and innovation of supply chain logistics, improve the efficiency of industrial operation and meet the needs of social growth, it is needed to form a correct understanding of AI and make scientific use of it [7, 8]. In AI, through the open source of related algorithms and the characteristics of big data internet, it can embody brand-new solutions, complete standardization and high efficiency in the whole supply chain logistics, and become a brand-new development source [9]. AI can be combined with supply chain to provide valuable information for human beings and simplify the decision-making about the logistics process. This AI solution can not only predict the direct impact of delayed supply, but also analyze the impact on other stages in the supply chain. In addition, enterprises can reasonably apply the multi-agent collaborative supply chain logistics management system based on AI to their daily work in combination with the needs of multi-agent management, so as to comprehensively improve the efficiency and quality of management. This can promote the common growth of multi-agents while exerting the role of supply chain logistics management. At present, many supply chain logistics enterprises have applied AI, and built an intelligent chain system and logistics nodes on the logistics platform. This article discusses the application of AI in the optimal scheduling of urban logistics service supply chain through relevant investigation and analysis of literature, and puts forward a new optimal scheduling algorithm of logistics service supply chain. In order to provide a reliable basis for staff to carry out their work.

18.2 Methodology

18.2.1 Application of AI in Logistics Supply Chain

In the AI era, various intelligent technologies are widely used in all aspects of supply chain logistics. In order to ensure that the logistics system of AI supply chain can be fully developed and effectively applied, it is needed to analyze its urgent improvement and find out the improvement direction to ensure its more effective development [10]. With the promotion of AI, the logistics industry began to appear smart supply chain. The advantage of AI is that it can integrate different data sources in the supply chain logistics system. For example, in the transportation of goods, it can complete the screening of time information according to the overall delivery time data and transportation data [11]. With the flexible use of AI, the security of inventory management will be significantly improved, which can guarantee the overall benefit of enterprises. Urban supply chain presents different temporal and spatial characteristics due to the change in consumption, production and circulation mode, which requires the reorganization of time and space resources in the process of supply chain, that is, the adjustment of time and space allocated to activities such as warehousing, transportation and distribution, and thus produces multi-level and multi-dimensional pressure on urban space–time system. AI supply chain logistics has unique advantages, which can be divided into the following three advantages: basic advantages, technical advantages and application advantages. At present, the growth of AI in the field of supply chain logistics is unstoppable. In this article, the AI algorithm is integrated to achieve the intelligent growth of supply chain logistics industry, the optimal scheduling problem of urban logistics service supply chain is studied and discussed, and a new optimal scheduling algorithm of logistics service supply chain is proposed.

18.2.2 Optimal Scheduling of Urban Logistics Service Supply Chain

In the supply chain logistics management system, enterprises can input information such as capital information, business scope, business qualification, address information, enterprise name, etc., and can also use AI to design barcode labels for different products, and then input barcodes into the system to lay the foundation for data information management. For logistics enterprises, enterprises can conform to the development trends of all walks of life and vigorously promote the whole logistics infrastructure in order to realize intelligent transformation and ensure that their supply chain logistics system can be built in an intelligent direction. Generally, the urban logistics spatial system is divided into three levels according to the hierarchy: the logistics portal area at the city level, the distribution center at the district level and the urban distribution station at the community level. In terms of intelligent transportation and distribution, it is mainly manifested in two aspects, namely distribution equipment and transportation routes. In practical work, AI can make scientific use of scheduling algorithm and path optimization algorithm, and fully combine the data information of data center to carry out dynamic route planning to ensure that the transportation route is reasonable and scientific. In addition, intelligent distribution equipment will comprehensively consider the planned transportation routes, so as to implement automatic distribution and comprehensively improve distribution efficiency. The operation mode diagram of the supply chain is shown in Fig. 18.1.

It is an effective way to reduce the distribution cost by reducing the number of trains and the total distance of transportation as much as possible on the premise of ensuring that the goods arrive at the place required by customers on time. Each target point corresponds to a unique delivery vehicle, and it is needed to ensure

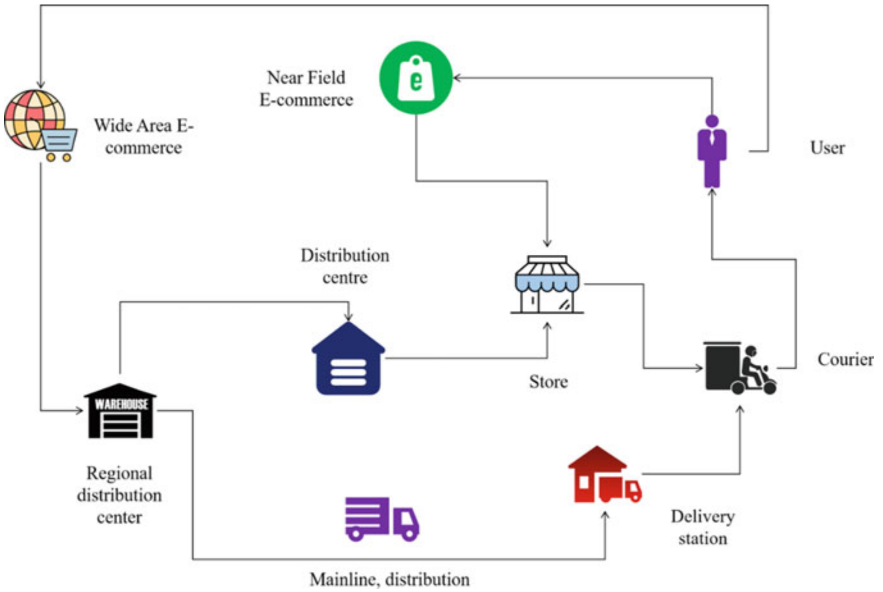


Fig. 18.1 Operation mode diagram of supply chain

that the driving of delivery vehicles between target points cannot be repeated, that is, each target point only corresponds to one delivery service of a delivery vehicle. Establish a single-objective optimization model with the objective of minimizing the total vehicle transportation cost:

$$Z = \min \sum_{i} \sum_{j} \sum_{k} C_{ij} x_{ijk} (i \neq j) \tag{18.1}$$

where Z stands for the total cost; i, j represent any two points; k represents any transport vehicle; C_{ij} represents the transportation cost from point i to point j.

Considering the actual situation, three optimization objectives are put forward. That is, the shortest delivery time, the shortest driving distance of all delivery vehicles, and the least cost:

$$\begin{cases} \min Z_1 = \sum_{i} \sum_{j} \sum_{k} T_{ij} x_{ijk} (\text{The shortest time}) \\ \min Z_2 = \sum_{i} \sum_{j} \sum_{k} D_{ij} x_{ijk} (\text{The shortest distance}) \\ \min Z_3 = \sum_{i} \sum_{j} \sum_{k} C_{ij} x_{ijk} (\text{The least cost}) \end{cases} \tag{18.2}$$

Population renewal is a problem in deciding how to preserve the new population produced by the selection, crossover and mutation operators of a population as a new population to prepare for the next reproduction. If none of the offspring chromosomes can be more adaptive than the one with the greatest fitness in the parent, the fitness of the offspring chromosomes will be degraded. The optimization scheduling of urban logistics service supply chain carries out AI construction and innovation, and reduces the intermediate logistics links as much as possible, which makes the logistics transportation efficiency of the supply chain higher, and can also ensure the accuracy. In intelligent inventory management, intelligent inventory management can upgrade the whole management method based on cost management or rely on manual electronic file management to break away from the conventional workflow. Real-time dynamic management is completed in the mechanism of inventory, inventory type, article storage and market. Transform traditional inventory management into intelligent inventory management, and combine Internet technology and visualization technology to make the whole warehouse data fast.

18.3 Result Analysis and Discussion

In the application of AI in the field of logistics chain, as the most important technical resource, AI drives the efficiency of the whole supply chain to ensure that it can become a powerful boost for the subsequent transformation of logistics industry chain. Therefore, relevant personnel must form a correct understanding of AI and make scientific use of this technical means to ensure the effective combination of AI and supply chain logistics. Based on this, the logistics industry will achieve high-quality development. In this article, a mathematical model of supply chain logistics distribution path with constraints is established by combining AI algorithm, and a new optimal scheduling algorithm of logistics service supply chain is proposed. In order to verify the effectiveness of the algorithm, this section carries out simulation experiments and analyzes the experimental results. In order to improve the efficiency of logistics operation, and at the same time provide a higher level of logistics services for merchants in the supply chain, so as to promote the growth of modern supply chain logistics in the direction of intelligence and more efficiency. Before starting the experiment, it is needed to preprocess the initial data. The algorithm in this paper is mainly based on Tensor Flow platform, and NVIDIA-2080Ti GPU graphics card is used in hardware equipment. The use of GPU is helpful to accelerate the training of neural network. The number of rounds of training is set to 70, the number of cycles of each round of generating network is 100, and the batch size is 40. The environmental configuration parameter requirements of batch system are shown in Table 18.1.

In this section, firstly, the algorithm is trained and tested, and the preprocessed data is used for simulation, and the convergence of the algorithm is shown in Fig. 18.2.

It can be seen that the convergence of the proposed algorithm is better. Using the optimal scheduling algorithm of logistics service supply chain proposed in this article can solve some requirements put forward in practical cases. For example, according

Table 18.1 Requirements for environmental configuration parameters of the system

Project	Version
Operating system	Windows 11
CPU	Intel(R) Core(TM) I7-13700K
Internal storage	16GB
Hard disc	1TB
GPU	RTX 2080Ti
Memory	11G
DL framework	TensorFlow 2.6
Database administration	Navicat for SQLite
Compiler	Python 3.8
Interface development	Qt designer

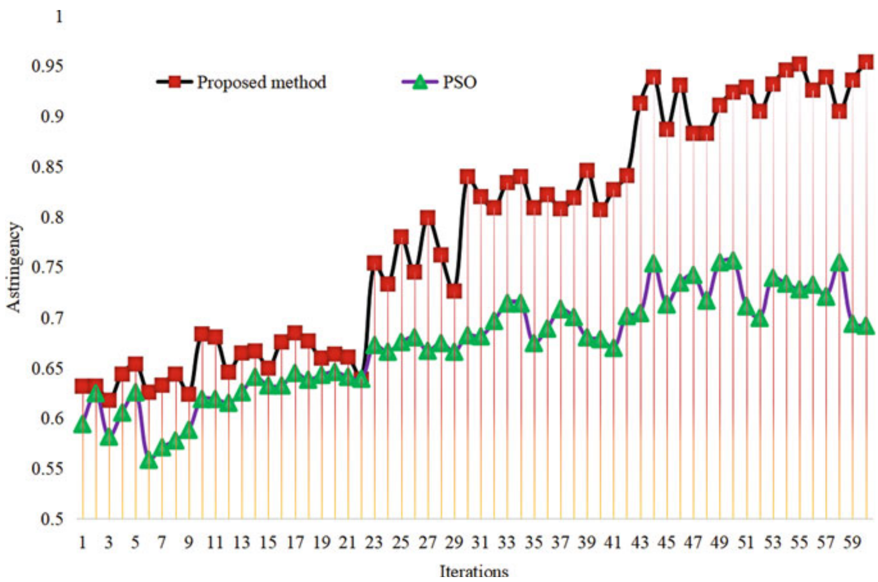


Fig. 18.2 Convergence of the algorithm

to the demand freight volume, the delivery is integrated and carried out at the same time; The mileage is as small as possible, which has better results compared with traditional methods; meets the time window constraint. The algorithm is reasonable and feasible. In order to further verify the performance of the analysis algorithm, this article records the accuracy rate and the recall rate. The formulas defined by these two indicators are as follows:

$$P = TP/(TP + FP) \tag{18.3}$$

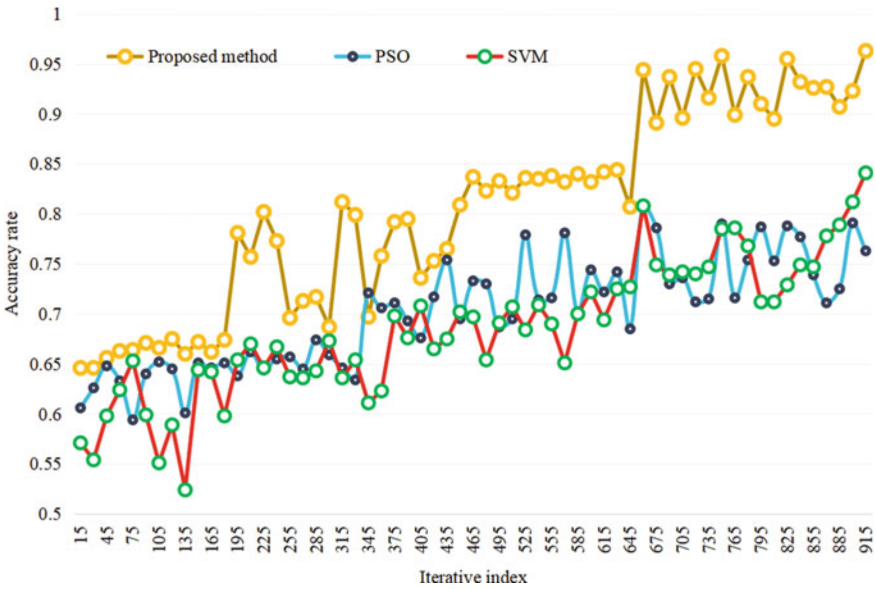


Fig. 18.3 Accuracy of the algorithm

$$R = TP/(TP + FN) \tag{18.4}$$

Through experiments, the comparison of the accuracy of several algorithms is shown in Fig. 18.3. The recall rate of several algorithms is shown in Fig. 18.4.

Generally speaking, the shorter the path of logistics distribution in the supply chain, the lower the distribution cost and the shorter the distribution time. Therefore, the shortest path is achieved and the cost and efficiency of the supply chain are controlled. The proposed algorithm determines to add a point that is not on the line to the line every time according to a certain discriminant function, so as to meet the constraint requirements and get a satisfactory distribution line until all the points are arranged in the line. In addition, the factors considered in the traditional saving method are single, ignoring the customer's requirements for delivery time. Based on this deficiency, this article adds the restriction condition of time window to the traditional saving method. This method not only meets the customer's requirements for delivery time, but also calculates the vehicle departure time window and organizes the smooth implementation of delivery work. The comparison of simulation running time of the algorithm is shown in Fig. 18.5.

The experimental results in this section show that the algorithm converges quickly, and its accuracy rate can reach 95.14% and recall rate can reach 94.97%. The algorithm realizes multi-objective and multi-constraint scheduling optimization and obtains good optimization results. And the simulation running time of the algorithm is short, which shows that its performance is better. The optimized algorithm can quickly and accurately provide a scientific and accurate logistics distribution path

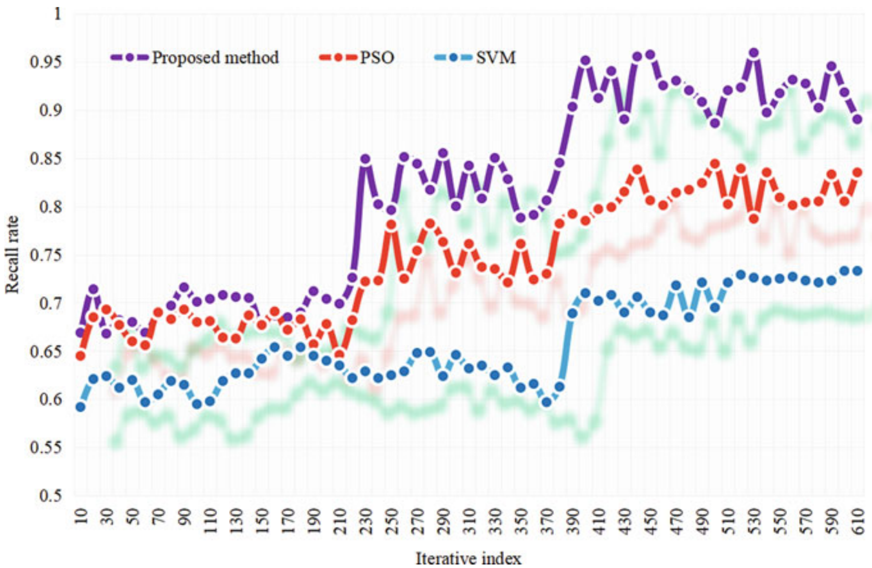


Fig. 18.4 Recall rate of the algorithm

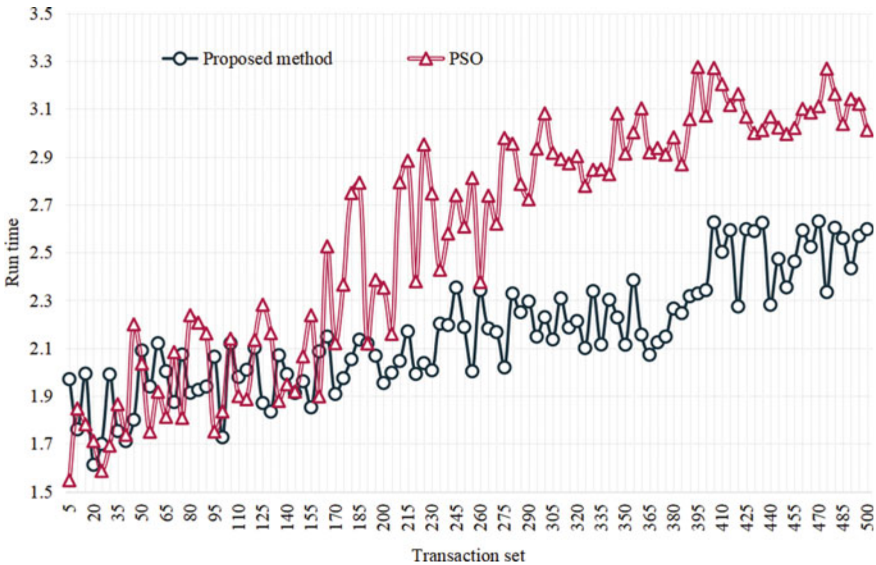


Fig. 18.5 Comparison of simulation running time of algorithms

planning scheme for the supply chain, and the supply chain can realize a higher degree of overall planning for the distribution system.

Intelligent development is the inevitable trend of the future growth of supply chain logistics, which can significantly improve the efficiency of industrial operation, reduce costs and meet the needs of the times. This method improves the efficiency of supply chain logistics transportation, and can promote the mutual integration and coordinated growth of AI and supply chain logistics. All complete information chains can form an information network, from which historical data required by supply chain logistics can be obtained. When there is a problem in the process of logistics transportation, the function of intelligent traceability can be used to quickly determine the cause of the problem and the person directly responsible, so as to solve the problem, thus improving the efficiency of supply chain logistics and ensuring the safety of all aspects of logistics transportation.

18.4 Conclusion

Logistics transportation system is playing an increasingly important role in modern production and life. The development degree of logistics industry has become one of the important symbols to measure the modernization degree of a city or a country. In the logistics system, the supply chain is an important link that closely connects the supply and demand sides, and the rationality of the supply chain structure affects the efficiency and cost of the supply chain operation. In this article, the optimal scheduling problem of urban logistics service supply chain is studied by combining AI algorithm. In this article, a mathematical model of supply chain logistics distribution path with constraints is established, and a new optimal scheduling algorithm of logistics service supply chain is proposed. Through the supply chain optimization scheduling algorithm, the overall optimal solution of logistics distribution path is solved, and finally, the performance of the algorithm is verified by simulation experiments. A number of simulation studies show that this algorithm converges quickly, and its accuracy rate can reach 95.14% and recall rate can reach 94.97%. Moreover, the simulation running time of this algorithm is short, and its performance is excellent. The algorithm realizes multi-objective and multi-constraint scheduling optimization, and obtains good optimization results. Generally speaking, this research can provide some theoretical basis and technical support for the optimal scheduling of urban logistics service supply chain; Give full play to the advantages of AI and improve the efficiency of supply chain logistics and transportation. In order to promote the modern supply chain logistics to develop more efficiently in the direction of intelligence.

Although the VRP problem of fixed starting and ending points proposed in this paper is common in daily life, this paper only studies the constraint condition of fixed starting and ending points, and does not involve the conditions such as the number of vehicles, load capacity, cargo demand and time constraint, so it needs further study.

References

1. Zhao, M., Liu, X.: Growth of decision support tool for optimizing urban emergency rescue facility locations to improve humanitarian logistics management. Saf. Sci. **102**(11), 110–117 (2018)
2. Su, Y., Qin, J., Yang, P., et al.: A supply chain-logistics super-network equilibrium model for urban logistics facility network optimization. Math. Probl. Eng. **2019**(16), 1–12 (2019)
3. Bu, S.: Logistics engineering optimization based on machine learning and artificial intelligence technology. J. Intell. Fuzzy Syst. **40**(3), 1–12 (2020)
4. Yáez-Sandivari, L., Cortés, C.E., Rey, P.A.: Humanitarian logistics and emergencies management: New perspectives to a sociotechnical problem and its optimization approach management. Int. J. Disaster Risk Reduct. **52**(1), 101952 (2020)
5. Akkad, M.Z., Nyai, T.B.: Multi-objective approach for optimization of city logistics considering energy efficiency. Sustainability **12**(18), 7366 (2020)
6. Kuznietsov, K.A., Gromov, V.A., Skorohod, V.A.: Cluster-based supply chain logistics: a case study of a Ukrainian food distributor. IMA J. Manag. Math. **28**(4), 553–578 (2017)
7. Shankar, R., Bhattacharyya, S., Choudhary, A.: A decision model for a strategic closed-loop supply chain to reclaim end-of-life vehicles. Int. J. Prod. Econ. **195**(1), 273–286 (2017)
8. Nz, A., Hs, B., Mg, C., et al.: A novel multi-objective model for green forward and reverse logistics network design. J. Clean. Prod. **208**(10), 1304–1316 (2019)
9. Hu, H., Lin, T., Wang, S., et al.: A cyberGIS approach to uncertainty and sensitivity analysis in biomass supply chain optimization. Appl. Energy **203**(10), 26–40 (2017)
10. Shi, Y., Huang, Y., Xu, J.: Technological paradigm-based construction and demolition waste supply chain optimization with carbon policy. J. Clean. Prod. **277**(9), 123331 (2020)
11. Guo, X., Voogt, J., Annevelink, B., et al.: Optimizing resource utilization in biomass supply chains by creating integrated biomass logistics centers. Energies **13**(22), 6153 (2020)

Chapter 19
Construction of “Double-Position” Teachers’ Teaching Ability Evaluation System in Higher Vocational Colleges Based on Big Data Analysis

Xiaoting Sun and Wenbo Peng

Abstract “Double-position” teachers, as an important part of HVC (Higher Vocal Colleges) teaching staff, scientific performance evaluation of “double-position” teachers directly affects the quality of the whole teaching staff construction in HVC. It is of great significance to scientifically evaluate the teaching ability of “double-position” teachers for subsequent targeted training and comprehensively improving the quality of HVC personnel training. In order to improve the accuracy of the evaluation of teaching ability of “double-position” teachers, this paper puts forward the evaluation method of teaching ability of “double-position” teachers based on big data on the basis of constructing the evaluation system of teaching ability of “double-position” teachers in HVC. The evaluation results of the proposed method are reliable and can reflect the real ability of teachers. The evaluation results are in good agreement with the real situation of “double-position” teachers’ teaching ability. Through quantitative assessment, “double-position” teachers can be encouraged to keep learning and actively participate in teaching and scientific research, thus promoting the development, promotion and play of double-position teachers’ ability.

19.1 Introduction

In recent years, many documents issued by the Ministry of Education have mentioned the requirement of training “double-position” teachers. The evaluation of teachers’ teaching ability is the core category of China’s vocational education evaluation system, and it is also an important work that vocational education theory and practice pay close attention to. However, the related research on the connotation and composition of vocational education teachers at home and abroad is relatively lacking at

X. Sun (✉) · W. Peng
Krirk University, Bangkok, Thailand

W. Peng
e-mail: 1379397871@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_19

present [1, 2]. The teaching ability of “double-position” teachers in HVC (Higher vocational colleges) directly affects the training quality of high-skilled and practical talents, and then affects the recognition of HVC graduates by industry enterprises and society.

“Double-position” teachers, as an important part of the HVC teaching staff, scientific performance evaluation of “double-position” teachers directly affects the quality of the whole HVC teaching staff construction. However, at present, there are great differences in the understanding of the evaluation standard of “double-position” teachers, and there is no consensus, and there is no complete quantitative evaluation system [3]. How to use big data technology to objectively, comprehensively, timely and accurately evaluate the teaching situation of HVC teachers has become an important link to promote the development of teachers’ teaching; How to build an evaluation system of HVC teachers’ teaching development based on big data technology analysis has become an important content to scientifically establish an evaluation system of teachers’ teaching development and improve teaching quality.

At present, big data has become an important strategic resource and has an important impact on global production, circulation, distribution, consumption activities and economic operation mechanism [4]. It is of great significance to scientifically evaluate the teaching ability of “double-position” teachers for subsequent targeted training and comprehensively improving the quality of HVC personnel training. In order to improve the accuracy of the evaluation of teaching ability of “double-position” teachers, this paper puts forward the evaluation method of teaching ability of “double-position” teachers based on big data on the basis of constructing the evaluation system of teaching ability of “double-position” teachers in HVC.

19.2 Research Method

19.2.1 Principles for the Construction of Evaluation System

The “double-position” teachers in this study refer to teachers who have both theoretical knowledge and practical ability in the subject or field they are teaching. In this study, the “double-position” type is no longer determined by whether it has a certificate or not, and it is required to be recognized as a “double-position” type by actually “possessing” theoretical knowledge and practical ability. For the latter type, the “double-position” type in this study locates theoretical knowledge and practical ability in the disciplines or fields taught, so as to facilitate the grasp of standards and indicators [5, 6]. This makes it possible that the theoretical knowledge base of the “double-position” type as mentioned above must be the theoretical knowledge base of a specific discipline or field, rather than a comprehensive or comprehensive theoretical knowledge base, which makes the theoretical knowledge or practical ability specific to a certain discipline or field, avoids vagueness and abstraction, and facilitates the operation and implementation at the practical level.

The principle of combining science with foresight. Scientifically, we should grasp the principles of personality, interests, hobbies and the development of teaching ability of "double-position" teachers in HVC, respect the facts, and combine the development reality of HVC at the same time. It is necessary to measure teachers' teaching development ability from the perspective of long-term development [7]. Only by combining scientificity with foresight can we formulate the testing standards for the development of HVC "double-position" teachers' teaching ability more objectively and scientifically.

The principle of combining systematicness with dynamics [8]. The systematic principle is to stand in the overall perspective of the development of teaching ability of HVC "double-position" teachers, not only to consider the testing standards from multiple levels and dimensions, but also to pay attention to the internal relationship between various evaluation indicators and standards. Only by comprehensive and systematic evaluation can it be fair and credible. Only by combining systematicness with dynamics can we formulate a more comprehensive, fair and scientific testing standard.

19.2.2 Construction of "Double-Position" Teachers' Teaching Ability Evaluation System

Compared with basic education and ordinary education, vocational education has its own unique law of development. Compared with basic education teachers and ordinary college teachers, "double-position" teachers in vocational colleges face unique training objects, working environment and their own development paths. Judging from the teaching ability of "double-position" teachers in vocational colleges, its dynamic characteristics make it possible and space for development. Its practical characteristics determine that it is difficult for "double-position" teachers to develop their teaching ability only at the theoretical level, and they must be demonstrated, developed and promoted with the help of teaching practice activities. Its systematic characteristics make the teaching ability of "double-position" teachers present in a systematic way instead of just one or one link.

Teachers' professional development theory regards teachers as a profession, and holds that teachers' professional development is characterized by persistence, stages, diversification and initiative. Based on the scientific performance evaluation system, determining the salary, welfare, appointment and promotion rules of "double-position" teachers will help to mobilize the enthusiasm of "double-position" teachers for self-improvement, attract more high-quality teachers and comprehensively improve the level of "double-position" teachers in HVC. At present, the number of teachers in the establishment of HVC is small, and the proportion of "double-position" teachers in the establishment is obviously small, which is not conducive to stimulating the enthusiasm of "double-position" teachers for the development of teaching ability. This restricts the development of "double-position" teachers themselves.

Combine indicators with the same connotation, and exclude indicators that do not meet the evaluation principles or evaluate non-teaching ability. Modify the description of some evaluation standards to make them more operable. This study mainly adopts the revised Delphi method to consult experts, aiming at collecting and providing information in a specific field in order to obtain the consensus of experts in this field [9, 10]. In view of the cross-border nature of this study, 30 experts with high academic attainments, macro-vision and foresight in the fields of vocational education theory and practice, information education and teacher development were invited. In order to ensure the heterogeneity of members, the gender, education background, professional title and teaching experience of experts are also taken into account.

Combined with the core competence and professional characteristics of HVC "double-position" teachers, the evaluation standard of HVC "double-position" teachers was established [11]. The standard includes four aspects: basic qualifications, teaching ability (including theoretical teaching ability and practical teaching ability), teaching reform and scientific research ability, and professional curriculum construction ability. A score of 60 is a pass, a score of 70 is a primary, a score of 80 is an intermediate, and a score of 90 or more is advanced.

In order to improve the reliability and validity of the evaluation, according to the basic principles of modeling, taking into account the basic requirements of the model, such as comprehensiveness, generality, simplicity and operability [12], the project team applied big data analysis to mine nearly 1,000 related documents, and obtained 56 categories of influencing factors, with a total of 38 influencing factors. Based on these influencing factors, four types of questionnaires were designed for teaching managers, professional leaders, teachers and students. 1000 questionnaires were sent to HVC in the province, and 927 valid questionnaires were recovered. Through analysis and statistics, the original six categories were divided into 16 factors according to significance. The specific evaluation index system of "double-position" teachers' teaching ability is shown in Fig. 19.1.

"Double-position" teachers' teaching ability includes professional theoretical knowledge, teaching practice ability and teaching technical ability. Professional theoretical knowledge mainly emphasizes teachers' mastery of professional theory and whether they can teach professional theory reasonably in class. Teaching practice ability can be evaluated by observing the interaction between teachers and students, the choice of teaching strategies and classroom rules. Teaching technology mainly refers to the ability of teachers to use modern information technology to assist teaching, and usually examines the ability of teachers to use information technology to promote students' learning after teaching and class. Moral education infiltration teaching refers to the cultivation of students' good conduct and behavior habits in the process of professional teaching, which can usually be observed from the aspects of teachers' requirements for students' behavior norms and the teaching method of combining arts and sciences.

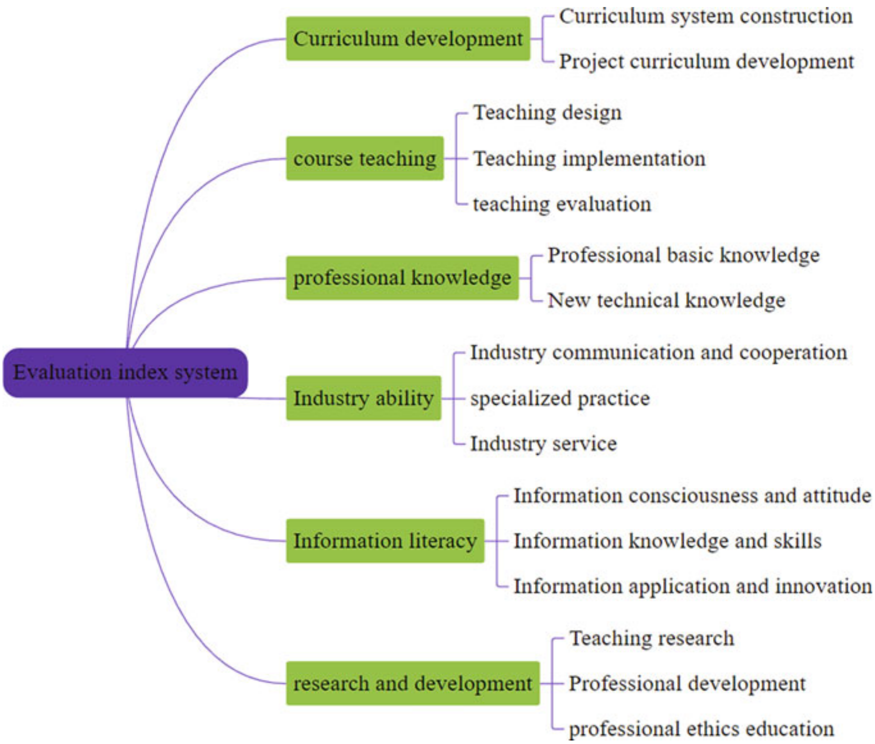


Fig. 19.1 Evaluation index system of "double-position" teachers' teaching ability

In their own development, teachers' personal development goals are still unclear. Reasonable goals are an important driving force to promote their own professional development. But at present, many teachers have not designed proper development tasks in their own development, which makes their own development blind. Based on the theory of teachers' professional development, perfect system design is an important guarantee for the development of "double-position" teachers' teaching ability. Combine the overall interests of colleges and universities with the personal interests of teachers, and give "double-position" teachers support in project declaration, salary and welfare. At the same time, the classification evaluation system should be gradually implemented, and teachers should be evaluated according to the characteristics and needs of different majors and jobs, so as to effectively mobilize the enthusiasm of teachers to improve their professional ability.

19.3 Application of "Double-Position" Teaching Ability Evaluation System

Quantitative analysis and qualitative analysis are not only the most common research methods in scientific research, but also practical means to detect the development of teaching ability of HVC "double-position" teachers. Quantitative analysis is generally observable, measurable, and objective with data and facts, while qualitative analysis is generally characterized by generalization and subjective speculation. Both quantitative analysis and qualitative analysis have their own shortcomings and defects, and if they are organically combined, an ideal evaluation effect can be achieved.

With the general attention paid by managers at all levels in HVC to the development of teachers' teaching and the massive accumulation of teaching big data, it has become an important part to explore how to establish a developmental evaluation system of teachers' teaching based on big data, which comprehensively and accurately reflects the teaching status of teachers, promotes teachers' development and improves teaching quality. Data integration and feedback can realize immediate feedback according to the pre-designed function of the system, with less human intervention. Using the "step by step" system, the evaluation management department needs to manually summarize, classify, filter and integrate the data before feeding it back, and the teaching evaluation department needs someone to be responsible for it. Of course, the data feedback of teaching analysis and evaluation results should be effective, so that managers and teachers can find problems in teaching in real time and adjust and improve them in time.

In order to comprehensively and objectively reflect the influence of each evaluation index on the teaching ability of "double-position" teachers and improve the reliability and validity of the evaluation results, the project team adopted a fuzzy comprehensive evaluation method combining fuzzy set theory and analytic hierarchy process.

$$\begin{gathered}\sum_{i=1}^{6} r_i = 1(i = 1, \cdots, 6) \\ r_i = \sum_{j=1}^{m_i} r_{ij}\end{gathered} \tag{19.1}$$

where r_{ij} is the weight of the j index of the i component. $i = 1, \cdots, 6$ respectively represent the six evaluation indexes in Fig. 19.1.

In order to reduce the amount of calculation, this section carries out rough and intensive processing on the teaching data of "double-position" teachers. Use the following formula to describe the "double-position" teacher teaching information system.

$$R = \{S, B = D \cup E, H, f\} \tag{19.2}$$

Among them, S represents a set of finite objects, that is, the universe, D represents a set of conditional attributes, and E represents a set of decision attributes. If $D \cup E$ is an empty set, R is considered as a decision system, and H represents the range of D and E; $f(a, b) \in H_d, \forall d \in B, a \in S_.$

Evaluate the teaching ability index of "double-position" teachers, and the greater the evaluation index value, the stronger the teaching ability of "double-position" teachers. The calculation formula of evaluation index value in this section is as follows:

$$X_i = \sum_{j=1}^{q} Q_j * P_{ij} \tag{19.3}$$

where: X_i is used to describe the i-th "double-position" teacher's teaching ability evaluation value; Q_j is used to describe the comprehensive weight of the j index; P_{ij} is used to describe the evaluation value of the j index of the i th "double-position" teacher after normalization; q is used to describe the number of double-position teachers.

According to the "double-position" evaluation index system of teachers' teaching ability, the research group designed four types of questionnaires: teaching managers, professional leaders, teachers and students, sent 1000 questionnaires to HVC in the province, and recovered 927 valid questionnaires. Through analysis and statistics, the original six categories were divided into 16 factors according to significance. Applying the statistical results to the established model, the scores of "double-position" teachers' teaching ability components are obtained as shown in Fig. 19.2.

The closer to 1, the better the performance of teachers in their constituent elements. Intuitively, the evaluation results show that there are three items in the [4, 5] interval:

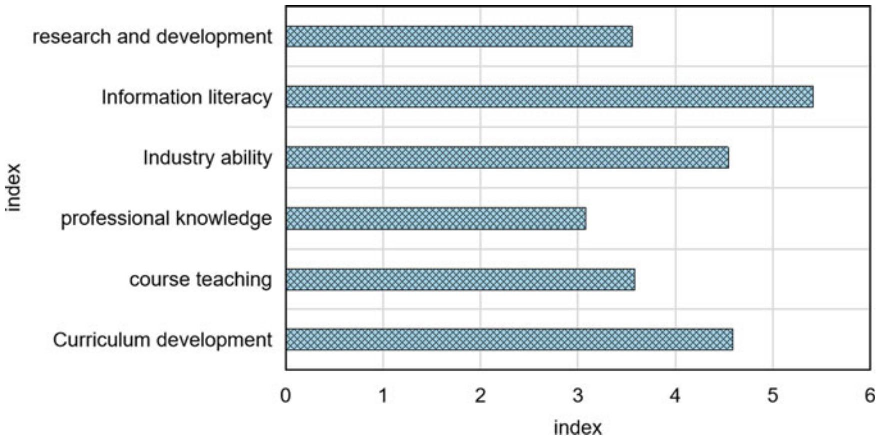


Fig. 19.2 The "double-position" teacher's teaching ability component scores

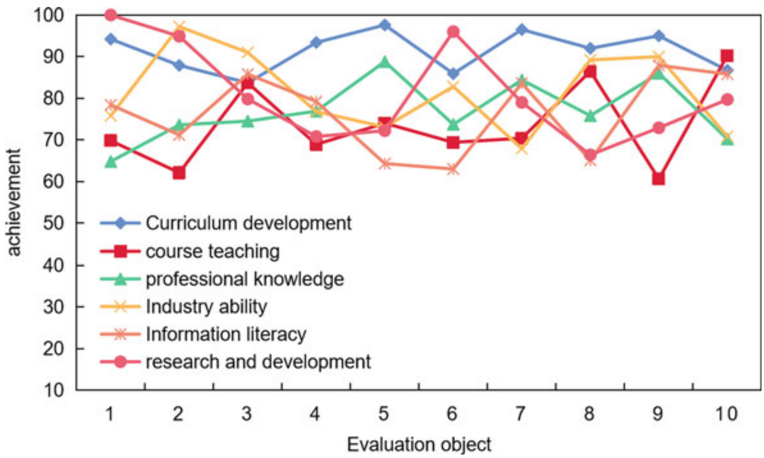


Fig. 19.3 "Double-position" evaluation results of teachers' teaching ability

currency development, Industry ability and Information literacy; There are three items in [1, 3] interval: from high to low, the order is course teaching, research and development and professional knowledge.

According to the weights and evaluation data, the evaluation results of the teaching ability of 10 "double-position" teachers are obtained, as shown in Fig. 19.3.

Teachers with more than 90 points account for 20%, teachers with more than 80 points and less than 90 points account for 30%, teachers with more than 70 points and less than 80 points account for 40%, and teachers with less than 70 points account for 10%. The evaluation results of the proposed method are reliable and can reflect the real ability of teachers.

The ability of teaching reform and innovation is the core ability to measure the development effect of HVC "double-position" teachers' teaching ability. It mainly includes the ability of teaching understanding, teaching design, teaching organization (implementation), teaching practice and teaching research. From the perspective of the research and development of teaching design ability, it was originally a professional skill possessed by educational technicians. With the development of learning theory, teaching theory and new technology, the elements of teaching structure have undergone earth-shaking changes. Under the new teaching environment, all teachers, including HVC teachers, are required to have the ability of teaching design. The reason is that the objectives advocated by basic education are comprehensive and operable, so it is suitable for the formulation of HVC curriculum objectives.

The construction of incentive mechanism for teachers' teaching development is the key link of system construction. Teaching evaluation of HVC teachers is an important guarantee for improving teaching quality and an important basis for HVC education and teaching decision-making. The sustainable development of teachers' teaching must be guaranteed through a complete incentive system and measures. Determine the proportion of self-evaluation results and other evaluation results through investigation and study to ensure the fairness and rationality of “double-position” performance evaluation. For example, actively guide outstanding graduates to work in the enterprise, select excellent “double-position” teachers to carry out employee training for the enterprise, and share some experimental facilities and latest research results with the enterprise, so as to attract enterprises to actively participate in the performance evaluation of “double-position” teachers and guide “double-position” teachers to pay attention to the improvement of practical ability.

The ultimate goal of performance evaluation is to promote the growth of “double-position” teachers. Therefore, vocational colleges should make rational use of the results of performance evaluation and gradually improve the relevant incentive mechanism. After the performance evaluation work is finished, the results of the performance evaluation should be publicized on the campus website in time to ensure the openness and transparency of performance evaluation and arouse the attention of “double-position” teachers. For the “double-position” teachers who have made great progress, the school should give them corresponding honorary titles and encourage them to constantly improve themselves. Whether the “double-position” teacher performance evaluation index is set scientifically and the evaluation method is reasonable is directly related to the future development of vocational colleges.

19.4 Conclusion

The ability of teaching reform and innovation is the core ability to measure the development effect of HVC “double-position” teachers’ teaching ability. It mainly includes the ability of teaching understanding, teaching design, teaching organization, teaching practice and teaching research. In order to improve the accuracy of “double-position” teachers’ teaching ability evaluation, this paper puts forward a “double-position” teachers’ teaching ability evaluation method based on big data on the basis of constructing HVC “double-position” teachers’ teaching ability evaluation system. The evaluation results of the proposed method are reliable and can reflect the real ability of teachers. Through quantitative assessment, “double-position” teachers can be encouraged to keep learning and actively participate in teaching and scientific research, thus promoting the development, promotion and play of double-qualified teachers’ ability; On the basis of clarifying the construction value of “double-position” teacher performance evaluation system, it is helpful for HVC to build a higher level of “double-position” teachers’ performance evaluation system and adopt diversified implementation strategies. On the whole, the research results expand the knowledge system about the connotation and standards of teachers’ teaching ability in theory, and provide a basis and an operable tool for the evaluation and diagnosis of teachers’ teaching ability in vocational colleges in practice.

References

1. Xie, Y.Y.: Project curriculum reform and "double-position" teacher training research. Educ. Occup. **000**(026), 54–56 (2015)
2. Yuan, X.: Research on the "double-position" teacher evaluation system in transitional universities-Comment on "research on the dilemma of on-the-job training of teachers in higher vocational colleges from the perspective of" double-position. Sci. Technol. Manag. Res. **40**(23) (2020)
3. Cai, H.Y., Xiong, K.H.: Action learning: the training path of "double-position" teachers' teaching ability in applied undergraduate colleges. Heilongjiang High. Educ. Res. **2018**(6) (2018)
4. Li, Z.Y.: Research on the development of "double-position" teachers' teaching ability in higher vocational colleges from the perspective of action learning. Heilongjiang Anim. Husb. Vet. **2018**(14) (2018)
5. Ma, L., Cao, Y.Q.: Research on the three-dimensional evaluation of "double-position" teachers in higher vocational colleges. Educ. Occup. **2022**(4), 89–95 (2022)
6. Shi, W.S.: Dual structure teacher group: teachers' cooperative teaching and training "double-position" teachers. China Vocat. Tech. Educ. **2019**(14) (2019)
7. Li, Y.X., Zhang, D.K.: A new theory of "double-position" teacher's level management in higher vocational education. High. Eng. Educ. Res. **2018**(1) (2018)
8. Jiang, D.F.: Research on the development of "double-position" teachers' teaching ability in application-oriented colleges. China Adult Educ. **2020**(7) (2020)
9. Zhou, C.G., Dang, Y.G., Ye, L., et al. Performance evaluation of "double-position" teachers in higher vocational colleges based on improved grey relational analysis model-taking Jiangsu tourism vocational college as an example. Vocat. Tech. Educ. **41**(3) (2020)
10. Zeng, S.Y., Lu, S., Ai, Q.Z.: Research on the policy of "double-position" teacher evaluation in higher vocational colleges. China High. Educ. Res. **37**(10), 102–108 (2021)
11. Jeff. "Double-position" research on the development of teachers' teaching ability. China Vocat. Tech. Educ. **38**(26) (2022)
12. Wu, H.Y., Hou, L.M.: University "double-position" teaching evaluation mechanism is perfect. Heilongjiang High. Educ. Res. **37**(7) (2019)

Chapter 20
Design of Intelligent Transportation Assistant Decision System Based on Traffic Big Data

Min Sun

Abstract The increasing traffic "big data" brings new challenges to traffic management innovation and puts forward new requirements for traffic management. Traffic information construction will inevitably step into the intelligent application stage brought by big data. Using big data to crack the bottleneck of urban traffic will become the only way for the future traffic industry. Big data technology can send accident information to nearby vehicles to prevent accidents from happening again, and encourage other vehicles to change their driving routes. It can effectively improve the operational capacity of traffic roads and maximize the traffic infrastructure. The intelligent traffic assistant decision system designed in this paper processes and analyzes the data, and finally forms a traffic state prediction method based on the dynamic reliability of multi-source traffic data to predict the traffic state of the whole road network. The main goal is to make use of the information, models, methods and rules provided by traffic data warehouse and information intelligent analysis module to solve reasonable answers to problems in an acceptable time, such as the optimal control parameters and control strategies for regional signal control, traffic allocation and route guidance, to make an optimal decision on the traffic command and guidance scheme.

20.1 Introduction

The rise of big data and artificial intelligence has set off an upsurge of data mining and utilization, and also marked the arrival of a new era. With the continuous development and deepening of big data and artificial intelligence, all walks of life want to effectively use the existing large-scale data through the big data method, and tap its intrinsic value to create better development for the industry. Big data has been used by researchers to solve various problems in transportation, and some achievements

M. Sun (✉)
Wuhan Institute of Design and Sciences, Wuhan, China
e-mail: mm.sunny@126.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_20

have been made. However, there are still some shortcomings in the application of big data theory. Due to the complexity of the intelligent transportation system itself and the increasing interdependence of various subsystems, isolated subsystems will be difficult to make a difference. Due to the increase of vehicles and the limitation of land resources, optimizing the transportation system is a top priority. Developed countries, represented by Europe, the United States and Japan, have gradually shifted from mainly relying on expanding the scale of the road network to solving the growing traffic demand to using high-tech to improve the existing road transportation system and its management system [1]. Big data technology can send accident information to nearby vehicles to prevent accidents from happening again and encourage other vehicles to change their driving routes. It can effectively improve the operation capacity of traffic roads and maximize the traffic infrastructure.

The growing traffic "big data" brings new challenges to the innovation of traffic management and puts forward new requirements for traffic management. The construction of traffic informatization will inevitably step into the smart application stage brought by big data, and using big data to solve urban traffic bottlenecks will become the only way for the future traffic industry [2]. The application of big data technology can timely notify drivers in case of traffic congestion, and propose feasible solutions to improve the traffic operation capacity. Big data is also real-time. It can process and apply static data, and intelligently use big data to make traffic operations more reasonable. Static data refers to relatively stable data for a period of time, such as infrastructure, transportation tools, etc. [3, 4]. Dynamic data refers to data with continuous changes in time and space, such as traffic.

The key to intelligent transportation is to improve people's travel efficiency and reduce the travel cost and the operation cost of the transportation system on the premise of meeting people's travel needs. The transportation information has a wide range of sources, types, forms and volumes [5]. Therefore, the storage, organization, presentation, and query optimization of information have become the problems to be solved. This paper builds an intelligent transportation aided decision-making system based on transportation big data. Promoting the construction of intelligent transportation through this system is an effective way to promote the development of modern comprehensive transportation system. The basic, forward-looking and service-oriented functional positioning of transportation determines that intelligent transportation is an important part of smart city construction [6].

20.2 Comprehensive Monitoring of Urban Traffic Operation

20.2.1 Present Situation of Comprehensive Traffic Operation Monitoring

With the increasingly serious traffic congestion, how to make full use of the existing traffic facilities, efficiently develop the limited traffic resources, and relieve the traffic congestion through scientific management means has become a prominent and important issue in the social development, and the traffic flow theory came into being to solve this key problem. In order to realize the transformation of transportation industry from independent operation of various industries to comprehensive coordination, an authoritative, comprehensive and real-time transportation data center, a monitoring and early warning center for traffic conditions, an operation coordination center for various modes of transportation, and a unified external information release center will be built to realize the functions of data management, industry monitoring, decision support, information release and collaborative service of transportation industry [7]. Every day, the vehicle data generated on the traffic roads are very huge, and it takes a long time to search for accurate vehicle information from the huge database. At the same time, the existing traffic control system does not have the ability to analyze and process the information stored by itself, and the screening process can only be realized by manual labor.

Traffic flow has the characteristics of complexity, dynamics and randomness. The theoretical research of traffic flow is mainly to establish a mathematical model that adapts to it by analyzing the actual traffic flow characteristics, and to establish a normalized information reporting and sharing mechanism with civil aviation, railway, meteorology, tourism and other units through numerical simulation and theoretical analysis. At the same time, relying on comprehensive traffic monitoring data, we actively explored the coordination and linkage of various modes of transportation, and realized the passenger flow and capacity allocation of taxis in airports, railway hubs, etc. Especially during major holidays, provide one-stop information reporting and sharing services with cooperative units. Thereby revealing the inherent characteristics of traffic flow, and providing theoretical basis for traffic network optimization and management and traffic guidance.

Comprehensive radio and television media, websites, Weibo, mobile phone clients and other ways are used to regularly publish all kinds of traffic analysis reports, traffic forecasts and real-time road condition broadcasts, providing unified comprehensive traffic information and service support for motor vehicle driving groups and public transportation groups in the city [8]. Time series theory is one of the powerful tools to study the spatio-temporal evolution of traffic flow. Using time series theory, we can analyze the traffic flow changes of road sections in a period of time, extract the congestion of road sections, and deduce the traffic state of the whole road according to the spatial propagation characteristics of traffic state.

20.2.2 Existing Problems

With the continuous improvement of the traffic management department's demand for traffic information management and services, the urban traffic geographic information integrated management system needs to be designed according to the idea of modularization and expandability. The modularization process should follow the principle of high cohesion and low affinity, so that the module can complete certain functions independently as far as possible, and maintain the operation independence of the module to a certain extent [9]. The traffic data is composed of multiple branch single tables, and the statistical information of each single table is different. For example, the speed detection and vehicle information shooting are completed by different systems and are reflected in different single tables. The functions of the comprehensive traffic monitoring platforms in major cities are similar, basically realizing the core functions of the city's comprehensive traffic monitoring, coordination and linkage, emergency dispatching, operation analysis, and comprehensive services, and providing data support for the smooth, safe, and efficient operation of urban traffic through accurate monitoring and active services; However, the following problems exist in data based auxiliary decision-making and data accuracy and integrity [10]. The problems in urban comprehensive traffic operation are shown in Fig. 20.1.

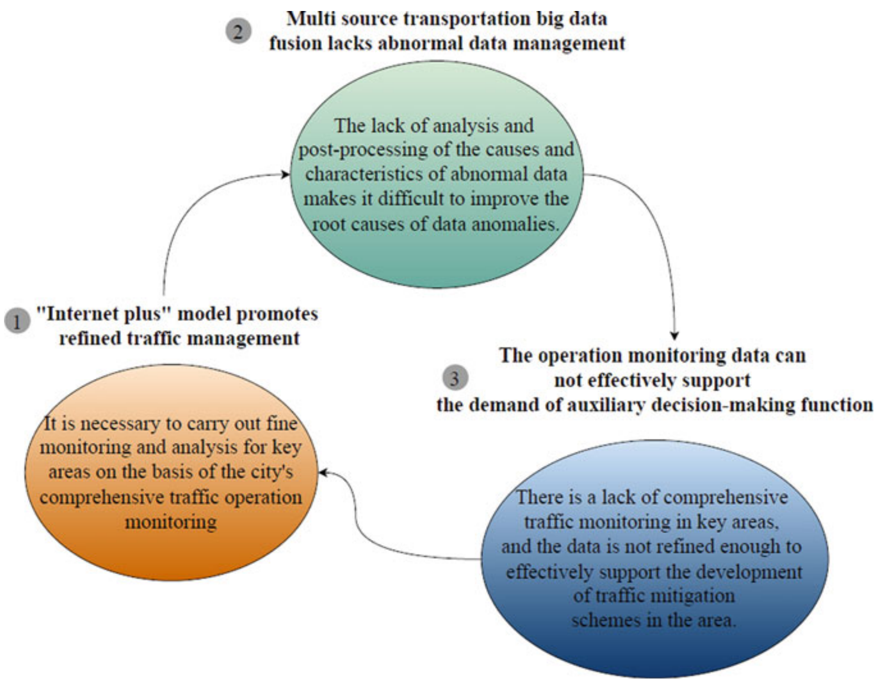


Fig. 20.1 Analysis of problems in urban comprehensive traffic operation

The integrated management system of urban traffic geographic information involves interface display, man–machine interaction, organization and access of traffic data, query and statistics of traffic information, call of assistant decision-making functions, etc. Because the traditional two-tier architecture adopted in system design has certain defects in operating efficiency and functions [11]. With the technical upgrade of traffic control system, the forms of collected data information are changeable, the proportion of pure video data begins to decrease, and text information and picture information begin to increase. The big data intelligent system can be compatible with various forms of information, and at the same time, the unified analysis platform can realize various forms of free switching.

In the field of road survey and design, special map of survey information, auxiliary map of route selection and three-dimensional display map of route selection can be extracted; In the field of road network monitoring and emergency, geological hazard level map, traffic land monitoring map, post-disaster road section assessment map, damage emergency assessment map and major disaster investigation can be provided [12]. The modern traffic control platform mainly relies on surveillance to collect data, that is, the information of vehicle license plate and cab can be captured by camera. The processing conditions of video data are higher than those of literal or digital information, and the changes of each frame in the video are accompanied by the changes of time, which further enhances the requirements of information processing system, which not only needs to analyze the data, but also needs to ensure the accuracy of time.

20.3 Design of Intelligent Transportation Aided Decision System

20.3.1 System Design

The intelligent transportation aided decision-making system is to effectively and comprehensively apply advanced satellite positioning and navigation technology, computer technology, image and graphics processing technology, sensor technology, information technology, electronic control technology, data communication technology, operational research, artificial intelligence and other high-tech technologies to transportation, service, control and vehicle manufacturing. The physical model of data warehouse mainly solves the problems of data indexing strategy, data storage strategy and storage allocation optimization. The main purpose of physical model modeling is to improve system performance and more effectively manage stored data. The access frequency, data capacity and storage medium configuration will affect the final result of the physical design.

On the one hand, the spatio-temporal correlation of traffic information provides strong support for traffic control, prediction and research; On the other hand, it puts forward new requirements for information processing technology. The time series analysis, spatial data mining and other modern information processing technologies have strong adaptability to such data, can accurately predict its change trend, and provide a reference for real-time traffic control. Based on the transportation big data warehouse and transportation big data integration into the system, it provides a multi-layer application architecture, including data layer, algorithm layer, analysis layer and application layer. The multi-layer architecture of transportation big data can be applied in a cross platform and network environment. The system can adopt flexible ways as required. The system operation environment includes several subsystems, including data area, transportation application area, GIS area, communication area, and terminal area. The database server, transportation application server, GIS server, communication server, and a series of operating terminals are arranged respectively. Each subsystem realizes different functions. The system network topology is shown in Fig. 20.2.

The intelligent transportation assistant decision-making system can adjust the entire intelligent transportation assistant decision-making system through the analysis and processing of real-time data, with the strong support of communication facilities and policies and regulations, so that the entire system is in the best state, and can respond to emergencies in a timely manner and solve them in a timely manner. Based on the above considerations, our guiding ideology for the modeling of intelligent transportation information analysis and auxiliary decision-making system is to always treat intelligent transportation system as an “intelligent system”. For the characteristics of transportation system, transportation information, transportation

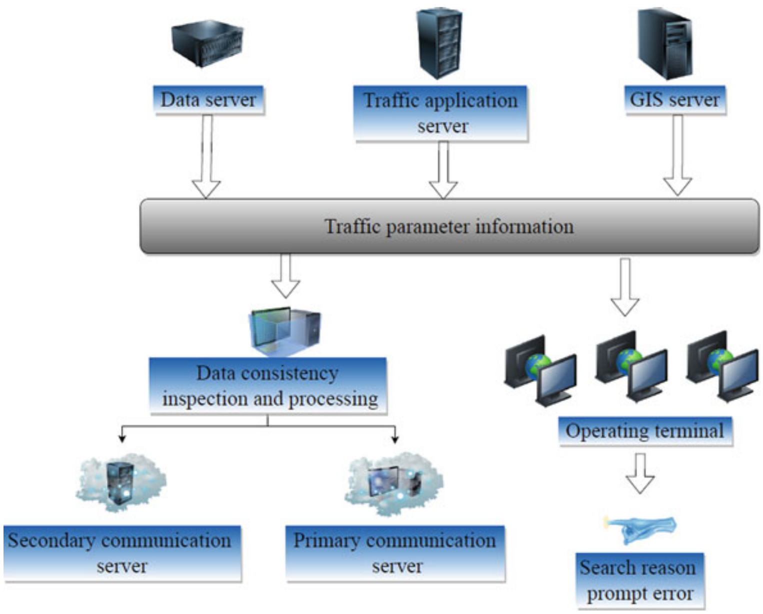


Fig. 20.2 System network topology

behavior, etc., transportation big data adopts a data warehouse on the basis of traditional decision support system based on database, knowledge base and model base. Traffic guidance pursues dynamic optimization. In order to minimize the total travel cost of each traveler's instantaneous choice of path in the system, it is defined as:

$$\pi(t) = \min\left\{\tau_k^{rs}(t)\right. \tag{20.1}$$

where $\pi(t)$ represents the minimum actual travel cost between time pairs; $\tau_k^{rs}(t)$ represents the travel cost of route k that it takes to select a route at t time.

Let the starting stop of a traveler's trip be A and the ending stop be B. Where i represents the number of lines that travelers need to transfer to, l_i represents the distance between two stations for quality inspection. In this way, the following model can be established:

$$\min Z = \sum l_{ij} \tag{20.2}$$

D represents the density of direct passenger flow, and p_{ij} represents the passenger flow between stations i and j, l_{ij} represents the distance between stations i and j The route is planned by the method with the highest direct passenger flow density, specifically expressed as follows:

$$\max D = \frac{p_{ij} x_{ij}}{l_{ij} \Delta_{ij}} \tag{20.3}$$

$$\omega_i = \frac{V_i}{f_a x_a} \tag{20.4}$$

where f_a represents the joint departure frequency on the edge, and V_i represents the passenger flow on the node, ω_i represents the total waiting time of node i.

In order to ensure that the system can resume work in time when the data processing server is down, the dual-computer redundancy guarantee mode is adopted. The data processing server is divided into master and slave, which run the same data processing program and communicate with each other through heartbeat protocol. At the same time, because the amount of data is related to the number of clients, which may be large, the load balancing mode is adopted.

20.3.2 Implementation of the System

The intelligent transportation assistant decision-making system includes dynamic system and static system. Among them, the dynamic intelligent transportation subsystem includes a traffic flow monitoring system, information control system, high-definition video monitoring system, etc., with various data sources and static systems such as environmental road data. The data collection of transportation big data refers to the transmission of structured or unstructured data generated by various systems or equipment to the big data platform for storage. A road consists of many sections and intersections. The road table contains the following attributes: road ID, road name, speed limit, and road length. The detailed design is shown in Table 20.1.

Based on the data collected by the system designed in this paper, the traffic index of busy streets on October 1, October 4, and October 7 during the National Day holiday was analyzed. It was found that the traffic congestion on October 1 in the National Day holiday lasted for a long time, and the accumulated congestion duration was more than 10 h. The traffic congestion decreased slightly on October 4, and returned to normal on October 7. It can be seen from the time distribution map of the traffic index in busy streets that the holiday congestion time is concentrated between 8:00 and 16:00, and the duration is long; According to the data collected by the fixed detectors of the surrounding roads, the holiday traffic volume is significantly higher than the daily level of weekdays and weekends, about 42% higher than the annual level of weekdays, and about 30% higher than the level of weekends. See Figs. 20.3 and 20.4 for details.

From Figs. 20.3 and 20.4, it can be seen that the optimized traffic scheduling method significantly shortens the duration of interception congestion compared to other methods, improving the smoothness of traffic flow. For transportation hub

Table 20.1 Detailed design of road table

Serial no	Column name	Data type	Length	Explain
1	LENGTH	Date	18	Road length
2	SPEED	Double	31	Speed limit
3	AREA_ID	Varchar		Region ID
4	ROAD_NAME	Int	39	Road name

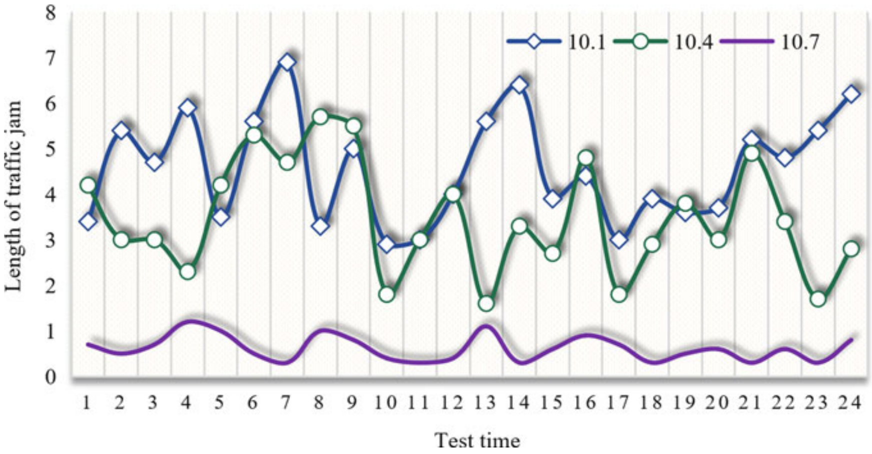


Fig. 20.3 Congestion time in busy neighborhoods

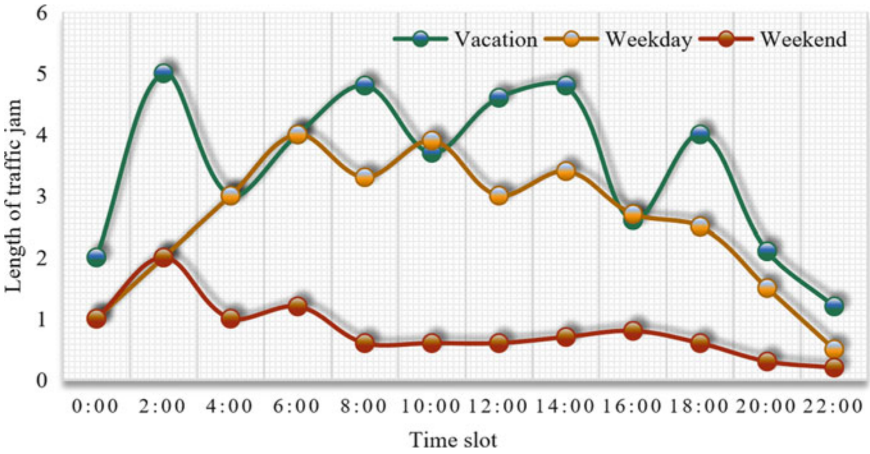


Fig. 20.4 Time distribution of traffic index in busy neighborhoods

areas, in conjunction with railway operation schedules before and after holidays, measures such as increasing ground bus numbers and increasing departure intervals during peak hours are taken to improve passenger flow distribution and transportation capacity, and meet passenger travel needs. In terms of rail transit, a flow restriction plan is formulated during peak hours, with additional station staff and civilized guides dispatched between stations to strengthen passenger flow guidance and reduce passenger flow interweaving. By analyzing the existing traffic state index, the intersection level traffic state index and its calculation method are proposed. The application of regional traffic state index can better reflect the basic situation of regional, road and intersection traffic, and provide a reference for the traffic management department to comprehensively master the traffic congestion situation. For the collection and processing of video data of video monitoring and vehicle detector data, the collection, transmission and storage of traffic big data are realized by integrating the current open-source log collection technology, data flow processing technology, message queue technology and distributed storage technology.

Therefore, the important data information of various comprehensive traffic models and real-time traffic models in the intelligent transportation aided decision-making system is processed and comprehensively applied through the traffic big data technology, so as to improve the efficiency of traffic management and road use, improve the service level of road traffic system, and reduce road traffic accidents.

20.4 Conclusions

The arrival of big data really facilitates people's travel. It not only solves many problems of difficulty in purchasing tickets and taking taxis, but also improves people's travel efficiency and reduces operating costs. And intelligent transportation will continue to solve the problem that is only being alleviated now. The intelligent traffic assistant decision system designed in this paper processes and analyzes the data, and finally forms a traffic state prediction method based on the dynamic reliability of multi-source traffic data to predict the traffic state of the whole road network. The main goal is to make use of the information, models, methods and rules provided by traffic data warehouse and information intelligent analysis module to solve reasonable answers to problems in an acceptable time, such as the optimal control parameters and control strategies for regional signal control, traffic allocation and route guidance; Make optimal decisions on traffic command and guidance schemes, etc. Using this technology in the field of intelligent transportation can make the generated traffic model more realistic; all kinds of traffic flow information are analyzed and predicted by time series analysis and sequence pattern mining. The static data, real-time information collection data and Internet data are integrated to realize the cleaning, extraction, integration and standardization of urban traffic information resources. On this basis, the construction of a monitoring data application question bank, abnormal data management question bank and customized application question bank are realized, and the visual analysis of all kinds of data is realized through the data visualization tool set.

Acknowledgements This article is supported by the 2021 school-level scientific research project of Wuhan Institute of Design and Engineering: "Narrative Research and Visual Animation Interpretation of Dunhuang's 'Deer King Bensheng' (No.: K202107)".

References

1. Darwish, T., Bakar, K.A., Kaiwartya, O., et al.: TRADING: traffic aware data offloading for big data enabled intelligent transportation system. IEEE Trans. Veh. Technol. **58**(49), 33–62 (2020)
2. Jia, Y., Xin, W.: Intelligent traffic decision analysis system based on big data mining. IOP Conf. Ser.: Mater. Sci. Eng. **392**(6), 062187–062196 (2018)
3. Li, X., Hsiao, J.: Big data oriented intelligent traffic evacuation path fuzzy control system. J. Intell. Fuzzy Syst. **35**(28), 41–67 (2018)
4. Huang, M., Yan, X., Bai, Z., et al.: Key technologies of intelligent transportation based on image recognition and optimization control. Int. J. Pattern Recognit. Artif. Intell. **53**(18), 39–55 (2020)
5. Shi, Q., Wang, L., Wang, T.: Traffic big data in the development of intelligent transportation system. Electron. Res. Appl. **74**(19), 38–49 (2018)
6. Putra, K.T., Wang, J.D., Prasetyo, E., et al.: Modeling traffic congestion with spatiotemporal big data for an intelligent freeway monitoring system. J. Theor. Appl. Inf. Technol. **56**(26), 56–69 (2019)

7. Wu, S., Dong, Z.: An auxiliary decision-making system for electric power intelligent customer service based on Hadoop. Sci. Program. **26**(11), 19–31 (2022)
8. Zhang, S., Lu, C., Jiang, S., et al.: An unmanned intelligent transportation scheduling system for open-pit mine vehicles based on 5G and big data. IEEE Access **102**(99), 45–81 (2020)
9. Zhou, Y., Chen, J.: Traffic change forecast and decision based on variable structure dynamic bayesian network: traffic decision. Int. J. Decis. Support. Syst. Technol. **13**(2), 45–61 (2021)
10. Kumar, P.K., Gopal, A.V.: Developing big data frame works for efficient intelligent transport system for safety and security. J. Theor. Appl. Inf. Technol. **97**(6), 1833–1842 (2019)
11. Wang, R.: Research on short-term traffic flow forecast and auxiliary guidance based on artificial intelligence theory. J. Phys: Conf. Ser. **1544**(1), 012164–012177 (2020)
12. Jung, J., Jayakrishnan, R.: Simulation framework for modeling large-scale flexible transit systems. Transp. Res. Rec. **2466**(1), 31–41 (2018)

Chapter 21
Construction of an Interactive Computer Aided English Language Teaching Based on Multimedia Technology

Wei Zeng

Abstract As a new teaching method, computer multimedia technology provides students with a large amount of information and materials from a visual and auditory perspective, stimulating their interest and initiative in learning English. The ELT (English Language Teaching) established in this article overcomes the closed and fixed nature of traditional computer teaching software, improves the interactivity and flexibility of learning, and is an open teaching system that integrates learning, testing, real-time tutoring, and management. The results show that from the perspective of students' Basic English ability, the average score is 16.15, the average score of oral ability is 14.04, and the average score of conversational communication ability is 8.13. In the test of standard deviation, it can be found that the standard deviation of students' Basic English ability is 13.4, the standard deviation of oral ability is 1.91, and the standard deviation of conversational communication ability is 0.88. It can be seen that the students' Basic English ability is better than the other two test modules, whether in the average score test or the standard deviation test. Learning at the right time and place can achieve various communication and interaction between students and computers, students and students, and teachers.

21.1 Introduction

With the rapid development of vocational education in China, all kinds of advanced teaching resources and equipment provide great convenience for improving teaching and scientific research in schools. As far as computer courses and design courses are concerned, using multimedia and network technology to set up a multimedia network experimental room can effectively solve the problem of difficult information interaction between teachers and students and improve the teaching effect of practical links [1]. In English teaching, teachers adopt various flexible teaching methods and

W. Zeng (✉)
Yibin Vocational and Technical College, Yibin 644003, Sichuan, China
e-mail: 2023236027@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_21

means to improve teaching efficiency and students' English application ability. As a new teaching method, computer multimedia technology provides students with a lot of information and materials from an audio-visual perspective, which stimulates students' interest and initiative in learning English, enriches the teaching content and improves the teaching effect [2]. These advantages of multimedia network technology are very useful for English teaching. It plays a great auxiliary role in optimizing English classroom teaching structure, mobilizing students' diverse senses, simulating real-life dialogue and improving students' attention. If multimedia technology is abused in teaching without combining teaching content and objectives, it will form a misunderstanding of "emphasizing form but neglecting content" [3, 4]. Students only get visual enjoyment, but their English application ability is not improved, which goes against the original intention and purpose of teaching with multimedia technology. Interactive teaching method was first put forward by Palinscar in 1982. It is a creative teaching theory and strategy. In the interactive teaching method, interaction, as a teaching method, refers to taking teachers as the leading factor and students as the main body, and guiding students to discover with teachers' inspiration in teaching. Compared with traditional teaching methods, network multimedia-assisted English teaching has obvious advantages: it greatly expands the learning space and content, and mobilizes learners' learning enthusiasm, initiative and creativity [5]. It also provides a good language carrier and a relaxed language environment for learners. The ELT established in this paper overcomes the closure and fixity of traditional computer teaching software and improves the interactivity and flexibility of learning. It is an open teaching system integrating learning, testing, real-time tutoring and management [6]. Studying at the right time and place can realize various exchanges and interactions between students and computers, students and students, and students and teachers. It provides the best environment for teachers and students to realize constructivism teaching and learning together, which is the mainstream of future education model.

21.2 Systems Design

21.2.1 Interactive English Teaching Mode

Interactive teaching emphasizes the educational concept of "people-oriented". Teaching activities should clarify that students are the center of the entire language activity, and the role of teachers is to create scenarios that are conducive to language. Cultivate students' self-learning ability and strengthen their interest in learning through encouragement, stimulate their enthusiasm for learning, and thus improve teaching effectiveness. By reviewing on the computer, students have consolidated their knowledge and gained an understanding of their mastery of the knowledge. At the same time, the computer also retains records of each student's practice or test, which can be regularly or irregularly reviewed and analyzed by teachers and students themselves, as well as their level of knowledge mastery and existing problems at a certain stage [7]. In terms of teaching methods, we can enable students to use various means such as online visits, online reading, online searches, online discussions, online displays, online Q&A, and online classrooms to enhance their interest in learning, broaden their knowledge, increase their vocabulary, and consolidate their proficiency in knowledge.

Through the operation of the teacher on the console, students can freely and easily engage in English conversations without seeing each other, increasing their chances of speaking and confidence. After continuous practice, students can speak English boldly and bravely learn from it. Go behind the scenes to the front desk to communicate in English and improve the comprehensive application ability of the English language. More questions can be asked in limited classroom hours to impart knowledge to students. But doing too much is not enough. In the practice of foreign language teaching, some teachers use this modern teaching method to digitize a large amount of content and continuously transmit knowledge to students [8]. English is a highly practical subject because the ultimate goal of learning a language is to use it. Therefore, if students cannot practice with a large amount of language input, there will be no language output and they will not be able to learn English well.

21.2.2 Design Philosophy

Multimedia English teaching provides students with vivid and realistic learning situations, stimulates many organs such as eyes, hands, ears and brain at the same time, and makes students have a strong interest in English, which is conducive to the development of students' thinking and the overall improvement of students' quality. Nowadays, people say that computer multimedia includes man–machine information communication means using other media besides text teaching, which usually refers to information media such as sound, graphics, images, moving images and animations [9]. In traditional comprehensive English teaching, the sentences used by teachers to explain words, sentence patterns and exercises are usually examples found in grammar books or dictionaries. Due to the limitation of teaching space, the examples given in dictionaries and grammar books are generally short and lack sufficient context, so they cannot objectively and comprehensively reflect the language usage context and frequency. This paper studies the design of interactive computer-aided ELT based on multimedia technology, because the design of ELT based on multimedia technology must strictly follow the design principles of software engineering and adopt unified multimedia technology standards [10, 11]. It has the characteristics of large information, diversified information input means and strong interaction.

Using computer multimedia to guide comprehensive English teaching can inject new vitality into English classes and achieve unexpected results. Nowadays, students like watching TV, and traditional teaching methods can't provide intuitive communication situations. However, using modern information technology can completely break through this point, so that all communication situations can be visualized in the classroom, and students can be more attracted to enter the role and participate in communication if they are immersive and empathetic. In the previous classroom teaching mode. The way we adopt is mainly to present the shape of the words we teach through cards, to teach the sounds of words through listening to recordings or teachers' reading, and to understand the meanings of words through teachers' explanations. It is impossible to achieve the organic unity of sound, shape and meaning. The ELT designed in this paper must complete all the functions of daily multimedia teaching, and the functions should be simple and convenient, so that teachers and students do not have to pay too much learning costs. At the same time, the design of the system should adopt the most advanced and mature framework system and technical design at present to ensure the advanced nature of the system.

21.2.3 System Architecture Design

Due to the fact that an ELT based on multimedia technology includes data access code, business logic code, and frontend presentation code, simply adding these codes together without any design pattern will result in a strong correlation among various parts of the system, which violates the principle of "high aggregation, low coupling" in software engineering, The implementation and maintenance of the system can also become exceptionally complex. The system records the user's usage, locks the network segment, and opens corresponding services as needed. When not needed, they can be turned off in a timely manner, effectively preventing hacker attacks [12]. In addition, due to the use of a hierarchical directory management mode in this system, the scalability and convenience of the system are improved.

The process of students learning English is no longer a passive acceptance process, but an active participation process. A strong interest in learning and active participation have formed an internal psychological process of student optimization. This psychological process interacts with external stimuli of optimization, enabling students to achieve different achievements based on their own characteristics in the process of learning English [13]. This system is based on the Unix operating system and can be mainly divided into application layer, service layer, and data layer. Each functional module has relative independence, which facilitates the design, implementation, maintenance, and expansion of the entire software. Ensure seamless connection between ELTs and external systems, leaving room for software expansion and integration [14]. Realized real-time interaction between teachers and students, as well as real-time self-assessment. The architecture is shown in Fig. 21.1.

The ELT established in this paper can call your prepared content through any networked computer. The teacher management module is mainly aimed at students' access to the system and the completion of homework, such as when was the last visit. The administrator module is mainly designed for system expansion and account management. In addition, there are all kinds of materials in the computer or we have some practical software, which is very convenient to use, and it is 100 times faster than looking through books or making some content by ourselves, which can save

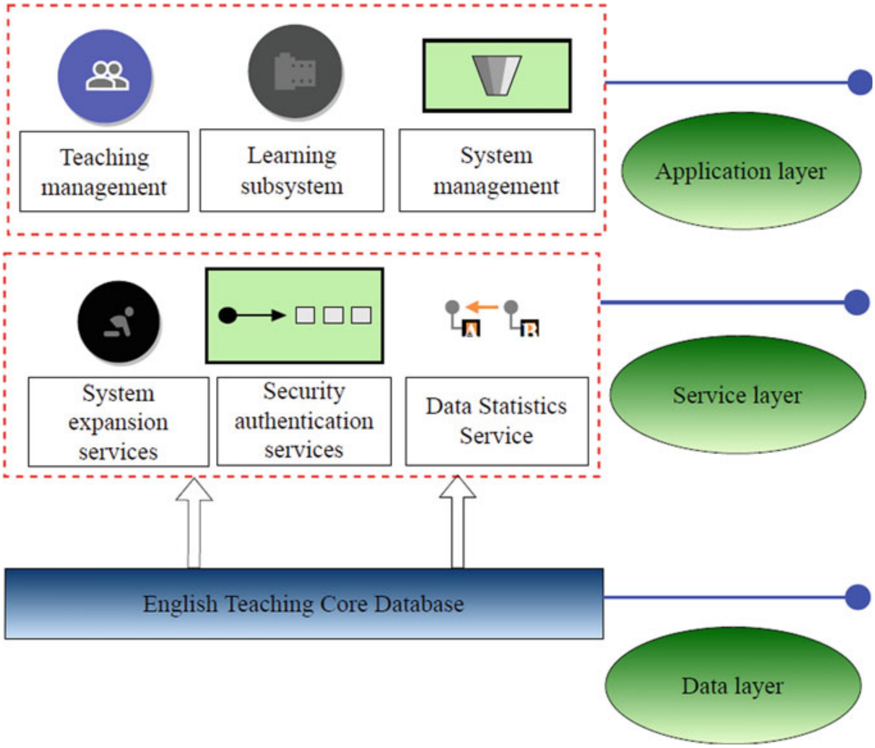


Fig. 21.1 Architecture diagram of ELT

a lot of time. The quality education of students' subjectivity frees students from the bondage of passive learning, makes students become the main body of learning, allows students to participate in teaching independently and actively, and gives full play to their initiative and enthusiasm.

21.3 Application of the System

21.3.1 Provide Materials that Meet the Teaching Content

In order to ensure better teaching effectiveness, teachers should pay attention to the language authenticity, content compatibility, practicality, and technicality of the selected materials when using multimedia technology for teaching. Remember not to present English materials that do not align with the teaching content and objectives in order to reflect the diversity of forms and the fancy content. Learning is a process of active construction, where a person's overall knowledge is formed through continuous interactions and changes in the environment. Constructivism believes that the learning environment is a place where learners can freely explore and learn autonomously.

The interactive function of multimedia technology reflects the two-way communication between teachers, students, and online learning materials. Through information exchange, the three constantly adjust themselves and create an optimal teaching environment. Maximize students' learning efficiency under the best possible conditions. The interactive content of English teaching in a multimedia environment is shown in Fig. 21.2.

Using computer network for teaching can solve this problem and realize personalized learning. With the help of the internet, students can easily read English and American electronic newspapers and magazines, learn standardized and popular English words and their usage on various English websites, and download reading materials that are both read aloud in the original sound and compared with the original text. These are irreplaceable for the cultivation of students' communicative competence, the ultimate goal of language learning.

We use the two-way nature of the computer network to encourage students to question and ask questions online in time during the learning process. Teachers

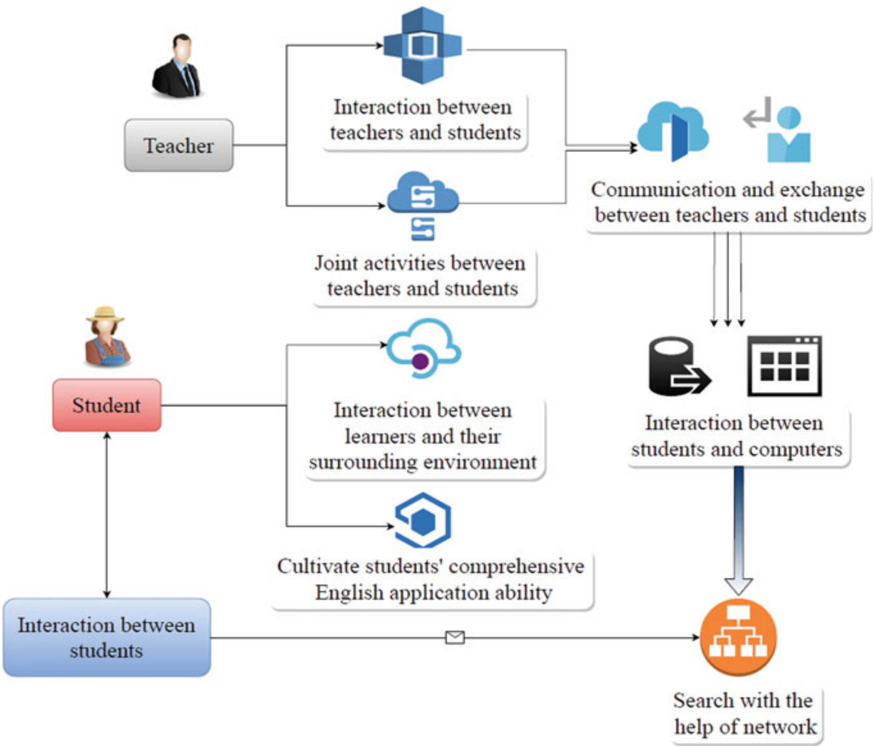


Fig. 21.2 Interactive English Teaching in a Multimedia Environment

give necessary instructions or answers to students' doubts. The use of computer network in English teaching greatly improves students' interest in learning English, broadens students' English knowledge, helps to develop students' various abilities and the healthy development of students' psychology, and improves students' English learning achievements. Interactive teaching modes can satisfy students' thirst for knowledge, expand their knowledge, facilitate personalized teaching for students, and greatly improve the learning efficiency of students at all levels.

21.3.2 Result Analysis

In the traditional English teaching environment, many exciting response scenes inspired by inspiration are fleeting, and the innovative thinking activities displayed during them cannot be reproduced to inspire future generations; The use of computer networks in English teaching can save many discussions and homework situations for other teachers and students to analyze, comment on, learn, and learn from, becoming a very precious educational resource. Teachers use the internet to assign learning tasks to students or raise unresolved issues; Provide students with resource catalogs, websites, and data collection methods and channels related to learning topics, allowing them to search for relevant materials and information online after class, in order to guide them to gradually break away from the dependence on face-to-face teaching. In order to further verify the effectiveness of the system established in this article, simulation experiments will be conducted to verify it. Taking the application in hotel English teaching as an example. As shown in Table 21.1, the assessment adopts post competence assessment, which is mainly based on the scores of foreign guests staying in the hotel. Students' Basic English ability, oral ability and conversational communication ability are assessed.

Adopt the above methods for teaching, and at the end of the academic year, collect the evaluation scores of 20 people for paired sample analysis. At the end of the experiment, gifts will be distributed as souvenirs to students, business instructors, foreign teachers, and foreign guests participating in the experiment. Students with excellent works will receive a 2 point quality development score. According to the above test modules, the average scores of students' Basic English ability, oral ability

Table 21.1 Content of English comprehensive skills test

Test module	Test content	Scoring criteria	Score
Basic English	Basic expressions and sentence application in hotel English	Objective evaluation	40 points
Oral proficiency	Conversation with foreign teachers	Proficiency and content in spoken language	30 points
Conversational communication skills	Cosplay	Role playing degree	30 points

and conversational communicative ability are shown in Fig. 21.3, and the standard deviation is shown in Fig. 21.4.

From the perspective of students' Basic English ability, the maximum score is 20.8, the average score is 16.15, the average score of oral ability is 14.04, and the average score of conversational communication ability is 8.13. In the test of standard deviation, it can be found that the standard deviation of students' Basic English ability is 13.4, the standard deviation of oral ability is 1.91, and the standard deviation of conversational communication ability is 0.88. It can be seen that the students' Basic

Fig. 21.3 Average score test

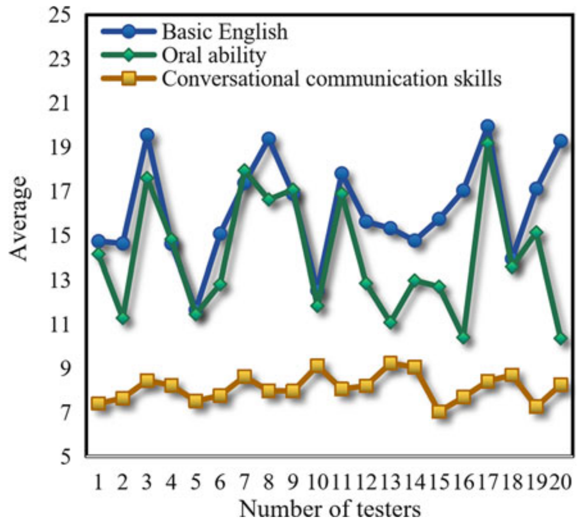
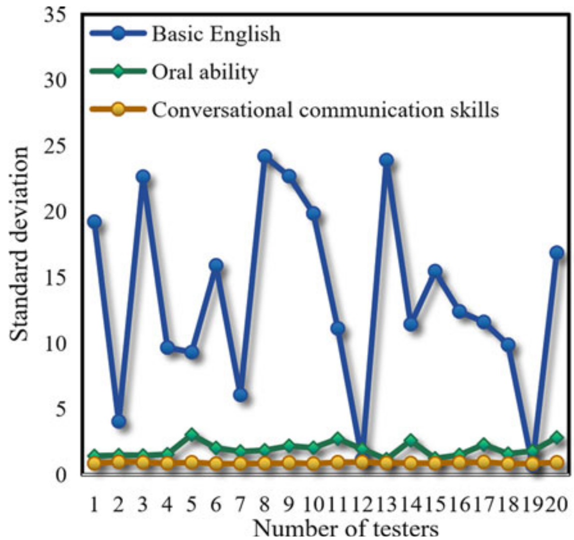


Fig. 21.4 Standard deviation test



English ability is better than the other two test modules, whether in the average score test or the standard deviation test.

21.4 Conclusions

Computer-assisted language teaching has great potential, but it does not mean that it can replace classroom teaching and create miracles. Multimedia can help to teach by going up one flight of stairs well, but it can't solve the problem of improper teaching. In English classroom teaching, if we can use the computer multimedia teaching system reasonably and timely, create English learning situations, and solve English teaching audio-visual and other related teaching problems, so that students can perceive English language materials through audio-visual and develop their audio-visual ability. The results show that the average score for students' basic English ability is 16.15, the average score of oral English ability is 14.04, and the average score of conversational communication ability is 8.13. In the standard deviation test, it can be found that the standard deviation of students' basic English ability is 13.4, the standard deviation of oral English ability is 1.91, and the standard deviation of conversational communication ability is 0.88. It can be seen that students' basic English ability is superior to the other two test modules in both the average score test and the standard deviation test. In addition, multimedia teaching requires students not only to have a certain level of English, but also to have computer knowledge and operational skills. Foreign language teachers are required to master certain multimedia software production techniques, be familiar with the use of various devices and be familiar with computer operation. Teachers can better understand students' learning style and learning ability, so as to formulate teaching activities that are more suitable for learners. Teaching has changed from "indoctrination" to discussion, and students have become the center of teaching activities.

References

1. Wu, H.: Multimedia interaction-based computer-aided translation technology in applied English teaching. Mob. Inf. Syst. **68**(20), 41–55 (2021)
2. Xiao, M.: Application of multimedia technology in primary school English classroom teaching. CAMPUS ENGLISH **62**(43), 19–38 (2022)
3. Chen, J., Qian, Y.: Research on the application of computer multimedia technology in ELT and teaching. J. Phys: Conf. Ser. **1915**(3), 032018–032031 (2021)
4. Li, P., Zhang, H., Tsai, S.B.: A new online and offline blended teaching system of college English based on computer internet technology. Math. Prob. Eng.: Theory Methods Appl. **68**(50), 89–97 (2021)
5. Guan, N., Song, J., Li, D.: On the advantages of computer multimedia-aided English teaching. Procedia Comput. Sci. **131**(50), 727–732 (2023)

6. Hindeme, U.: Fostering students writing skills through the use of computer and its multimedia system in English as a Foreign language class in Benin. Res. English Lang. Teach. **2**(4), 68–84 (2021)
7. Norbekova, G., Safarova, D., Bazarov, B., et al.: Interactive learning environment in the English language teaching with the help of computer technology. J. Phys: Conf. Ser. **66**(1), 58–72 (2019)
8. Qurat-Ul-Ain, Q., Shahid, F., Aleem, M., et al.: A review of technological tools in teaching and learning computer science. Eurasia J. Math. Sci. Technol. Educ. **15**(11), 41–57 (2019)
9. Rao, L.: On the exploration of college English teaching management mode under computer multimedia technology. J. Phys: Conf. Ser. **1915**(4), 042005–042068 (2021)
10. Sun, Y.: Improved design of ELT based on multimedia technology. Mod. Electron. Tech. **25**(10), 46–67 (2022)
11. Wu, Z., Cao, X.: Application characteristics of multimedia technology in Internet times in higher vocational English teaching. China Comput. Commun. **15**(10), 44–59 (2019)
12. Wu, Y.: A brief talk on how to use multimedia technology to construct a diversified evaluation system in English teaching. J. Phys: Conf. Ser. **1744**(4), 042058–042068 (2021)
13. Zhao, Y.: A personalized English teaching design based on multimedia computer technology. Int. J. Emerg. Technol. Learn. no. iJET) **15**(8), 210–245 (2020)
14. Li, C., Zhang, X.: Abnormal Signal Recognition Method of Wearable Sensor Based on Machine Learning. IoT and Big Data Technologies for Health Care, pp. 322–338. Springer Nature Switzerland, Cham (2022)

Chapter 22
Construction of Tool Life Dynamic Prediction Model Based on On-Line Learning

Jicheng Duan and Kechang Zhang

Abstract A tool life dynamic prediction model based on on-line learning is proposed in this paper. The model selects signal features as input data and regards tool wear and residual life as data labels, a tool residual life prediction model based on the improved long-short time memory network can predict the tool residual life dynamically under different working conditions.

22.1 Introduction

The manufacturing industry is leading our country's national economic development, vigorously developing the manufacturing industry is the foundation of building a manufacturing power. The proposal of Industry 4.0 has clarified relevant requirements for the global manufacturing industry to develop towards intelligent manufacturing. In short, Industry 4.0 is the "Internet plus manufacturing". It optimizes the integration of multiple factors of production and connects information, resources, goods, equipment and people, and makes it realize better network, information, automation and intelligence [1].

In the field of machining and manufacturing, the state of cutting tool directly affects the machining efficiency and quality of NC machine tool, and indirectly affects the level of production technology and economic benefit. If the tool life of the machine tool does not reach its expected life, it will lead to tool waste, cause the manufacturing cost to rise, and cause the production efficiency to decline, will lead to the reduction of product processing accuracy and productivity, and may even lead to machine damage, causing a major loss of equipment. Therefore, it is important to predict the residual life of NC machine tools, whether in reducing the cost of

J. Duan · K. Zhang (✉)
Hunan Railway Professional Technology College, ZhuZhou 412001, China
e-mail: 545939952@qq.com

J. Duan
e-mail: 305748030@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_22

production or in ensuring the accuracy of products and the safety of machining equipment.

In the process of NC machining, the abnormal state of cutting tool often occurs, which will lead to various quality problems in product processing. A great deal of data shows that there are more than 40 broken tools in a large-scale machining shop on average every day, most of which are caused by the tool not being changed after its service life [2]. The statistical analysis of KENNAMTAL company in the United States shows that the cutting tool monitoring system can save about 30% of the processing cost for NC machine tools, and W. Koening's statistical analysis data show that the tool condition monitoring system can increase the production efficiency of CNC machine tool by 10–60%, and reduce the time of failure by 75%, the efficiency of CNC machine tool is increased by more than 50% [3].

The tool life management of NC machine tool can effectively predict the remaining useful life (Rul) of NC machine tool, and judge the abnormal state of the tool in advance [4], the utility model can avoid the reduction of the product processing quality or the appearance of the waste products caused by the tool failure, and can effectively guarantee the normal processing of the product and promote the improvement of the product processing efficiency.

22.2 The Process of Tool Wear

In NC machining, the cutting tool wear is increasing with the cutting time, as shown in Fig. 22.1, the cutting time curve of the cutting tool wear in milling process, the wear state can be divided into three stages: initial wear, normal wear and sharp wear.

In Fig. 22.1, I is tool initial wear, II is tool normal wear, III is tool sharp wear.

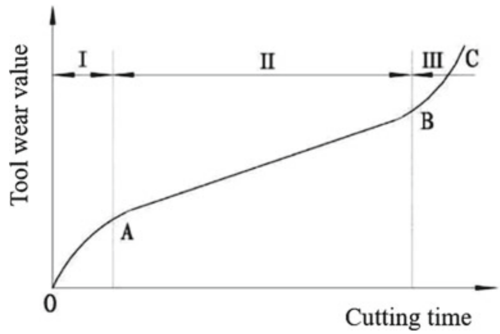


Fig. 22.1 Tool wear process curve

22.3 The Status of Tool Residual Service Life Prediction of NC Machine Tools

There are mainly four kinds of prediction methods in the research of residual service life prediction of NC machine tools, it mainly include methods based on empirical model, AI model, statistical model and hybrid methods [5, 6]. The following are the NC MT of the residual life of the prediction methods for the status quo analysis.

22.3.1 Forecasting Methods Based on Empirical Models

Coble et al. [7] used linear regression model to build a tool residual service life model. Karandikar et al. [8] regard the root mean square of spindle power as a health factor, and fit the quadratic function of health factor and cutting time, the trend of health factors is updated by Bayesian inference and its posterior distribution is obtained. The residual tool life can be obtained by extrapolating the trend to the threshold. Kumar et al. [14], based on the milling force signal of NC milling machine, used AdaBoost algorithm to integrate the trajectory similarity and differential evolution support vector regression, and established the tool residual life integrated prediction model. Yang et al. [9] extracted the features of tool monitoring signals, used principal component analysis to reduce the dimension of feature vectors, constructed health factor vectors, and combined with hierarchical clustering to predict the residual service life of milling cutters.

22.3.2 Prediction Method Based on AI Model

Huang et al. [10] put forward the idea of tool residual life prediction based on edge and cloud, the task of data collection and feature extraction is realized at edge, and the tool residual life prediction is realized by LSTM in cloud. Zheng et al. established a tool life model based on machining parameters, and used LSTM to predict residual tool life. Niu et al. [11] proposed an integrated DWT method to predict tool life by CNN-LSTM model.

22.3.3 Statistical Model Based Forecasting Methods

This method takes the trend of current health factors as the basis to predict the future health factors, and combined with the threshold of health factors to predict the tool life, mainly including three types of models, namely Stochastic Process model, Time Series model, State Space model. Based on the risk model and multi-layer perceptron, Li and Liu [12] constructed the state transition and observation probability of HMM, and accomplished the prediction of tool life under different machining conditions. Liu et al. [13] extracted the characteristics of vibration signals as tool health factors, and used ARIMA model to predict tool life.

22.3.4 Mixed Methods

The hybrid method combines different models to predict the residual service life of cutting tools in order to obtain higher prediction accuracy. Kumar et al. monitor the drill signal and use HMM to cluster the time series. Sun et al. get the sensitive features of cutting force and vibration, input them into BP neural network, and then use the network to predict the tool wear value, and get the predicted wear sequence, then the tool wear state is monitored and updated in real time by using the linear Wiener process, and the tool life can be predicted [14].

22.4 The Establishment of Tool Life Prediction Model

22.4.1 Overview of Tool Life Prediction Models

Modified Long-Short Term Memory (MLSTM) is an improvement of Long-Short Term Memory (LSTM), which optimizes part of the network, structure, and gate structure. The tool life prediction model proposed in this paper consists of MLSTM layer and regression layer. MLSTM layer processes time series information, including extracting information from input data to hidden layer, and outputting multi-dimensional vector for prediction, the regression layer is used to fit the function to output the predicted values of tool wear and tool residual life. The network model is expanded as shown in Fig. 22.2.

The output of the tool life prediction model is the tool wear amount and the tool remaining life, where the tool wear amount is the maximum tool back face worn and uniform tool back face worn of the tool wear at the current time, and the tool remaining life is the remaining life when the tool reaches the wear limit under the working condition information of the current time step. In this paper, we use VB to express the tool back face worn.

22.4.2 MLSTM Layer Structure

MLSTM is an improvement of Long-Short Term Memory (LSTM), which optimizes part of the network, structure and gate structure. The MLSTM retains the cell state cell

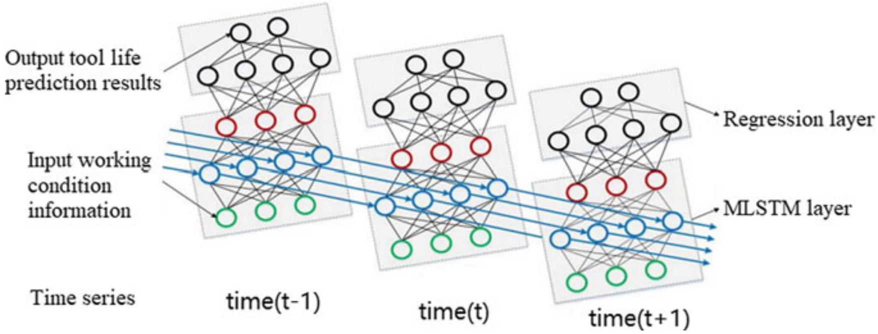


Fig. 22.2 Tool life prediction model based on MLSTM

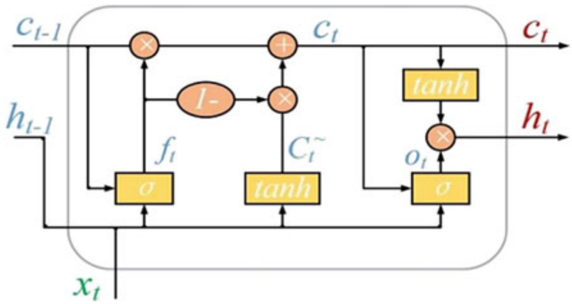


Fig. 22.3 MLSTM internal structure

for recording long-term information, and uses the three structures of forgetting gate, input gate and output gate in the hidden layer to control the transmission and recording of short-term information, the cell state of the previous moment is integrated into the current, and the input of the moment is helpful to control the information flow reasonably according to the cell state. The internal structure of MLSTM is shown in Fig. 22.3.

In Fig. 22.3, C_t is the description of the current input cell state, and " × " represents the multiplication by bits to finally complete the cell state update from C_{t-1} to C_t. f_t is the forgotten gate state, O_t indicates the output gate and h_t indicates the hidden layer information of the current status.

The forgetting gate in LSTM controls the trade-off of historical information in the cell state, which is characterized by the gate structure output control coefficient to the last, cell state and the current state input. In MLSTM, the last cell state information is integrated into the input node, and the input matrix data is expanded.

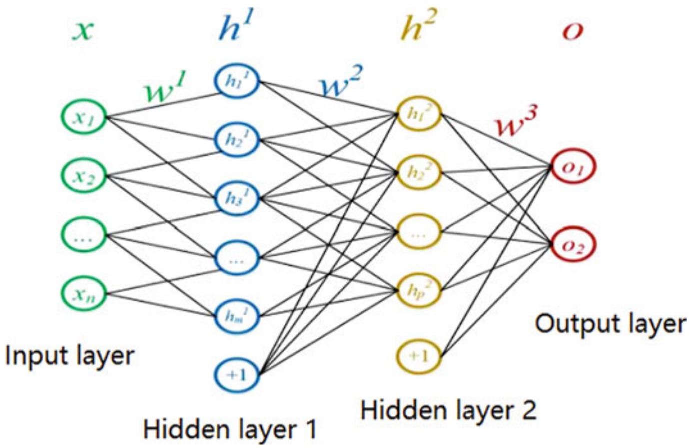


Fig. 22.4 Structure of regression layer

22.4.3 The Structure of Regression Layer

MLSTM processes the complicated working condition information input and establishes the relation in time series, but does not get the tool residual and life prediction value directly, the output information obtained by MLSTM is translated into the required residual life, value and tool wear. It has been proved that a single hidden layer neural network can approach an arbitrary shape function infinitely through a nonlinear activation function without limiting the number of neurons, the number of single-layer neurons and the number of network weights can be effectively reduced by increasing the number of neural network layers. Based on the above theory, the regression layer is designed as a fully connected network with double hidden layers to approximate the curve of tool residual life.

The output of MLSTM is a one-dimensional vector of length M, and the input of the first hidden layer is defined as the matrix X (X_1, X_2, X_3, … X_n) of N Rows 1 column (Fig. 22.4).

22.4.4 Hyperparameter Adjustment of the Prediction Model

The training method of the regression layer is the back propagation algorithm, which calculates the error between the predicted value and the true value, obtains the weight gradient based on the chain derivation rule, completes the weight update, and realizes the training of the regression network. The calculation formula is as follows:

$$W = W + \eta \frac{\partial E}{\partial W} \tag{22.1}$$

In the formula, W is the weight matrix, E is for energy, η is the learning rate.

After the MLSTM output data is processed by the regression layer, the output is the tool wear and the tool remaining life.

In the training of network model, it is necessary to adjust the hyperparameters to shorten the training time and improve the accuracy of the model. The number of neurons in each layer, the number of network layers and the regularization parameters in the hyperparameters mainly affect the accuracy of the network, the learning rate, the choice of activation function types, the definition of loss function, the weight initialization method mainly affects the network training speed, the number of training iterations, the size of batches, the time step length mainly affects the network training time and the network accuracy. The over-parameter adjustment is mainly through the empirical analysis method and the trial-and-error method, that is, according to the existing knowledge theory to set a better default value initially, and then through the trial-and-error method to gradually adjust the single parameter. This method is a local optimization method, each greedy to get the optimal solution of a parameter, and then check and tune the other parameters, and get the final adjustment results. First, the number of layers and the number of single-layer neurons were determined, then the loss function and the activation function were determined, then the batch size and the time step length were determined after the learning rate was approximately determined, and finally, the regularization parameters and the number of training iterations are checked.

22.4.5 Tool Life Dynamic Prediction Model

The on-line learning module is parallel to MLSTM network and independent of each other. The on-line learning module does not affect the loop of MLSTM in time step, and only assists in the final output of the predicted value. Therefore, it is considered that the tool life prediction value is still near the global best when the short time actual condition information changes, therefore, the global optimal solution can be obtained by adding adjustment value to the final output, so the on-line learning module can improve the network model accuracy.

In Fig. 22.5, SV is Supplementary Vector (SV), o_t is the output gate state, U and V are weight matrices, and x_t is the input working condition information parameter at time t.

The structure of the tool life dynamic prediction model is shown in Fig. 5. The on-line learning module behaves as an SV, whose content is the tool wear amount automatically measured on-line based on machine vision, the supplementary vector is connected to the original prediction network through the weight matrix WSV, and acts together with MLSTM to the regression layer, and finally affects the output of the prediction value. Considering the problem of machining efficiency, the tool wear value measured automatically is a low frequency sequence with a long interval, but the measured value in the supplementary vector (SV) is a high frequency sequence with a short interval, in order to ensure the consistency of the time and frequency of

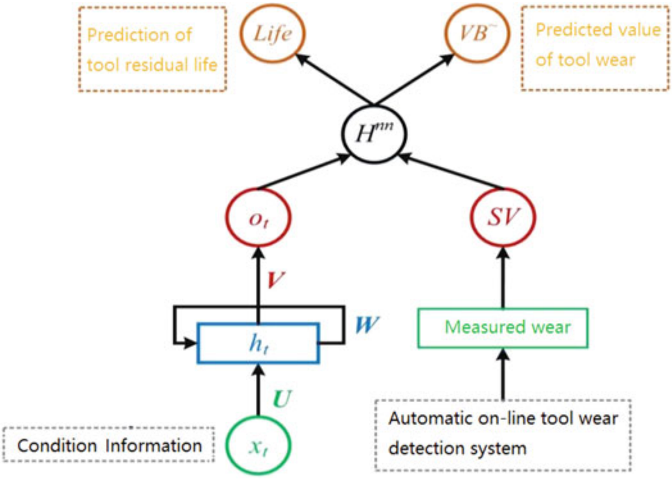


Fig. 22.5 Dynamic prediction model of tool life

the two-part network, the part of the wear quantity sequence of the supplementary vector (SV) which is short of the measured value is marked with −1, if the value of −1 means that there is no measured value at this time, then the model is not updated, if the value of −1 means that the value is measured wear, then update the network model accordingly.

22.5 Application Cases

22.5.1 Sample Data

The data used to train the model and verify the performance of the model came from the actual machining process. The cutting experiment was conducted on the machine tool of the turn-milling complex machining center, using the carbide integral milling cutter. The label of the data at each time point came from the results of the on-line automatic measurement of the tool wear based on machine vision.

The sample collected is the processing data of the full life cycle of each tool, that is, from no wear to the wear limit, and records the monitoring data of the entire processing process, the tool wear amount data and the use time data. The tool uses time data to be processed, and the remaining life of each time point is marked, thus completing the remaining life labeling of all samples.

All the data within 1 s is converted into a set of eigenvalues, which are the input to the prediction model. When collecting training data, if the machine vision system is invoked every 1 s to measure the actual tool wear, it is too time-consuming, and the wear amount can be approximately regarded as a linear increase between consecutive seconds. Therefore, in actual processing, the actual wear value is measured every 10 s, and the sample data between the two measurements is supplemented by the linear difference of the wear amount of the adjacent two points, so as to complete the wear amount labeling of all sample data.

In the cutting experiment, 16 tools were blunted, and the data of the first 10 tools were selected to train and verify the model. The cumulative service life of the tools was about 100 min. 90% of the samples were used for training the model, and the remaining 10% were used for testing the model.

22.5.2 Training Result

Many experimental data are used to train the dynamic prediction model of tool life and determine the relevant hyperparameters. Because the model needs to accurately fit the information of time series in the tool wear process, two-layer MLSTM network superposition method is used to enhance the expression ability of time series information, and the regression layer adopts double hidden layer structure. The training results of the tool life dynamic prediction model based on on-line learning are as follows.

For the 10 tools worn in the experiment, 90% of the data of each tool is used to train the model, and then the remaining data of the 10 tools are tested in turn. Finally, the mean square error of the test results on each tool is calculated. The calculated mean square error includes the Mean Square Error (MSE) of the average tool wear VB. MSE of maximum VB and MSE of remaining life. The MSE of tool life dynamic pre-measurement model based on on-line learning predicted results on the test data of 10 tools is: the mean MSE of the average VB is 3.32652e-3, the mean MSE of the maximum VB is 2.27653e-3 and the mean MSE of the remaining life is 2.57656e-3.

22.6 Conclusion

A dynamic on-line learning prediction model of tool life is proposed, in which signal features are selected as input data and tool wear and residual life are regarded as data labels. It can complete dynamic prediction of the tool remaining life under different working conditions. Finally, the model is used to calculate and analyze the tool remaining life, and the reliability of the model is verified.

Acknowledgements This work was supported by Scientific Research Fund of Hunan Provincial Education Department (No. 22C1097).

References

1. Li, W.B.: Study on the essence and influence of Industrial Revolution 4.0. China University of Mining and Technology. Master's thesis, Xuzhou (2019)
2. Yan, S., Hu, X.F., Liu, Y.C., et al.: Study on the self-adaptive prediction method for the life of semi-finish milling cutter for wheel groove. Mod. Mach. Tools Autom. Mach. Technol. **2016**(9), 95–98 (2016)
3. Wang, Z.W., Zhou, Y.J.: Research on key technology of tool wear monitoring system for NC machine tool. Internal Combust. Eng. Accessories **2021**(1), 59–60 (2021)
4. Tian, Y., Wang, W.H., Yang, L.M., Shao, W.T.: Multi-objective optimization of machining parameters based on tool wear state identification. J. Chin. Tianjin Univ. **55**(2), 166–173 (2022)
5. Wang, G.F., Dong, Y., Yang, K., et al.: Tool life prediction based on depth learning and particle filter. J. Tianjin Univ. Sci. Eng. **52**(11), 1109–1116 (2019)
6. Chan, W.: Research on the method of equipment health evaluation and residual life prediction based on LSTM network. University of Science and Technology of China, Hefei (2019)
7. Coble, J., Hines, J.W.: Incorporating prior belief in the general path model: a comparison of information sources. Nucl. Eng. Technol. **46**(6), 773–782 (2014)
8. Karandikar, J., Abbas, A., Schmitz, T.: Remaining useful tool life predictions in turning using Bayesian inference. In: 2013 International Manufacturing Science and Engineering, MSEC2013–1152. ASME, America (2013)
9. Yang, Y., Guo, Y., Huang, Z., Chen, N., Li, L., Jiang, Y., He N.: Research on the milling tool wear and life prediction by establishing an integrated predictive model. Measur.: J. Int. Measur. Confederation **145**, 178–189 (2019)
10. Huang, Q., Kang, Z., Zhang, Y., et al.: Tool remaining useful life prediction based on edge data processing and LSTM recurrent neural network. In: 2020 IEEE International Conference on Prognostics and Health Management (ICPHM), pp. 1–5. IEEE, America (2020)
11. Niu, J., Liu, C., Zhang, L., et al.: Remaining useful life prediction of machining tools by 1D-CNN LSTM network. In: 2019 IEEE Symposium Series on Computational Intelligence (SSCI), pp. 1056–1063. IEEE, America (2019)
12. Li, W., Liu, T.: Time varying and condition adaptive hidden Markov model for tool wear state estimation and remaining useful life prediction in micro milling. Mech. Syst. Signal Process. **131**, 689–702 (2019)
13. Liu, Y., Chang, Y., Liu, S., et al.: Data-driven prognostics of remaining useful life for milling machine cutting tools. In: 2019 IEEE International Conference on Prognostics and Health Management (ICPHM), pp. 1–5. IEEE, America (2019)
14. Kumar, A., Chinnam, R.B., Tseng, F.: An HMM and polynomial regression based approach for remaining useful life and health state estimation of cutting tools. Comput. Ind. Eng. **128**, 1008–1014 (2019)

Part IV
Mathematical Approaches and Applications

Chapter 23
Research on Power Quality Data Compression Algorithm Based on Complex Wavelet Transform

Taiqing Tang and Xiwen Wei

Abstract With the development of smart grid and new energy grid-connected technology, a large number of nonlinear, sensitive, and impact loads are widely used in modern power systems. In order to improve the shortcomings of large amount of data and low recognition rate in power quality signal disturbance identification based on wavelet, this paper proposes a power quality data compression algorithm based on complex wavelet transform. Through the improved energy threshold method, the threshold value is set by using the energy average correction coefficient to keep the compressed energy above 99%, thus ensuring that the distortion of the reconstructed signal is very small and adaptively eliminating the noise added to the disturbed signal. This method can perform real-time multi-resolution analysis on the data signal of power quality disturbance, and filter and retain the wavelet transform coefficients related to the disturbance at all scales, so as to realize real-time compression. The simulation results show that the operation efficiency of the power quality data compression algorithm based on complex wavelet transform has been greatly improved. At the same time, the recognition accuracy of the algorithm reaches 95.49%, which proves the feasibility of the data compression method. It is hoped that this research can provide some reference for the field of power quality data compression.

23.1 Introduction

In recent years, with the popularization and application of sensitive power electronic equipment in power grids in industry, the power quality problem is becoming more and more serious [1]. The ideal waveform of voltage and current in a power system should be a sine wave at power frequency, but due to the existence of a large number of disturbances and nonlinear loads in the actual system, the waveform always has

T. Tang (✉) · X. Wei
School of Electrical and Information Engineering, Heilongjiang University of Technology, JiXi City 158100, Heilongjiang, China
e-mail: 790693032@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_23

different degrees of non-sinusoidal distortion, thus reducing the power quality [2]. In practical systems, power quality problems are reflected in voltage quality problems, mainly involving voltage rise, voltage drop, voltage flicker, harmonic distortion, and so on [3]. Due to the wide application of frequency conversion speed regulation equipment, programmable logic controllers, various automatic production lines, and other electrical equipment sensitive to power quality, power quality problems such as overvoltage, undervoltage, voltage sag and bulge, voltage discontinuity, voltage fluctuation and flicker, various voltage transients, and harmonics have attracted extensive attention of power departments and users [4]. Compressed sensing (CS) has the characteristics of parallel compression and sampling, less sampling data, and high signal recovery. In recent years, some scholars have carried out corresponding research on the compression and reconstruction of power quality signals based on CS theory and successfully realized the compression and sampling reconstruction of power quality signals [5]. CS combines sampling and compression and obtains a small amount of data by sampling and compression with a standard far lower than Nyquist sampling frequency, so it can reconstruct the signal through the corresponding reconstruction algorithm, which has been highly valued in the field of signal detection and analysis.

In China, with the rapid development of the national economy and the formation and promotion of the power market in the future, the power quality problem has attracted and will increasingly attract the attention of electric workers in China [6]. At present, a large number of scholars have analyzed and studied the identification of power quality disturbance signals and the compression of power quality data. The effectiveness of wavelet transform in solving the data compression problem of power quality detection signal has been confirmed [7]. In order to make full use of the sparse compression ability of CS theory and the time–frequency localization ability of wavelet transform theory, in recent years, some scholars have combined wavelet transform with CS for data compression. Wavelet transform is the transformation of signal from the time domain to the time-scale domain. The wavelet transform analysis method has good time–frequency localization characteristics, it can focus on any details of the signal by adopting gradually fine sampling steps for different frequency components and can deal with weak or abrupt signals well, especially suitable for analyzing unsteady distorted waveforms [8]. Wavelet transform has the characteristics of time–frequency analysis, which can not only capture the time-domain changes of waveforms but also characterize the frequency-domain characteristics of signals. It is an effective tool for power quality analysis and data compression. In the past, some scholars used the Fourier sparse basis to extract the characteristics of power quality signal CS and the perceived sparse vector to identify it [9]. Because this method uses the Fourier sparse basis, it can't identify non-stationary signals such as transient oscillation and pulse well [10]. The application of wavelet transform can overcome the limitations of traditional Fourier transform and solve the problem that the Fourier transform is difficult to deal with. The commonly used real-valued wavelet transform can only provide the amplitude-frequency characteristics of the signal, but can't make full use of the phase information of the signal, which affects the accuracy of detecting the signal disturbance. In contrast, complex wavelet transform contains abundant phase information, which can provide not only the amplitude-frequency characteristics of the signal but also the phase-frequency characteristics of the signal, so it can better extract the characteristic information of the disturbance. Based on the above, in order to improve the shortcomings of large amounts of data and low recognition rate in power quality signal disturbance identification based on wavelet, this paper proposes a power quality data compression algorithm based on complex wavelet transform and makes simulation analysis.

23.2 Methodology

23.2.1 Relevant Theoretical and Technical Basis

The disturbance in the power system usually lasts for tens of ms, so a disturbance monitoring record will produce a large number of bytes of data. As a result, the recorded data capacity increases rapidly, which makes the storage cost increase obviously, so the compression problem of disturbed data becomes urgent [11]. Fourier transform is a tool to study the singularity of functions. It has many advantages, such as orthogonality and completeness, but when using FFT, two conditions must be met: the sampling theorem and the waveform being analyzed are stably changing with time period. In addition, in the past, people usually used the "wrapping" strategy to reduce data storage. This method used new data to overwrite the original data to recycle the memory, which kept a certain amount of storage space, but it was not compression in the usual sense of data compression. This paper presents a new power quality data compression algorithm based on complex wavelet transform. Different wavelets should be selected for different application problems. Among them, Daubechies wavelet [12] has been widely used in the analysis of power quality problems because it is sensitive to irregular signals. Wavelet transform has the property of spatial localization, and it is more effective to analyze the singularity position and singularity of signals by wavelet transform. In the case of a large amount of data, it is unrealistic to only improve the hardware transmission rate. It is necessary to find a data compression method with high compression performance, that is, the data compression ratio is high and the signal characteristics should be kept unchanged during compression, transmission, and recovery, so as to make further power quality analysis. The diagram of wavelet compression technology is as follows Fig. 23.1.

$c_0(n)$ and $c_0(n)$ are original disturbance signal and reconstructed signal, respectively; $d_1(n)$, $d_2(n)$, and $d_3(n)$ are the compressed signals of detail signals $d_1(n)$, $d_2(n)$, and $d_3(n)$ of wavelet decomposition at various scales.

Nyquist sampling theorem needs to satisfy that the sampling frequency is more than or equal to 2 times of the signal frequency so that the signal can be reconstructed without distortion [13]. For sparse signals, CS theory mathematically describes that all information of signals can be obtained by linear measurement far below Nyquist sampling frequency. The data compression of the power quality detection signal

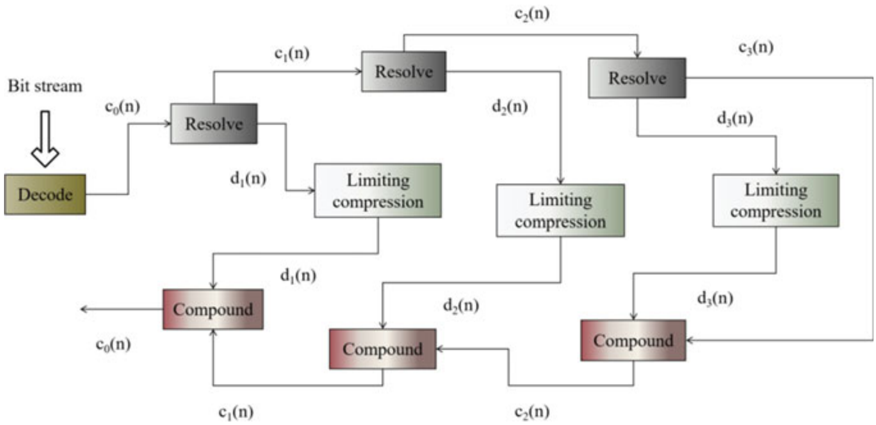


Fig. 23.1 Wavelet compression technique diagram

should meet the following two conditions: (1) There is a large compression rate under the condition of ensuring small signal distortion. (2) There is inevitably noise in the signal in the system, and the compression method should have the effect of eliminating noise, and the compressed signal should have a suitable signal-to-noise ratio. When multi-scale wavelet analysis is used, with the increase of scale, the modulus maximum of wavelet transform caused by disturbance is revealed, which can represent the local singularity of the signal, so wavelet analysis can be used to detect the disturbed signal. Using the properties of wavelet transform, this paper decomposes a given power quality disturbance data signal at different resolutions and scales. On each scale, the wavelet coefficients corresponding to a specific disturbance are much larger than those unrelated to the disturbance. In this way, the wavelet transform coefficients related to the disturbance on each scale are preserved together with their positions, while other irrelevant ones are discarded, thus achieving the purpose of compression. The original power quality data signal can be reconstructed and restored by these preserved wavelet transform coefficients.

23.2.2 *Power Quality Data Compression Algorithm Based on Complex Wavelet Transform*

Data compression mainly uses the time–frequency localization property of wavelet transform [14]. Different from other transforms, there are many choices for wavelet transform kernel-analysis mother wavelet. Using different analyses mother wavelet will produce different wavelet transform coefficients, and each analysis mother wavelet has its advantages and disadvantages. Therefore, reasonable selection should be made according to the occasion in practical application. Impulse interference may be the oscillation transient phenomenon caused by the excitation of natural frequency points in the power network. Because an impulse contains high-frequency components, its waveform will rapidly decay due to the characteristics of circuit components, and it will show different characteristics because of different observation points in the system. CS theory shows that when the sensing matrix meets the restricted isometry property (RIP) criterion, the sparse vector can retain the corresponding characteristics of the original signal, and the original signal can be accurately reconstructed from the sparse vector. Wavelet transform can decompose a given signal into multiple scales. On each scale, the information corresponding to a certain disturbance is retained, while other information is discarded. This can greatly reduce the storage data capacity. When using a one-dimensional wavelet to compress data, the signal compression rate is high when there is some noise, so the signal-to-noise ratio can't meet the requirements. There is a contradiction between them. In order to improve the compression ratio, this paper proposes a data compression method combining two-dimensional discrete wavelet and energy thresholds. This method uses a two-dimensional discrete wavelet to decompose the high-frequency and noise signals of the detection signal into eight directions: horizontal, vertical, and diagonal, and the energy of the signal in the frequency domain is very concentrated, which greatly eliminates the noise, thus improving the compression ratio. Because about 90% of the wavelet transform coefficients are discarded in the compression process, the reconstructed data signal inevitably has information loss compared with the original data signal. However, because most wavelet transform coefficients contain noise, the reconstructed data signal quality is very high. Because the wavelet transform coefficients related to the disturbance are preserved, the disturbance is preserved accurately. In a sense, the reconstructed signal has better quality than the original signal because it contains less electrical noise.

The measurement matrix is related to whether the effective information of the signal can be collected during compression sampling. In this paper, a sparse cyclic-structured measurement matrix is adopted. The matrix has the characteristics of low computational complexity and easy hardware implementation. In this paper, the energy threshold method is to set the threshold for each layer of wavelet coefficients, which can better retain the wavelet coefficients containing important information than the global threshold method. Let $\psi \in L^1 \cap L^2$ and satisfy

$$C_\Psi = \int \frac{\left|\hat{\Psi}(\omega)\right|^2}{|\omega|} \omega d < \infty \tag{23.1}$$

Then Ψ is a wavelet function. Continuous wavelet transform is defined as

$$Wf(s, x) = Wf_s(x) = f * \Psi_s(x) = \frac{1}{s} \int f(t) \Psi\left(\frac{x - t}{s}\right) dt \tag{23.2}$$

Among them:

$$\Psi_s(x) = \frac{1}{s} \Psi\left(\frac{x}{s}\right) \tag{23.3}$$

The essence of data compression is to get rid of all kinds of redundancy, such as spatial redundancy, temporal redundancy, information entropy redundancy, and so on, and keep the really useful information. From a mathematical point of view, it is to transform the original data into a data set that is as irrelevant as possible from a statistical point of view. This transformation should be carried out before the data is stored, processed, and transmitted, and then the compressed data needs to be decompressed to reconstruct the original data. Let $f(t)$ be the time-domain signal to be analyzed, then the binary orthogonal wavelet transform $f(t)$ is defined as

$$W_{DWT,\Psi}[f(m,n)] = 2^{-\frac{m}{2}} \int_{-\infty}^{+\infty} f(t)\Psi * \left(\frac{t - 2^m n}{2^m}\right) dt \tag{23.4}$$

where $*$ stands for complex conjugate; $m, n \in Z$ is the scale and time translation parameter; $\Psi(t)$ is a given basis function. The power quality disturbance signal is analyzed by a two-dimensional discrete wavelet. Calculate the wavelet coefficient energy E_j on each scale j, and set its average value as the threshold, but usually, the value is small. In the actual calculation, a correction coefficient is multiplied before the threshold. The value of the correction coefficient is obtained by using the ratio of the energy E'_j to E_j of the compressed data greater than 99%:

$$\begin{cases} \mu = \dfrac{\alpha E_j}{n} \\ \alpha = \dfrac{u}{\min|l_i - l_j|} \end{cases} \quad i, j \in Z, i \neq j \tag{23.5}$$

where α is the correction coefficient; n is the number of wavelet coefficients; l is the value of wavelet filter; Z is the length of the filter coefficient; u is the effective value of the signal voltage.

Data compression is realized by saving the wavelet transform coefficient values related to disturbance and discarding other wavelet transform coefficient values unrelated to disturbance. Usually, the maximum compression ratio and the minimum mean square error are contradictory and need to be weighed in practical application. The whole multi-resolution decomposition process is an orthogonal linear operation, so the orthogonal discrete wavelet transform itself is also an orthogonal linear operation operator. The signal reconstruction process is the inverse orthogonal discrete wavelet transform. Because of the orthogonality and linearity of the inverse transform, the inverse operation can be realized simply by reversing the whole process. In the traditional algorithm, the signal data acquisition process has the characteristics of long sampling time and high frequency, which makes the data amount of feature processing huge and the processing complexity high. In order to avoid collecting too much redundant disturbance information, this paper adopts the power quality data compression algorithm based on complex wavelet transform to compress and sample the disturbance signal, hoping to obtain less but useful sampling data. The method in this paper ensures the feasibility of data compression of power quality disturbance signal by two-dimensional discrete wavelet transform and ensures that the reconstructed signal has small distortion.

23.3 Result Analysis and Discussion

In order to verify the compression effect of the power quality data compression algorithm in this paper, various common transient disturbance signal and steady disturbance signal models are built on the MATLAB 2022 platform to verify the compression method proposed in this paper. In the simulation, the effective voltage is 220 V, the frequency is 50 Hz, the sampling frequency is 1024 Hz, and the sampling time is 0.4 s. The simulation operating system is Windows 11, the processor is Core i7 13700 k, the graphics card is RTX 3070, the memory is 16 GB, and the hard disk capacity is 1 TB. The selection principle of wavelet base is that a wavelet with good localization in both the time domain and frequency domain and sensitive to irregular signals should be used in data compression. In order to ensure the stability and distortion of signal reconstruction, it is necessary to choose a wavelet with a linear phase. After experimental comparison, this paper chooses the db4 wavelet, which is sensitive to mutation point and has low order, as the base wavelet of Mallat decomposition [15]. It is decomposed into three layers in two-dimensional discrete wavelet decomposition. When the wavelet packet is decomposed, it is decomposed into four layers. Taking the voltage sag signal as an example, its duration is 0.25 s, its frequency is 50 Hz, and there are 10 cycles and 500 data points. The compression ratio (M/N) is 26%, and the compression sampling and reconstruction experiments are carried out. The final simulation result is shown in Fig. 23.2.

On different scales, when the modulus of a wavelet transform coefficient is greater than the value of two adjacent points and at least strictly greater than the value of one of them, write down the corresponding wavelet transform coefficient of that point; with the increase of the number of scales, the modulus maximum of wavelet transform generated by the singular point of the disturbance signal gradually increases, so the time positioning of wavelet transform is the most accurate on a small scale. Here, we take the modulus maxima corresponding to the high-frequency subband d1 under the scale to locate the mutation time. In this paper, through the improved energy threshold method, the threshold value is set by using the energy mean correction coefficient to keep the compressed energy above 99%, thus ensuring that the distortion of the reconstructed signal is very small and adaptively eliminating the noise added to the disturbed signal.

In order to verify that the data compression method proposed in this paper has good compression performance, several kinds of common power quality disturbance signals generated by programmable ACSource are compressed. The data of all kinds of signals are obtained by mathematical model, the fundamental frequency is 50 Hz, which lasts for 200 cycles, and all of them contain Gaussian white noise with 10 dB intensity. The time taken by data compression methods of different algorithms is shown in Fig. 23.3.

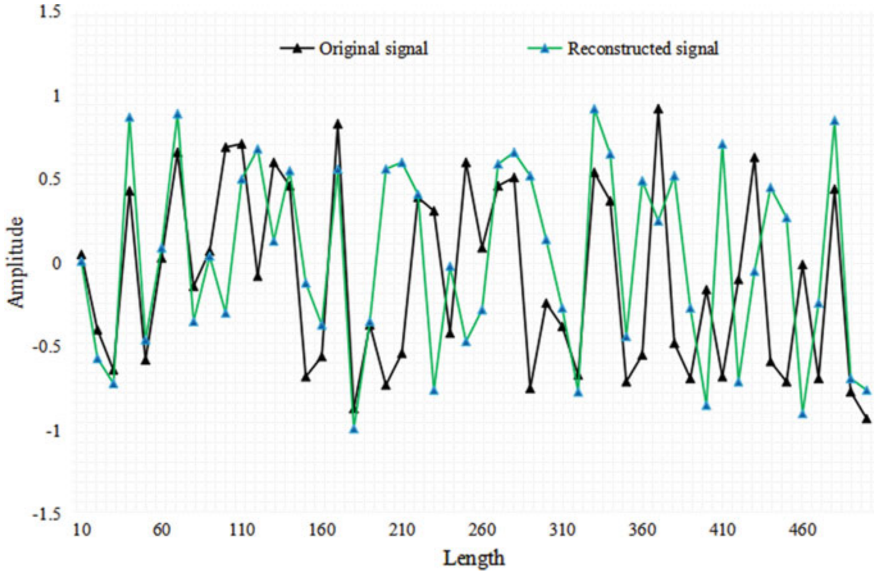


Fig. 23.2 CS reconstruction contrast diagram

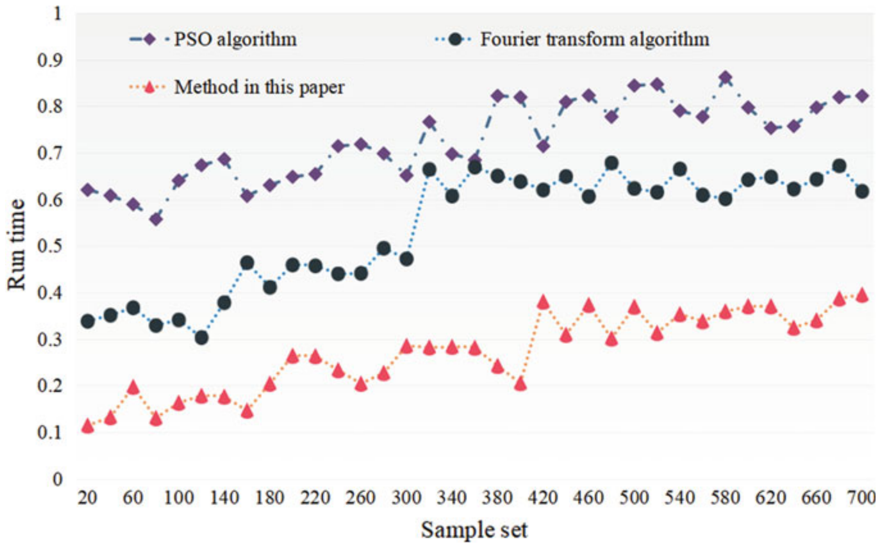


Fig. 23.3 Time for data compression of different algorithms

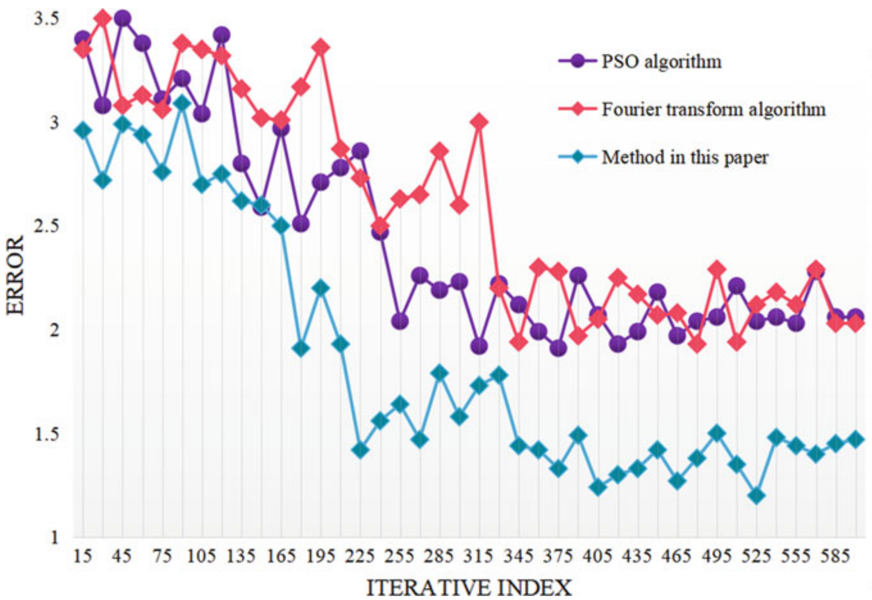


Fig. 23.4 Error comparison result

The compression ratio can be measured by the ratio of the number of bytes of disturbed data before and after compression. In order to evaluate the quality of the reconstructed signal, the mean square error between the original signal and the reconstructed signal can be calculated. The smaller the mean square error, the higher the quality of signal reconstruction. PSO algorithm, Fourier transform algorithm and this algorithm are used to compress power quality data. The result is shown in Fig. 23.4.

The experiment shows that the error of this method is small. In this paper, the disturbance information obtained from each scale of wavelet transform is taken as the characteristic quantity of the disturbance signal, and these characteristic quantities are taken as the input signals of the corresponding algorithm for the algorithm to identify the disturbance type. The compression ratio can be adjusted by controlling the compression rate according to the communication status of the power quality monitoring network. When the communication condition is good, the compression rate can be increased and more data can be transmitted. At this time, the signal-to-noise ratio of the reconstructed signal is high, and the characteristics of the original signal can be recorded accurately. When the communication condition is poor, the compression rate can be reduced and less data can be transmitted. At this time, the signal-to-noise ratio of the reconstructed signal is low, but the approximate characteristics of the original signal can still be recorded.

This paper makes full use of the correlation between two-dimensional matrix data, which makes the number of quantized bits in the frequency domain regular after wavelet transform, and greatly improves the energy concentration of high-frequency wavelet coefficients. In order to better verify the method proposed in this paper, the

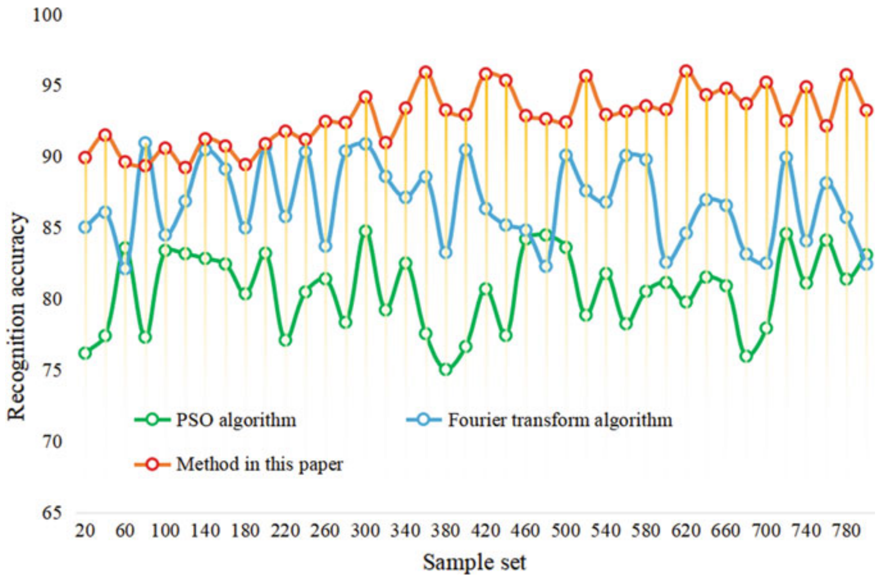


Fig. 23.5 Comparison results of algorithm recognition accuracy

data are obtained by simulation. The comparison results of recognition accuracy of different algorithms are shown in Fig. 23.5.

The simulation results show that the operation efficiency of the power quality data compression algorithm based on complex wavelet transform has been greatly improved. At the same time, the recognition accuracy of the algorithm reaches 95.49%, which has a certain superior performance. The length of data processed by the proposed method is only 30, which greatly reduces the amount of data to be processed, and it is more conducive to improving the accuracy of the algorithm because all the original useful feature information is saved with a very small amount of data.

23.4 Conclusions

With the improvement of sampling frequency and transmission accuracy, it is inevitable to record a large amount of data, which is very difficult to transmit to the dispatching center or store locally. Therefore, it is necessary to study a real-time compression method of power quality data with high efficiency and high compression ratio. Wavelet transform is widely used in the power system to suppress, detect, locate, quantitatively analyze, and classify various disturbance signals, compress, and store disturbance data, analyze and identify power system transients, and protect power system transients. In this paper, the application of wavelet transform in power quality analysis is summarized, and then a power quality data compression algorithm based on complex wavelet transform is proposed. This method has the characteristics of less sampling data, convenient processing, and simple feature extraction. Moreover, the length of data processed by the proposed method is only 30, which greatly reduces the amount of data to be processed, and it is more conducive to improving the accuracy of the algorithm because all the original useful feature information is saved with a very small amount of data. The simulation results show that the recognition accuracy of the algorithm reaches 95.49%, and it has strong robustness and good accuracy. The simulation results prove the feasibility of the data compression method, which has important practical significance. With the further development of wavelet technology and the appearance of wavelet basis functions with better performance, wavelet transform must be more widely used in power quality disturbance analysis and detection and data compression.

Acknowledgements Young Elite Scientist Sponsorship Program by Heilongjiang Province (2022QNTJ020).

References

1. Huang, N., Wang, D., Lin, L., et al.: Power quality disturbances classification using rotation forest and multi-resolution fast S-transform with data compression in time domain. IET Gener. Transm. Distrib. **13**(22), 5091–5101 (2019)
2. Silva, L., Kapisch, E.B., Martins, C., et al.: Gapless power-quality disturbance recorder. IEEE Trans. Power Deliv **32**(2), 862–871 (2017)
3. Acharya, S., Demarco, C.L.: Low-loss image-based compression for synchrophasor measurements. IET Gener. Transm. Distrib. **14**(20), 4571–4579 (2020)
4. Yin, L., Qi, M., Lei, G., et al.: Reconstruction of compressed sensing power quality data based on overcomplete dictionary. Power Syst. Prot Control **46**(8), 88–94 (2018)
5. Chakravorti, T., Patnaik, R.K., Dash, P.K.: Detection and classification of islanding and power quality disturbances in microgrid using hybrid signal processing and data mining techniques. IET Signal Proc. **12**(1), 82–94 (2017)
6. Zjavka, L.: Power quality statistical predictions based on differential, deep and probabilistic learning using off-grid and meteo data in 24-hour horizon. Int. J. Energy Res. **2022**(8), 46 (2022)
7. Wendy, M., Liu, A., Zakaria, A., et al.: Power quality and rooftop-photovoltaic households: An examination of measured data at point of customer connection. Sustainability **10**(4), 1224 (2018)
8. Jasinski, M., Sikorski, T., Borkowski, K.: Clustering as a tool to support the assessment of power quality in electrical power networks with distributed generation in the mining industry. Electr. Power Syst. Res. **166**(1), 52–60 (2019)
9. Zhu, L., Hill, D.J.: Data/model jointly driven high-quality case generation for power system dynamic stability assessment. IEEE Trans. Industr. Inf. **2022**(8), 18 (2022)
10. Guo, S., Liu, P., Li, Z.: Enhancement of performance monitoring of a coal-fired power plant via dynamic data reconciliation. Energy **151**(5), 203–210 (2018)
11. Xiao, F., Lu, T., Ai, Q., et al.: Design and implementation of a data-driven approach to visualizing power quality. IEEE Trans. Smart Grid **11**(5), 4366–4379 (2020)

12. Nastos, C.V., Saravanos, D.A.: Multiresolution Daubechies finite wavelet domain method for transient dynamic wave analysis in elastic solids. Int. J. Numer. Meth. Eng. **122**(23), 7078–7100 (2021)
13. Malik, H., Almutairi, A., Alotaibi, M.A.: Power quality disturbance analysis using data-driven EMD-SVM hybrid approach. J. Intell. Fuzzy Syst. **2021**(6), 1–10 (2021)
14. Liu, Y., Tang, W., Luan, Y.: A new reconstruction algorithm based on temporal neural network and its application in power quality disturbance data. Measur Control **54**(7–8), 1199–1207 (2021)
15. Yelampalli, P.K.R., Nayak, J., Gaidhane, V.H.: Daubechies wavelet-based local feature descriptor for multimodal medical image registration. IET Image Proc. **12**(10), 1692–1702 (2018)

Chapter 24
Cultural Heritage Protection of Qi Great Wall Based on Digital Cultural Relics Technology

Xiaoxu Li

Abstract With the rapid development of informatization and digitalization, the protection of the cultural heritage of Qi Great Wall has also begun to change from information preservation to information management. In this paper, the protection of the cultural heritage of Qi Great Wall based on the digital technology of cultural relics is studied. Firstly, the fast and precise 3D reconstruction technology of Qi Great Wall is analyzed, and the camera is calibrated by using the accurate feature points on the reference objects around Qi Great Wall. Rebuild the 3D feature points of Qi Great Wall, and carry out 3D modeling based on the feature points. Secondly, the construction and digital protection of Qi Great Wall information model are put forward, and the 3D point cloud data on its surface are collected by Farofocus3D phase LiDAR laser scanner. Based on the point cloud data of Qi Great Wall, the BIM model is established on Revit platform to realize 3D information retention and expression of Qi Great Wall. Finally, a digital Qi Great Wall exhibition system is established, which enables users to acquire online knowledge about related services, exhibits, cultural relics, and related documents of Qi Great Wall through the Internet.

24.1 Introduction

Qi Great Wall is a large-scale military defense project built by Qi in the Eastern Zhou Dynasty, and it is the largest architectural relic in Shandong. Qi Great Wall is the largest historical and cultural site of Qi State and the largest existing hardware of Qi culture. It is the product of Qi's unique geographical position, landscape situation, geopolitics, economic structure, military strategy, values, and developed science and technology, and it is an ideal perspective for us to re-examine Qi's history and culture. However, under the impact of all kinds of quick success and instant benefit

X. Li (✉)
Qilu Institute of Technology, Ji Nan City, Shandong 250000, China
e-mail: 790693032@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_24

economic activities, the present situation of Qi Great Wall site is not optimistic. Digitalization of cultural relics plays an extremely important role in protecting, studying, and displaying cultural relics, and establishing digital models for physical cultural relics is an important content of digitalization of cultural relics.

Heritage monitoring is an important content and mandatory requirement in the World Heritage Convention and Operational Guidelines, and it is also an important part of the virtuous circle heritage management mechanism, which mainly includes regular monitoring and reactive monitoring. Together with the heritage impact assessment, it is paid attention to by the international heritage protection field [1]. Nowadays, the digitalization of cultural relics has many characteristics, such as recording, preservation, reading, retrieval, sharing, restoration, simulation, reproduction, etc. Using these new technologies of digitalization of cultural relics will undoubtedly play a revolutionary role in excavating rich and colorful material culture and inheriting material culture and art [2, 3]. Digitalization of cultural relics refers to obtaining the data information of cultural relics such as shape, texture, texture, and material by some technical means, and storing it in the computer as the research data of corresponding professionals. With the help of advanced multimedia and Virtual Reality (VR) technology, it is of great practical significance to digitally display and effectively protect China's ancient cultural heritage. Serve the protection, development, and utilization of cultural relics. The development of information technology, especially the research and application of digital photography, 3D information acquisition, VR, multimedia, and broadband network technology, provides a solid technical foundation for the digital research of cultural relics [4, 5].

Qi Great Wall is a cultural heritage of Shandong, China, and the world, and it is our national treasure. Cultural and cultural relics departments along the Qi Great Wall should, in combination with local conditions, strengthen the publicity of national laws and regulations on the protection of cultural relics and historic sites in towns and villages on both sides of the Great Wall site in a planned and step-by-step manner, so as to convey the significance of protecting the Great Wall. The cultural heritage of Qi Great Wall can't live forever and can't be regenerated. In the long river of history, it has been eroded by the natural environment and destroyed by human activities. What we can do is to try our best to protect it, try our best to delay its disappearance, and inherit its spirit. With the rapid development of informatization and digitalization, the development of digital technology of cultural relics related to the protection of Qi Great Wall has been widely concerned, and more modern technologies have been applied to the protection of Qi Great Wall, which has greatly improved the work efficiency, and the protection of Qi Great Wall's cultural heritage has also begun to change from information preservation to information management [6]. In this paper, the protection of the cultural heritage of Qi Great Wall based on the digital technology of cultural relics is studied. Through the research of technical principles and the discussion of details, the purpose of digital protection of the cultural heritage of Qi Great Wall is achieved, and the life cycle management model of the cultural heritage of Qi Great Wall is created, which is beneficial to the repair and preservation of the damaged Qi Great Wall.

24.2 The Application of Digital Technology of Cultural Relics in the Protection of Cultural Heritage of Qi Great Wall

24.2.1 Fast and Fine 3D Reconstruction of Qi Great Wall

3D reconstruction is the preservation and virtual reproduction of 3D information, including all kinds of geometric data of the object. It is very important that Qi Great Wall has a long history of construction, and there are few records about the layout, so it is necessary to re-survey and obtain the topographic layout data of the scene through field measurement or looking for records about the surveying and mapping department. The geometrical features of a building mainly include the physical dimensions of the shape of the building. In 3D reconstruction of general modern buildings, relevant data can be directly obtained through CAD drawings or measurements [7, 8].

The method used in this paper is to shoot the details of buildings from multiple angles with a digital camera, take the structure of the known data in the scene as a reference, and estimate the relevant data according to the proportional relationship between buildings. In the actual modeling process, the scale size is often more important than the actual size. Qi Great Wall parametric modeling method is shown in Fig. 24.1.

Aiming at the cultural relics with complex modeling in Qi Great Wall Group, this method can quickly achieve a realistic modeling effect. To obtain the geometric model of an object in 3D space, it can be achieved by reconstructing spatial feature points. In this paper, the structure of Qi Great Wall is reconstructed, and photos are taken from all angles with digital cameras, and a certain number of feature points are marked. The camera is calibrated by using accurate feature points on the reference objects around Qi Great Wall. Rebuild the 3D feature points of Qi Great Wall, and carry out 3D modeling based on the feature points [9].

Let C_1, C_2 be the corresponding local coordinate system when the camera is located at two different points, and the transformation parameters with the world coordinate system are R_1, T_1, R_2, T_2, respectively. For any point P in the world coordinate system, let its coordinates in the world coordinate system be $X_w(x_w, y_w, z_w)$ and its coordinates in the C_1, C_2 coordinate system be $X_{c1}(x_{c1}, y_{c1}, z_{c1})$, $X_{c2}(x_{c2}, y_{c2}, z_{c2})$, respectively, and there is the following relationship:

$$\begin{cases} X_{c1} = R_1 X_w + t_1 \\ X_{c2} = R_2 X_w + t_2 \end{cases} \tag{24.1}$$

The transformation relationship between the two coordinate systems of C_1, C_2 can be obtained by eliminating X_w in the above formula:

$$X_{c1} = R_1 R_2^{-1} X_{c2} + t_1 - R_1 R_2^{-1} t_2 = R X_{c2} + t \tag{24.2}$$

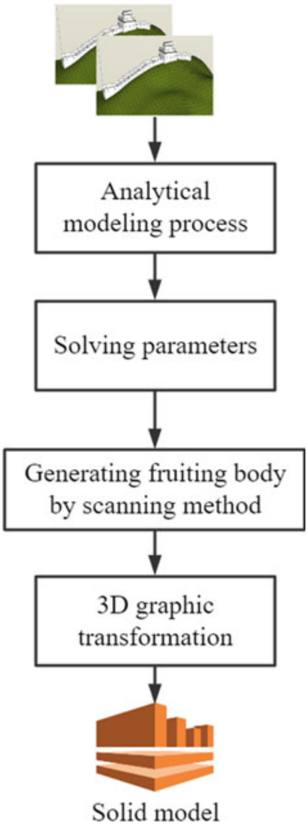


Fig. 24.1 Qi Great Wall parametric modeling method

The coordinates of any point in the world coordinate system in the C_1 coordinate system can be determined by setting one of the two coordinate systems as the standard coordinate system, its coordinate origin (0, 0, 0), and the directions of the three coordinate axes as the reference direction.

According to the geometric relationship of vanishing point, the vanishing point can be used to calibrate the image shot by a common digital camera, and the internal azimuth parameters and external azimuth parameters of the camera can be obtained, so as to restore the 3D scene of the two-dimensional image [10]. The principle of vanishing point calibration is shown in Fig. 24.2.

Eleven degrees of freedom in the projection matrix P can be decomposed into the rotation matrix R, translation matrix T and a 3*3 camera calibration matrix K of the camera relative to the spatial coordinate system.

$$P = M\left[I\middle|M^{-1}P_4\right] = KR[I|-C] = K[R, T] \tag{24.3}$$

In perspective projection, a group of parallel lines that are not parallel to the projection plane converge to a point after projection, which is called the vanishing

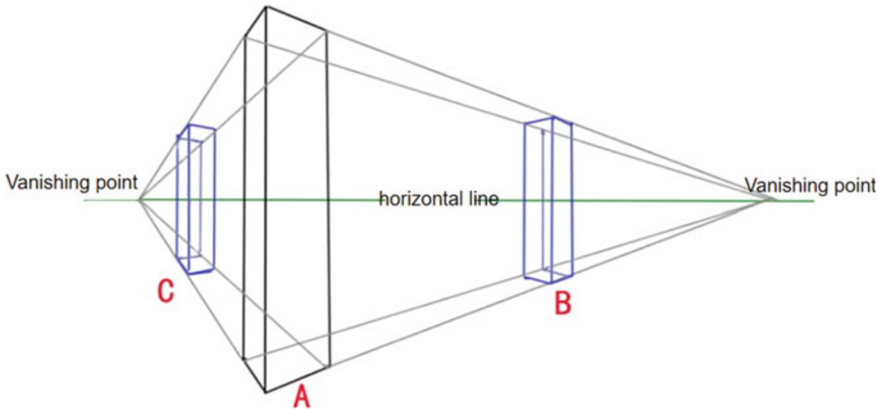


Fig. 24.2 Principle of vanishing point calibration

point. There can be infinitely many vanishing points. With the introduction of the 3D rectangular coordinate system, vanishing points parallel to the three coordinate axes are called principal vanishing points, and there are at most three principal vanishing points.

In fact, the 3D point cloud model of Qi Great Wall only contains the spatial coordinate information of the surface measuring points, and the corresponding 3D geometric model must be reconstructed from the surface to realize the digital preservation and virtual display of the site. At present, the gridding of large-scale point cloud data is mainly realized by triangulation of point cloud [11]. Point cloud data registration is a key step in the whole process of point cloud data processing. The point cloud registration process refers to two groups of point cloud data with repeated areas obtained by 3D laser scanning and sampling, and then an appropriate coordinate transformation is determined, so that the two groups of point cloud data are aligned and merged in a unified coordinate system to obtain a complete point cloud data model of the detected object.

In the process of surveying and mapping, due to the influence of weather factors, such as smog, and some problems such as inaccessible during the instrument erection, the collected point cloud data of Qi Great Wall are often accompanied by data unrelated to the characteristics of Qi Great Wall. These data points cannot be removed manually, so these redundant noise points must be removed better by algorithm processing.

In order to remove small-scale noise data, this paper adopts point cloud bilateral filtering method. In the process of image denoising, the bilateral filtering function can effectively denoise the surface of a 3D model in space, which can not only keep outstanding set feature information but also reduce the iterative calculation and avoid 3D point cloud data from being smoothed. The bilateral filtering algorithm is applied to the denoising process of 3D point cloud data, and the definition is shown in Formula (24.4):

$$g := g + am \tag{24.4}$$

Among them, g represents the data point, a represents the bilateral filtering weight factor, and m represents the normal direction of the data point g. If the normal direction is unknown, the principal component analysis method and the least square fitting plane normal direction can be used to estimate the normal direction of the data point g.

The 3D model based on point cloud data has good geometric accuracy, but in order to meet the needs of 3D model visualization and restore the surface details of the 3D landscape, texture mapping technology is needed to add real colors to the 3D model. Using Photoshop to stretch and deform the texture collected in the early stage to make it conform to the texture mapping specification, and the key point of processing is to make the texture size consistent with the real size. In order to make colors more uniform and harmonious, screen colors are extracted by using the color editor and applied to texture images with the same color.

24.2.2 Construction and Digital Protection of Qi Great Wall Information Model

As a professional building information management technology, BIM can effectively and accurately extract the component information of the cultural heritage of Qi Great Wall, which is convenient for the management and preservation of Qi Great Wall. Using BIM technology, the information of Qi Great Wall components can be extracted efficiently and accurately, and the scientific and systematic digital protection of Qi Great Wall can be improved. For the establishment of Qi Great Wall BIM model, the most important thing is the establishment of the geometric information model. Li DAR remote sensing can obtain point cloud data quickly and efficiently, which provides mature technical conditions for Qi Great Wall to obtain the high-precision geometric model.

Cultural heritage has its own life cycle, and inevitably faces the problems of weathering, erosion, and fading with the passage of time, and gradually loses its original appearance. Every time cultural relics protection and repair or other intervention measures such as archaeology occur, it is an opportunity to establish digital heritage information through the whole process record, and once it is missed, it will not be restored. Digital archives provide the possibility for cloud sharing based on computer science and internet technology and realize the functions of visual, structured, and online real-time entry and access of recorded archives on the basis of digitalization.

The classification system based on the digitalization of Qi Great Wall cultural heritage information is to present all the valuable information of Qi Great Wall cultural heritage in digital form, which is the first-hand material for the study of Qi Great Wall cultural heritage and the most basic metadata information. It is a key point in the research of digital protection of Qi Great Wall cultural heritage to build an information model with Qi Great Wall cultural heritage entities as the core. Qi Great Wall cultural heritage information can be divided into two levels of information model after digitization, one is a logical model, and the other is a physical model. With the information model based on the cultural heritage entity of Qi Great Wall, we can organize all kinds of related information on the cultural heritage of Qi Great Wall organically, and complete the extraction, sorting, storage, retrieval, and use of these information.

Based on the imported point cloud data, the built family components are loaded into the project, and the point cloud data provides the location reference for the "family" components [12]. According to the point cloud data, the precise location of family construction is determined, so as to build an accurate 3D model. The 3D point cloud data on its surface were collected by using the Faro Focus 3D phase LiDAR laser scanner. After data processing, a BIM model was established on the Revit platform based on the point cloud data of Qi Great Wall, and the 3D information of Qi Great Wall was retained and expressed.

Open the Qi Great Wall BIM model in Revit, and export the Qi Great Wall BIM model to a new data source by using plug-ins. After the data is completely exported, the 3DGIS software SuperMap is used to open the data source in the scene, so that the whole 3D effect can be browsed and the attribute information can be obtained by clicking on a single object. As shown in Fig. 24.3, the effect diagram after importing SuperMap is shown.

Through the research on the protection of the cultural heritage of Qi Great Wall, this paper summarizes the role of digital technology in each stage, and combined with the regional characteristics of Qi Great Wall, draws a set of ideas and methods for the research on the protection strategy of cultural heritage of Qi Great Wall based on digital technology, as shown in Fig. 24.4.

In the process of digital tracking and recording, in addition to the 3D spatial data of the Great Wall, the technical team also recorded the panoramic VR data regularly, so as to preserve the changes in the surrounding environment and take the recording process as the monitoring process of the Great Wall heritage. Content should also be open and inclusive. For example, the contents of public communication and media reports that were not included in the archives system should be included in the process of digital recording of cultural heritage. Let the public return to the state before archaeological excavation and some important changes after that, and understand the protection concept and research process of cultural heritage.

24.3 Digital Qi Great Wall Display System

VR technology is a new technology that has gradually attracted attention from all walks of life and has been further applied in the commercial field since the 1990s. In this research, OpenGL is selected to develop the corresponding digital Qi Great Wall display system. With the rapid development of multimedia technology and computer hardware technology, VR technology has a wide range of applications and influences

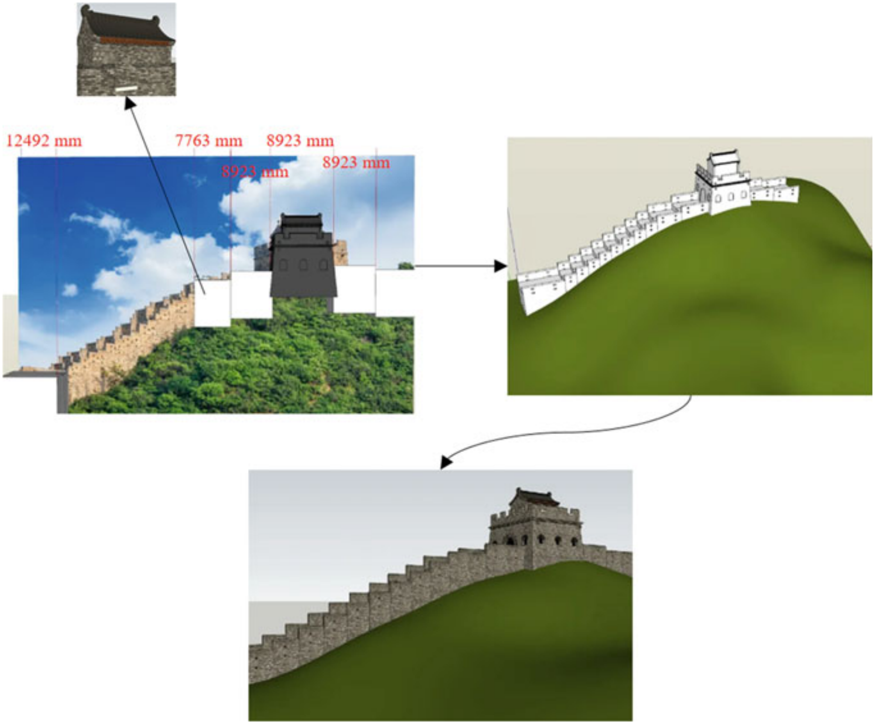


Fig. 24.3 Effect diagram after importing SuperMap

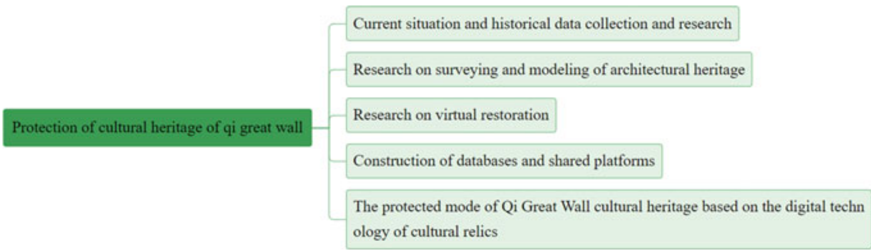


Fig. 24.4 Qi Great Wall cultural heritage digital protection framework

in the field of education. At present, it is more widely used in scientific research, virtual simulation campus, virtual teaching, virtual experiment, education, and entertainment. Because architecture is a combination of art and reality, its professional characteristics determine that virtual technology has an important influence in this field. At present, the architecture schools of major universities are widely developing the research and application of related technologies.

At present, the research on the protection strategy of the cultural heritage of Qi Great Wall based on the concept of digital protection is mostly preserved in the form of words, pictures, and videos. By combining these digital materials with architectural heritage entities, a large number of complex information in architectural heritage is integrated into a clear structure system. In the heritage space, the content and expression of history and culture are still very thin, and the expression method is relatively traditional, lacking the intervention of modern scientific and technological means. This case will be based on the actual needs of the protection of cultural heritage of Qi Great Wall, and use the technical characteristics of OpenG to establish a digital Qi Great Wall display system. Combined with the advantages of multi-layer architecture, it is very necessary to adopt this structural design for the cultural heritage protection and exhibition system of Qi Great Wall. The hierarchical structure diagram of this system is shown in Fig. 24.5.

Web application layer module is to provide users with a good operating interface. The main requirements for this module are user-friendly, concise, detailed information, and clear semantics. This layer communicates with the business logic layer module.

The construction of the web display system is based on the Qi Great Wall knowledge center. It enables users to acquire online knowledge of related services, exhibits, cultural relics, and related documents of Qi Great Wall through the Internet. It mainly

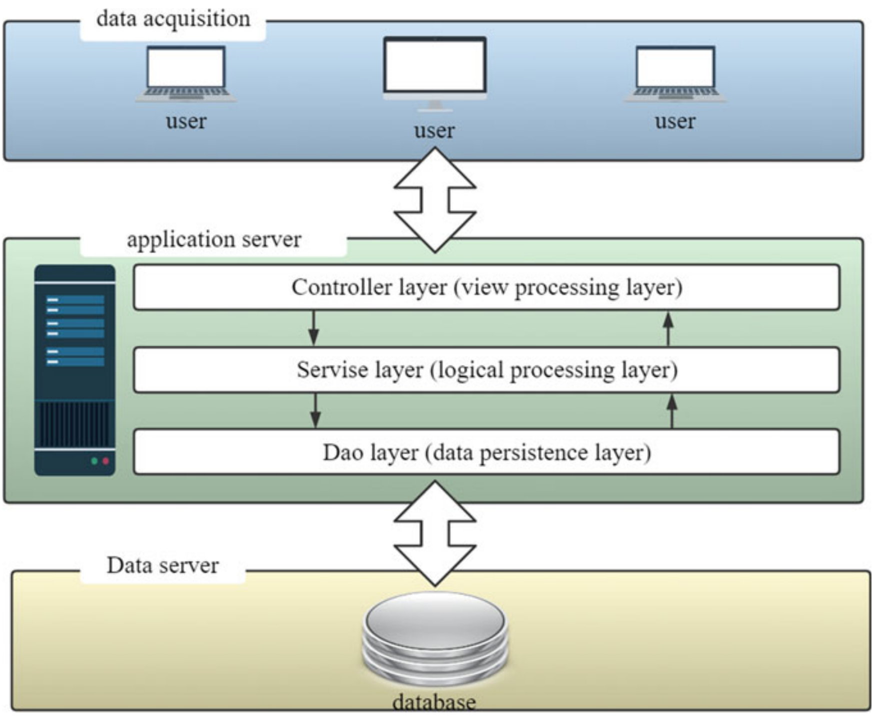


Fig. 24.5 System hierarchy diagram

includes Qi Great Wall category, exhibition, collection selection, service guide, and communication platform that can directly interact with users.

24.4 Conclusion

Qi Great Wall is the largest historical and cultural site of Qi State and the largest existing hardware of Qi culture. With the rapid development of informatization and digitalization, the development of digital technology of cultural relics related to the protection of Qi Great Wall cultural heritage has been widely concerned, and more modern technologies have been applied to the protection of Qi Great Wall, which greatly improves the work efficiency. In this paper, the protection of the cultural heritage of Qi Great Wall based on the digital technology of cultural relics is studied. Through the research of technical principles and the discussion of details, the purpose of digital protection of the cultural heritage of Qi Great Wall is achieved, and the life cycle management model of the cultural heritage of Qi Great Wall is created, which is beneficial to the repair and preservation of the damaged Qi Great Wall.

References

1. Yan, W.: On the construction of the heritage corridor of Qi Great Wall. Acad. J. (11), 7 (2015)
2. Sun, W.T.: Landscape painting of the ruins of Qi Great Wall. Art Watch (4), 4 (2018)
3. Tao, W., Wanli, Zhenjiang, L.: Study on protection scheme of long-span expressway tunnel undercrossing the ruins of Qi Great Wall. Highway **61**(8), 5 (2016)
4. Chen, H.Y.: Analysis of the current artistic protection of architectural cultural heritage. Art Hundred Schools **34**(4), 9 (2018)
5. Wu, Z.D., Guo, W.M., Ai, X.Q.: Exploration on digital protection of ethnic architectural cultural heritage. Guizhou Ethnic Stud. **37**(4), 4 (2016)
6. Stone, T.: Research on the architectural cultural heritage of tourist cities and the conservation and renovation strategies of historical blocks. Econ. Geogr. **30**(3), 6 (2010)
7. Dan, J.X.: The challenges faced by the renovation of ancient buildings in the Forbidden City. Outlook (49), 2 (2015)
8. Divya, U.J., Kim, H.S., Kim, J.I.: Animage-based approach to the reconstruction of ancient architectures by extracting and arranging 3D spatial components. Front. Inform. Technol. Electron. Eng. **16**(1), 12–27 (2015)
9. Castilla, F.J., Ramón, A., Adán, A., et al.: 3D sensor-fusion for the documentation of rural heritage buildings. Remote Sens. **13**(7), 1337 (2021)
10. Fais, S., Casula, G., Cuccuru, F., et al.: An innovative methodology for the non-destructive diagnosis of architectural elements of ancient historical buildings. Sci. Rep. **8**(1), 4334 (2018)
11. Bonali, E., Pesci, A., Casula, G., et al.: Deformation of ancient buildings inferred by terrestrial laser scanning methodology: the Cantalovo church case study (Northern Italy). Archaeometry **56**(4), 703–716 (2013)
12. Li, Y., Dai, A., Guibas, L.J., et al.: Database-assisted object retrieval for real-time 3D reconstruction. Comput. Graph. Forum **34**(2), 435–446 (2015)

Chapter 25
Design of Fault Monitoring Algorithm for Electrical Automation Control Equipment Based on Multi-Sensor

Xiao-Rong You

Abstract A single sensor can only monitor a single state or environmental parameter. The operation of electrical automation control equipment conforms to the barrel effect, that is, a certain factor exceeds the maximum operating limit that electrical automation control equipment can bear, and the fault will inevitably occur. The basic purpose of information fusion is to make full use of multiple sensor resources. Based on multi-sensor data fusion technology, this paper introduces data fusion technology into the field of fault monitoring of electrical automation control equipment. An end-to-end adaptive fault monitoring algorithm for electrical automation control equipment is proposed based on multi-head attention mechanism and CNN (Convolutional Neural Network). This algorithm has stronger robustness and, combined with CNN, improves the accuracy of network fault classification. The research results show that when the signal-to-noise ratio of the noise signal is 7 dB, the accuracy of multi-attention CNN prediction has reached 97% and has remained at 100% since then. The experimental results show that the multi-attention CNN algorithm has better noise adaptability and stronger robustness in noise interference environment compared with other mainstream algorithms. The fault monitoring algorithm of electrical automation control equipment proposed in this paper can help share the work pressure of staff.

25.1 Introduction

Due to the high complexity and variability of modern mechanical manufacturing system and the correlation and coordination between various parts of the machining process, the traditional single sensor single factor monitoring and single model processing evaluation have become insufficient. Automation equipment has developed to a wide extent, which has a great impact on the production and operation

X.-R. You (✉)
ChongQing Industry and Information Vocational College, Chongqing, China
e-mail: 872067079@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_25

efficiency of enterprises, and plays a vital role in both product quality and quantity. However, before the real functions of electrical automation equipment are fully displayed, common faults and maintenance skills have always been an urgent research topic.

A single sensor can only monitor a single state or environmental parameter. The electric control system operates according to the "barrel" phenomenon, and once it exceeds the limit that the electric control system can bear, it will lead to the failure of the electric control system [1, 2]. Multi-sensor information fusion technology refers to the information processing process of automatically analyzing and synthesizing information and data from multi-sensors or multi-sources according to certain algorithms and criteria through computer technology, so as to realize the required decision-making and estimation [3, 4]. Its fundamental goal is to maximize the role of multiple sensors, and reasonably combine these sensors, so that the information they can obtain is far greater than that provided by a single information source, so as to achieve the goal of multi-sensor cooperation and cooperation, thus improving the overall performance of the detection system [5]. Through wireless data transmission, the alarm signal is sent out in time and the working state of the water pump and dehumidifier is controlled, so as to realize remote monitoring and distributed control, thus solving the problem in real time, quickly, and effectively.

Electrical automation control equipment technology can improve equipment stability, improve equipment work efficiency, promote electrical reform, and promote the development of the electrical automation industry. Looking at the current operation of equipment, there are significant differences between automation equipment, so there must be some differences in the structural system, which leads to the system being too cumbersome [6]. At this time, once there is equipment failure, the relevant personnel must cut off the power supply device of the automation equipment at the first time, so as to avoid the occurrence of operational errors, thus inducing other equipment to fail at the same time. On the basis of multi-sensor data fusion technology, this paper introduces data fusion technology into the field of fault monitoring of electrical automation control equipment and makes full use of the redundancy between models to improve the reliability of fault prediction for complex systems.

25.2 Research Method

25.2.1 Multi-Sensor Network Node Design for Electrical Automation Control Equipment

Automatic operation of electronic control equipment, whose functions include monitoring, system detection, and automatic protection of electronic control equipment. In the production process of electronic control equipment, we should pay attention to the selection of components and the optimization of equipment performance, and at the same time strengthen the economic benefits of enterprises. Therefore, it is suggested that the classification inspection method should be applied in maintenance, and the maintenance personnel should classify the faults after understanding the principle of equipment operation. For example, if the equipment operation environment causes faults, the main factors include temperature and humidity, improper operation of staff, etc., the fault level can be determined by carrying out maintenance, and the targeted maintenance method for each level of faults is conducive to improving the efficiency of equipment fault maintenance [7, 8].

Electrical automation control equipment will be constrained and restricted by many unpredictable subjective and objective conditions when it is used on the right track. For example, the damaged insulator can't be used normally, the current value and voltage value can't be balanced with the average value between them, and there are problems in contact. In the process of actual operation of electrical automation control equipment, the phenomenon of sudden rise of voltage and current occurs from time to time, which will cause the chip temperature to reach an invaluable fault. At the same time, too frequent operation of electrical automation control equipment in the process of starting and stopping operation may lead to uncontrollable total current, resulting in major accidents.

WSN (Wireless sensor network) is composed of base stations and a large number of wireless sensor nodes, which play an important role in the whole network system. Sensor nodes must be closely combined with the perceived scene to perceive the changes in the external environment very carefully. It is the close combination of all sensor nodes and the perceived scene that makes the perceived object have a macro and micro understanding [9]. WSN nodes for electrical automation control equipment have high requirements in energy consumption, expansibility, flexibility, stability, security, and embeddedness. The design block diagram of this paper is shown in Fig. 25.1:

It consists of a sensor module, a digital signal processing module, and a wireless network operation module. Each module defines the corresponding interface, which

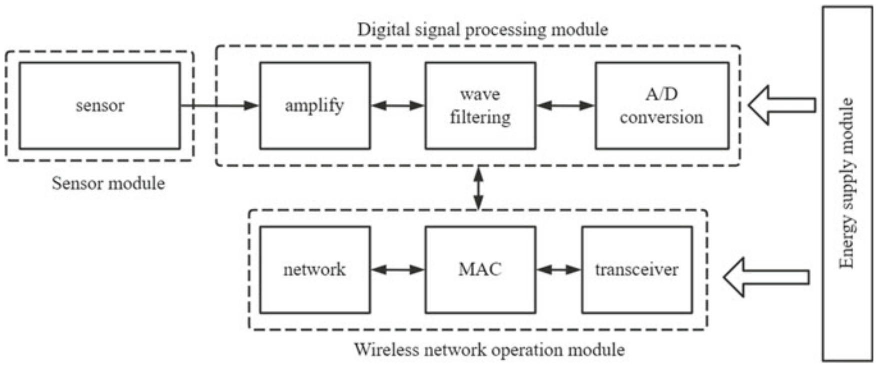


Fig. 25.1 Hardware framework design

makes the modules relatively independent, which is beneficial for network application layer developers to independently and completely apply research according to different application occasions.

According to the interface provided by the digital signal processing module, the sensor module is designed by professional researchers for different electrical automation control equipment and applications. So that this part of the researchers do not have to think too much about the problems related to wireless communication.

In the design concept, this paper completely separates the collection, processing, and storage of sensor data from the wireless communication module. The researchers of online monitoring of electrical automation control equipment design corresponding sensor modules for different electrical automation control equipment and applications, and configure and program the digital signal processing module, so as to realize the corresponding monitoring task of electrical automation control equipment. It makes the application development of wireless sensors more focused on the monitoring object, thus further simplifying the development process and development cycle, and greatly improving the scalability and flexibility of the node.

25.2.2 Multi-Sensor Information Fusion

Electrical automation control equipment faces many problems in normal startup, but when it can't be started normally, the faults are mainly manifested in mechanical and electrical faults. The complexity of electrical engineering is high, which involves many contents such as electric energy configuration, transmission, and manufacturing, and automation system plays an important role in it. In order to ensure the safety, stability, and reliability of automation equipment, it is often necessary to equip a lot of electrical automation control equipment as an assistant [10]. If the equipment has been used for a long time, the inspection and maintenance of key components should be done well during the overhaul process, and at the same time, attention should be paid to the standardization of operation. Once the problem of irregular operation occurs, measures should be taken to correct it immediately to avoid the failure caused by improper operation. Only in this way can the failure caused by excessive consumption of parts be avoided to the greatest extent.

The principle of multi-sensor information fusion technology is just like the process of the human brain processing all kinds of information comprehensively. Human beings instinctively synthesize the information obtained by various senses with transcendental knowledge and make estimates and judgments on the surrounding environment and the events that have occurred. When all kinds of modern information processing methods are used to realize this function by computer, multi-sensor information fusion technology is formed. Usually, the essence of information fusion is the synthesis of multiple information from low to high; the information processing process is abstracted level by level. In some cases, high-level information will produce feedback to low-level information, and high-level information will have an impact on low-level information. Low information content will also participate in the fusion of high information content, that is, mixed fusion mode, but it does not rule out the priority fusion of high information content.

Multi-sensor information fusion is a method to process information synthetically with the human brain. It can effectively use multiple or multiple types of sensory resources, and its advantage is that it can integrate various information, thus generating more information, thus achieving optimal synergy. In other words, it is not restricted by a single sensor or a few sensors and gives full play to the advantages of multiple sensors working together to improve the performance of the whole sensing system. In recent years, data fusion technology has been widely used in robots and intelligent instrument systems, battlefield observation and unmanned aerial vehicle systems, image analysis and understanding, automatic target recognition, and comprehensive processing of multi-source image data, making it a multidisciplinary comprehensive research [11].

In the field of mechanical equipment fault diagnosis, especially in data-driven methods, CNN (Convolutional Neural Network) has become an essential data processing tool. Because CNN can automatically extract features, and the extracted features have strong anti-interference ability, it is suitable for different load situations. CNN is a multi-layer neural network, which usually consists of four basic units: convolution layer, pooling layer, activation layer, and full connection layer. Its purpose is to fit a highly nonlinear function.

In this topic, we choose the CNN model to monitor the electrical automation control equipment at the same time. If the judgment results of the system are inconsistent, then there is a system failure of missing and misreporting. At this time, the decision-making layer data fusion is carried out, and its output results are used as the judgment criterion. If it is determined that there is a fault, the fault diagnosis will start. If it is determined that there is no fault, the system will not act and continue to monitor. Figure 25.2 is the workflow of the fault monitoring algorithm for electrical automation control equipment.

Based on multi-head attention mechanism and CNN, this paper proposes an end-to-end adaptive fault monitoring algorithm for electrical automation control equipment. The algorithm can adaptively select frequency data features with higher correlation with fault types to train the network model. The application of multi-head attention mechanism increases the diversity of extracted features and has stronger robustness. Combined with CNN, the accuracy of network fault classification is improved.

Multi-attention mechanism is improved from several parallel attention mechanisms, which not only improves the operation speed but also enables the network to adaptively select frequency data features with higher correlation with fault types to train the network model, thus increasing the diversity of extracted features. Figure 25.3 shows the network structure model of the multi-attention mechanism.

The structure of multi-head attention network mainly includes multi-head attention mechanism, normalization layer, full connection layer, and linear transformation layer. In the multi-head attention network, the performance of network feature extraction can be improved by stacking multiple multi-head attention layers. The multi-head attention CNN model uses two multi-head attention layers [12].

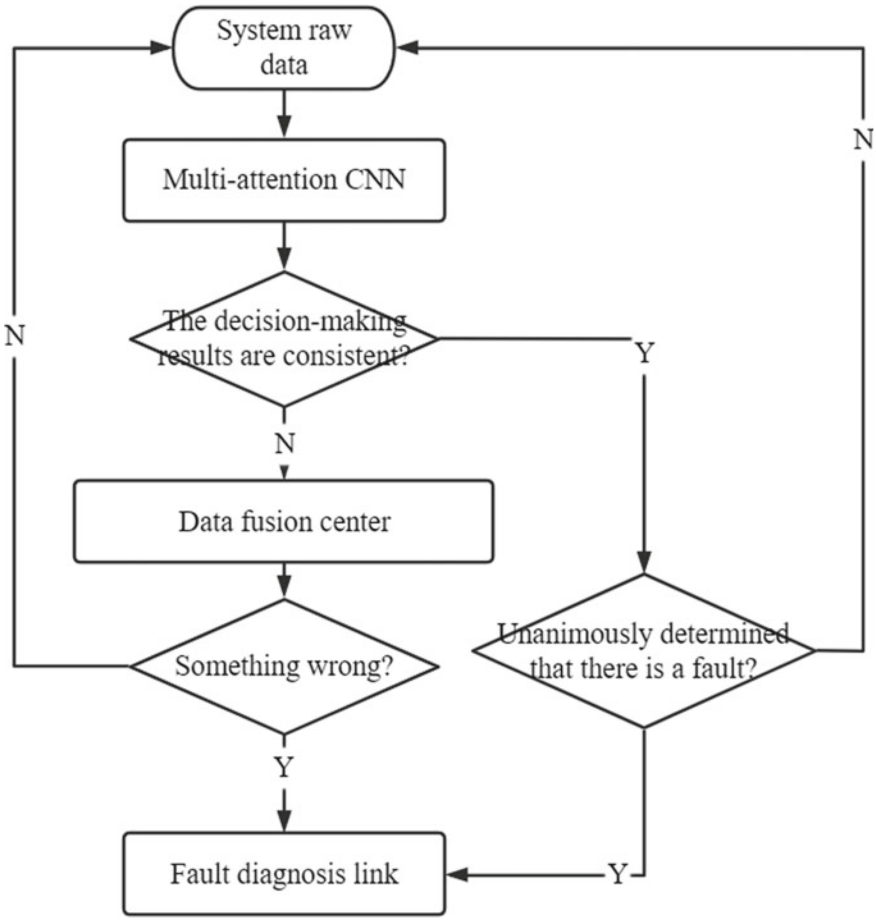


Fig. 25.2 Workflow chart of fault monitoring algorithm for electrical automation control equipment

The fully connected layer in each multi-headed attention layer structure contains two hidden layers, the middle activation function uses ReLU function, and its input and output dimensions are both d_{model}, and the hidden inner layer dimension is d_{ff}. The mathematical expression of the fully connected layer is

$$f(x) = \max(0, xW_1 + b_1)W_2 + b_2 \tag{25.1}$$

where W_1, W_2 is the weight matrix of the first and second layers respectively, and b_1, b_2 is the bias vector of the first and second layers, respectively.

In this method, the marked signal generated by the dynamic model is sent to the generator, and the generated signal obtained by the generator is sent to the recognizer for training together with the unlabeled signal actually collected, and the recognizer

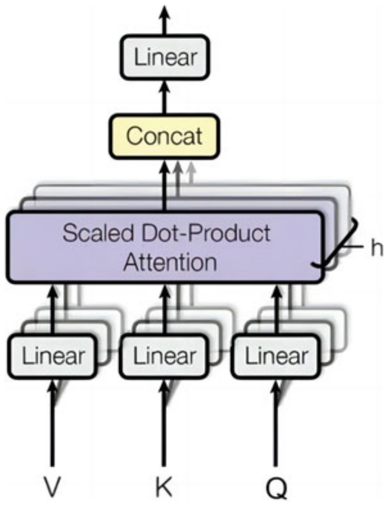


Fig. 25.3 Network structure of multi-head attention mechanism

is trained by judging whether the input signal comes from the generated signal or the truly collected signal [13, 14].

In this paper, a boundary self-balancing network is proposed, which increases the constraint between the generated signal $g(n)$ and the target domain signal n.

$$L_{gt}(n) = |G(n) - n| \tag{25.2}$$

Among them, Cov2D two-dimensional convolution layer has n convolution kernels, each kernel size is 3*3 and the step size is 2*2. After passing through each convolution layer, the batch specification layer is used to standardize the output characteristics. The activation function is chosen as Leaky ReLu, which is shown in Formula (25.3).

$$y = \max(0, x) + \alpha \min(0, x) \tag{25.3}$$

That is, the source domain signal can be converted into a generated signal with a similar signal distribution to the target domain signal.

The training process of the network model proposed in this paper is mainly to calculate the predicted value by forward propagation algorithm, then get the value of the loss function according to the error between the predicted value and the real value, and finally optimize the weight parameters in the network by backward propagation algorithm to minimize the loss function.

The optimization process is described by a mathematical formula as follows:

$$\theta^* = \arg\min_{\theta} L\left(f\left(x^i; \theta\right)\right) \tag{25.4}$$

where $L\left(f\left(x^{i};\theta\right)\right)$ is the objective function value; $f\left(x^{i};\theta\right)$ is the output value; θ is all parameters of CNN; θ^{*} is the optimal parameter of CNN; x^{i} is the input of CNN.

25.3 Experimental Analysis

In this paper, CWRU bearing data set will be used to verify the proposed end-to-end adaptive fault monitoring algorithm for electrical automation control equipment, and the experimental results will be evaluated. The test data CWRU in this paper is collected at the sampling frequency of 48 kHz, and the vibration signal of the motor is used as the test data. The same marked data is divided into a training set and a test set (7:3) according to a certain proportion. The training set data is used to train the network model, and the test set data is used to check the accuracy of the network model.

The test bench is a modular system including a driving motor (permanent magnet synchronous motor), a torque measuring instrument, and a load motor (synchronous servo motor). The whole set of measurement allows measuring a variety of mechanical damages. Such as worn component damage, is applied to the bearing. The bearing type is 6203, N203, or NU203. It contains man-made faults (EDM, drilling, and electrical corrosion), and natural faults are also collected in this data set.

In order to test the robustness of multi-headed attention CNN in a noisy environment, the fault diagnosis performance of different algorithm models in-15 dB to 15 dB noise environment is compared through experiments. The comparison algorithms are SVM and LSTM. The experimental results are shown in Fig. 25.4.

As can be seen from the figure, the prediction accuracy of all algorithms increases with the increase of SNR. When the SNR of noise signal is 7 dB, the prediction accuracy of multi-attention CNN is ninety-seven percent, and then it remains at 100%, basically without fluctuation. Experiments show that the algorithm has good noise adaptability and robustness.

According to the existing TPSN protocol, this paper writes a time synchronization protocol with nesC and connects the cross-layer routing protocol designed in this paper with the Multi-Hop cluster data transmission protocol to form a relatively complete protocol stack. The test environment is to place 35 nodes in a room of about 25 square meters of which node 0 is the sink node. Figure 25.5 shows the packet loss rate of 10 experiments.

From the data packets intercepted by the general base station, it can be seen that the node loses packets when receiving data, and after receiving timeout, it sends out the initiative to request data packets. Therefore, the fault tolerance in the protocol plays a role, as shown in Table 25.1.

After successful networking, the correct transmission rate of sensor data is 100%. The average packet loss rate of networking using the protocol is 0.51%, which shows that the protocol in this paper can really eliminate the inter-cluster interference problem of WSN based on cluster routing to a certain extent so that all clusters can coordinate data transmission in the same area.

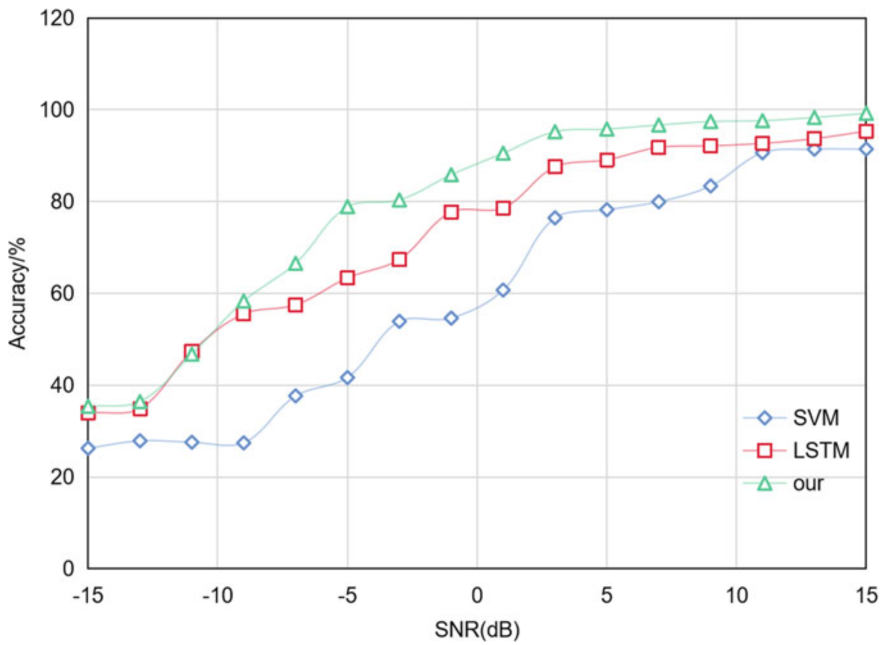


Fig. 25.4 Experimental comparison in different noise environments

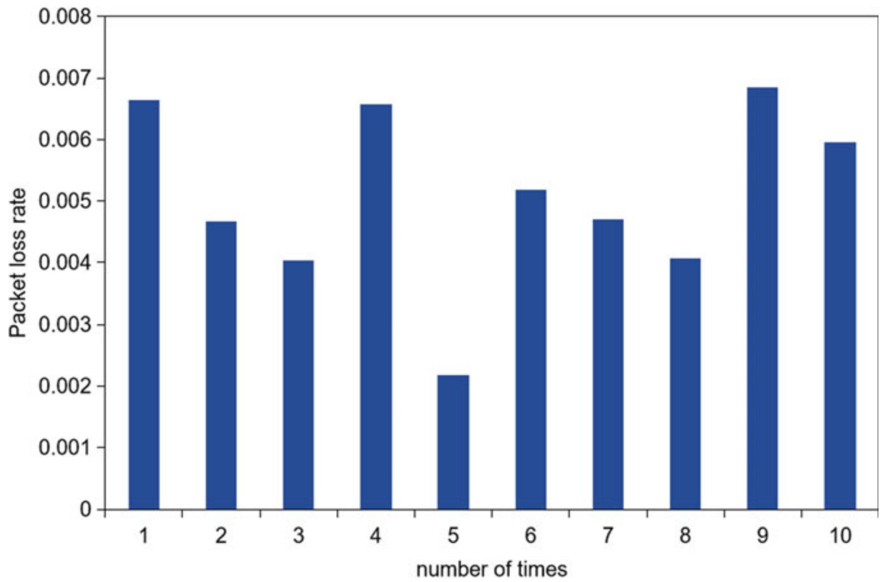


Fig. 25.5 Packet loss rate of networking

Table 25.1 Average value of test data

Average cluster head number	Average networking overhead (bit)	Average packet loss rate of networking (%)	Average sen data correct transmission rate (%)
3.853	6733.572	0.51	100

As an effective supplement to the online monitoring technology of electrical automation control equipment, the fault monitoring algorithm of electrical automation control equipment proposed in this paper can help to share the work pressure of staff and play an important auxiliary role in maintaining the normal operation of transmission and substation and improving the work efficiency of staff.

25.4 Conclusion

On the basis of multi-sensor data fusion technology, this paper introduces data fusion technology into the field of fault monitoring of electrical automation control equipment and makes full use of the redundancy between models to improve the reliability of fault prediction for complex systems. The research results show that when the signal-to-noise ratio of the noise signal is 7 dB, the accuracy of multi-attention CNN prediction has reached 97%, and has remained at 100% since then. The experimental results show that the multi-attention CNN algorithm has better noise adaptability and stronger robustness in noise interference environment compared with other mainstream algorithms. As an effective supplement to the online monitoring technology of electrical automation control equipment, the fault monitoring algorithm of electrical automation control equipment proposed in this paper can help share the work pressure of staff.

References

1. Heda, Z.: Fault diagnosis and life prediction of mechanical equipment based on artificial intelligence. J. Intell. Fuzzy Syst. **37**(12), 1–10 (2019)
2. Yuan, X.: The enhancement of fault detection for rolling bearing via optimized vmd and tqwt based sparse code shrinkage. J. Vibroengineering **24**(3), 15 (2022)
3. Shan, P., Lv, H., Yu, L., Ge, H., Li, Y., Gu, L.: A multisensor data fusion method for ball screw fault diagnosis based on convolutional neural network with selected channels. IEEE Sens. J. **20**(14), 7896–7905 (2020)
4. Yang, D., Liu, Y., Li, X., Zhou, W.: Gearbox fault diagnosis based on bacterial foraging algorithm optimization decisions. J. Central South Univ. (Science and Technology) **46**(4), 1224–1230 (2015)
5. Huo, J., Lin, D., Qi, W.: Intelligent fault diagnosis method of mechanical equipment based on fuzzy pattern recognition. J. Intell. Fuzzy Syst **38**(12), 1–8 (2020)

6. Wang, J., Li, J.: Multi-objective constrained optimizations based fuzzy fault tolerance algorithm of quarter vehicle suspensions. IEEE Access **2022**(10), 10 (2022)
7. Zhu, Y., Li, G., Tang, S., Wang, R., Su, H., Wang, C.: Acoustic signal-based fault detection of hydraulic piston pump using a particle swarm optimization enhancement cnn. Appl. Acoust. **2022**(8), 192 (2022)
8. Sahu, D., Dewangan, R.K., Matharu, S.P.S.: Fault diagnosis of rolling element bearing: a review. ECS Transactions (1), 107 (2022)
9. Li, Y.W.K.: A deep learning driven method for fault classification and degradation assessment in mechanical equipment. Comput. Ind. **104**(36), 65 (2019)
10. Zheng, B., Li, Y.F., Fu, G.Z.: Discretization algorithm based on particle swarm optimization and its application in attributes reduction for fault data. J. Shanghai Jiaotong Univ. (Science) **23**(05), 101–105 (2018)
11. Lee, K.M., Park, C.W.: Enhanced fault location algorithm for short faults of transmission line. Trans. Korean Inst. Elect. Eng. **65**(6), 955–961 (2016)
12. Chen, G.H., Qie, L.F., Zhang, A.J., Han, J.: Improved cica algorithm used for single channel compound fault diagnosis of rolling bearings. J. China Mechan. Eng. **29**(1), 8 (2016)
13. Xiong, F., Long, H.Y., Hu, X.M., Lu, X.L.: Application of momentum factor bp neural network algorithm in equipment fault prediction. Manuf. Autom. **33**(23), 4 (2011)
14. Li, Z., Liu, Y.B., Ma, Z.Y., Teng, W.: Application of adaptive resonance neural network combined with c- means clustering in fault diagnosis of wind turbine gearbox. J. Power Eng. **35**(8), 7 (2015)

Chapter 26
Traffic Classification Method Based on Graph Convolution Network

Zhaotao Wu and Zhaohua Long

Abstract More and more derivative technologies are emerging in tandem with the Internet's ongoing growth, and these derivative technologies are also continuously improving current productivity. As a result, there is an increasing amount of complex and varied data traffic that the Internet must manage. This puts a greater strain on the network and makes it more difficult to classify network traffic, and it also raises the bar for the classification speed and accuracy of the classifier. The primary focus of this article is on the flaws in conventional network traffic classification algorithms. This study suggests an enhanced graph convolutional network traffic categorization method. This model includes an autoencoder and a graph convolutional network, and it can use the latter to extract. In order to fully describe the network traffic data, the self-encoder is also used to extract node features; these two parts of features are then combined; the combined features are then fed into the second layer of graph convolution; and finally, the Softmax classifier is used. Sort the second layer of graph convolution's output data into categories.

26.1 Introduction

Begin the Introduction two lines below the Keywords. The manuscript should not have headers, footers, or page numbers. It should be in a one-column format. References are often noted in the text and cited at the end of the paper.

The idea of worldwide informatization [1, 2] has had a significant impact on how people live their lives. The entire society has been transitioning to an information society at an astounding rate over the last three decades. Today, the Internet has evolved into a crucial component of communication infrastructure, a force behind the new industrial revolution and industrial transformation, and a new engine for economic development.

Z. Wu · Z. Long (✉)
School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
e-mail: 1715633164@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_26

There is widespread use of network traffic categorization [3]. Traffic classification technology, for instance, can assist network managers in the analysis and optimization of the network, keep track of user usage patterns and traffic patterns, modify network policies in light of this information, and enhance network usage effectiveness. Network security experts can quickly identify and stop various types of network attacks, such as denial of service attacks (DoS), distributed denial of service attacks (DDoS), malware, network worms, etc., with the aid of network traffic classification technology. When an assault takes place, the classification technology can keep track of the traffic, spot anomalies, and send out an alert just in time to stop losses brought on by network attacks. In addition to monitoring and improving network service quality, network bandwidth and resource allocation, reaction time, and network stability can all be improved with the help of network traffic classification technology.

Therefore, it is essential to constantly enhance the precision and effectiveness of network traffic classification technology in order to better manage the difficulties and pressures brought on by network traffic. In order to investigate and analyze the current network traffic classification mechanism as well as to research and address the flaws of the current network traffic classification methods, this subject will integrate a wide range of domestic and international literature with practical applications.

26.2 Traditional Network Traffic Classification Methods

26.2.1 Port-Based Network Traffic Classification Method

Port identification, a technology to assign network traffic to specific ports, was the first scientific method for classifying different kinds of network traffic. This method can classify data packets as particular apps or services by looking at the destination port number. This method is easy to implement, inexpensive, and has excellent real-time performance. Although many applications and services may share the same port number due to the prevalence of peer-to-peer (P2P) applications with dynamic ports, port-based classification methods are unable to correctly direct traffic to the intended application or service. Additionally, the precision of the classification method has been significantly decreased by technologies like port masquerading, rendering the port-based traffic classification method ineffective.

26.2.2 Deep Packet Inspection-Based Network Traffic Classification Method

Aiming at the problem of unfixed allocation of specific ports to services and protocols brought about by dynamic ports, network traffic identification and classification technologies based on Deep Packet Inspection (DPI) [4, 5] began to appear, and were also adopted Traffic categorization based on load is called that. This technique uses in-depth network data packet detection and analysis to precisely classify and recognize network traffic. However, the DPI-based traffic classification technique has a relatively high application cost and higher hardware and memory requirements. Additionally, because the DPI-based traffic classification technique must thoroughly detect and analyze every bytecode of the data packet, it might raise privacy concerns for users.

26.2.3 Machine Learning-Based Network Traffic Classification

Numerous researchers have used machine learning algorithms to categorize network data in light of the machine learning field's ongoing maturation [6, 7]. Using unbalanced datasets and the SVM learning algorithm, Dong et al. [8] suggested an improved version of CMSVM to address the issue of data imbalance. To address various application unbalance issues, CMSVM employs a multi-class SVM with an active learning method. Last but not least, the algorithm was tried on various datasets, including Moore, and produced accurate classification outcomes. Semi-supervised learning is the foundation of the classification technique Erman et al. [9] proposed, which is adaptable to both known and unidentified applications. The model is appropriate for scenarios with high real-time requirements because it does not require a large portion of the label flow in the dataset during training, and it performs reasonably well in real time. Deep learning-based semi-supervised learning classification methodology was suggested by Aouedi et al. [10]. This approach enhances the model's capacity to handle labeled and unlabeled data by combining semi-supervised learning and stacked sparse autoencoder (SSAE), and it adds noise reduction and discarding methods to the model to increase the robustness of the extracted features. Avoid overfitting issues while exercising. The modified classification model outperforms the traditional model in terms of classification performance while keeping the complete classification process automated.

The process of feature selection, however, in the conventional field of machine learning, heavily depends on the professional expertise of those who define the features. Traditional machine learning methods place higher demands on the professional level and industry experience of feature engineers because choosing a good feature is frequently a challenging job. As a result, there is significant uncertainty in the performance of the final model.

26.3 Related Work

26.3.1 Design and Analysis of Model

In order to fully describe network traffic data, the model that combines GCN (Graph Convolutional Network) and Autoencoder (AE) [11] can use GCN to derive structural information between nodes and use AE to extract node features. Figure 26.1 depicts the model's design. To be more precise, the model's procedure can be summed up as follows: represent the network traffic data as a graph, with the nodes denoting the various kinds of data flow in the network and the edges denoting the relationships between the nodes; Each node's location in the graph is indicated by a coded vector created for it using GCN's Structural information of neighboring nodes and AE's feature vector compression. Each node's position in the feature space is indicated by a coded vector created for it using AE's feature vector compression. In order to accomplish the goal of traffic classification, the GCN encoding vector and the AE encoding vector are fused to create the node's final encoding vector. After that, the nodes are classified and mapped to various categories using the classifier.

Specifically, the steps of the model are as follows:

(a) Data preprocessing: The initial encrypted traffic data in PCAP format for the public datasets USTC-TFC2016 [12] and ISCX 2016 VPN-nonVPN [13] will be transformed into a classification model with a unified format at this point. Enter numbers. Remove the data from the original data that tampers with the classification findings. Tables 26.1 and 26.2 display the two data groups' original data in detail.

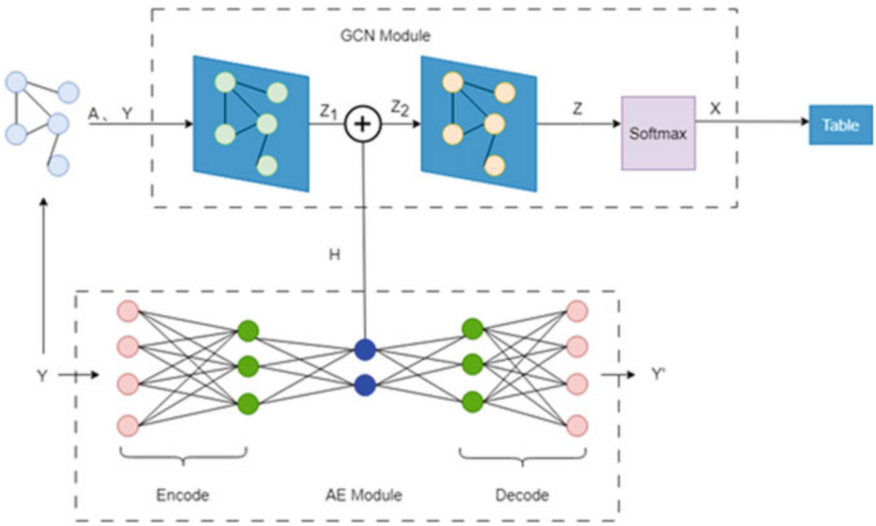


Fig. 26.1 Improved model architecture of graph convolutional networks

Table 26.1 Traffic types of the USTC-TF2016 dataset

App type	Size (MB)	Class	Malware type	Size (MB)
MySQL	22.3	Database	Geodo	28.8
Outlook	11.1	Email/Webmail	Cridex	94.7
Bittorent	7.33	P2P	Shifu	57.9
Gmail	9.05	Email/Webmail	Neris	90.1
Skype	4.22	Chat/IM	Zeus	13.4
Facetime	2.4	Voice/Video	Tinba	2.55
SMB	1206	Data Transfer	Virut	109
FTP	60.2	Data Transfer	Miuref	16.3
WOW	14.9	Game	Nsisay	281
Weibo	1618	Social Network	Htbot	83.6

Table 26.2 Traffic types of the ISCX 2016 VPN-nonVPN dataset

App type	Size (MB)	Class	App type	Size (MB)
ICQ	7	Chat/VoIP	Voipbuster	842
Email	28	Email	Gmail	12
SFTP	418	File/VoIP	Hangouts	3766
Youtube	251	Streaming/File	Skype	2872
Netflix	299	Streaming/File	FTPS	418
Spotify	40	Streaming/P2P	Torrent	70
Facebook	2502	VoIP	SCP	448
AIM	5	Chat	Tor	202
Vimeo	146	File		

We typically gather network traffic data, where each data point reflects traffic in a network, for tasks like network traffic classification. Each data point typically has a few important properties, such as five-tuple data, transfer rate, and so forth. The traffic in the PCAP file is then converted into a flow form based on the five-tuple data, and various application kinds are identified. The results are then normalized, and the processed results can be expressed as Y.

$$\mathrm{Y} = [\mathrm{y}_1, \mathrm{y}_2, \mathrm{y}_3, \ldots, \mathrm{y}_\mathrm{n}] \tag{26.1}$$

(b) Create a KNN graph: We can compute the similarity between two data streams, treating comparable streams as adjacent nodes, in order to use these data for a graph-like representation. The formula for determining the similarity between two m-dimensional flow vectors is as follows R_{ij}。 In the equation, m represents the dimension of the vector.

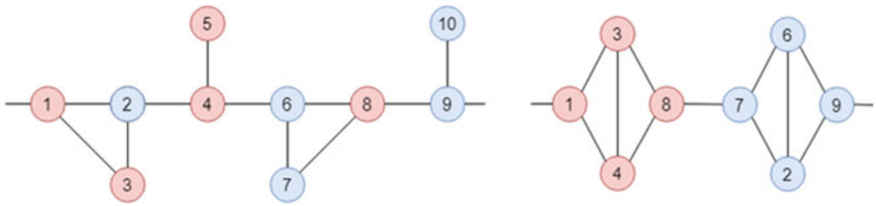


Fig. 26.2 Flow trajectory diagram and KNN (K = 3) diagram example

$$\mathrm{R_{ij}} = \frac{m\sum_{i=1,j=1}^{m} y_i y_j - \sum_{i=1}^{m} y_i \sum_{j=1}^{m} y_j}{\sqrt{(m\sum_{i=1}^{m} y_i^2 - (\sum_{i=1}^{m} y_i)^2)(m\sum_{j=1}^{m} y_j^2 - (\sum_{j=1}^{m} y_j)^2)}} \tag{26.2}$$

The K-adjacency graph of the stream is then obtained, as shown in Fig. 26.2, by selecting the top K most relevant streams of each stream as the current stream's adjacent nodes after getting the similarity between streams. Although some of the flows in Fig. 26.2 (left) pertain to the same application type, no edges have been established between them. This is so because the trajectory graph only has a few similar flows and is only connected by the corresponding IP addresses. information. The K-adjacency graph is more beneficial for subsequent traffic classification because, as shown in Fig. 26.2 (right), the KNN graph created by similarity calculation can create more links between the same application type of traffic. Additionally, the KNN graph makes it simpler to distinguish between different types of flows than the trajectory graph.

Finally, we merge all of the nodes and edges into a graph form, which is a matrix of adjacencies. The input for our network traffic classification algorithm is this graph.

(c) Autoencoder encoding: An autoencoder's training procedure primarily consists of two stages: encoding and decoding. We pass the input data into the encoder during the encoding step to create a smaller encoding vector that can represent the node's location in the feature space and be applied to subsequent classification tasks. In order to get reconstructed input data that is as close to the original input data as possible, we send the encoded vector to the decoder during the decoding step. The autoencoder can learn a compressed representation and a decompressed reconstruction of the input data by repeating this procedure.

Since the multi-layer autoencoder network structure is used in this study, the distinctive feature H_{l+1} discovered by the $l+1$ layer can be expressed as follows:

$$\mathrm{H}_{l+1} = \sigma(W_l H_l + b_l) \tag{26.3}$$

In Formula 26.3, $\mathrm{W_l}$ and $\mathrm{b_l}$ represent the weight and intercept of Layer l, respectively. The loss function is defined as

$$L_{ae} = \frac{1}{2m}\sum_{i=1}^{m} ||y_i - y_i'||^2 \tag{26.4}$$

The meaning of m in the formula is the same as that in Formula 26.2. In the network traffic classification job, the node feature vector can be compressed into a smaller encoded vector that represents the node's location in the feature space using the autoencoder's encoding component. This encoded vector can be used in future classification tasks as a new feature representation of the node.

(d) GCN encoding: In order to perform later classification and prediction tasks involving network traffic, it is necessary to represent each node as a vector. However, since network traffic is typically made up of connections and interactions between nodes, the relationship between nodes is frequently also a very essential piece of knowledge. As a result, we require a method for taking into account both the characteristics of components and their connections. The graph structure can be subjected to convolution processes by GCN, on the other hand, in order to extract the connections between nodes and produce an encoded vector that represents the node's location in the graph.

The input of the first layer GCN is the adjacency matrix A and the data flow Y, and the output can be expressed as Z_1:

$$Z_1 = \sigma(D^{-\frac{1}{2}} A D^{-\frac{1}{2}} Y W_0 + b_0) \tag{26.5}$$

In Eq. 26.5, D represents the degree matrix. W_0 represents the weight matrix, b_0 represents the intercept, and σ represents nonlinear operations. We can create an encoding matrix with each row corresponding to a node's position in the network by using GCN operations. This encoding matrix can be used for subsequent classification tasks because it includes structural data about the relationships between the nodes.

(e) Fusion encoding vectors: In network traffic classification tasks, we need to comprehensively consider the positions of nodes in the graph structure and feature space to acquire the final encoding vectors of nodes. We can combine autoencoder and GCN-encoded vectors to accomplish this.

The fused coded vector serves as the second layer of GCN's input in this paradigm. Fusion can occur in a variety of methods. The typical approach is to weigh the self-encoder-coded vector and the GCN-coded vector to produce a weighted average of the vectors. written as Z_2:

$$Z_2 = \eta Z_1 + (1 - \eta)H \tag{26.6}$$

The weight coefficient, η , is used to regulate the input ratio of the two vectors in Formula 26.6. H is the encoding vector generated by autoencoder.

The final encoded vectors of the nodes, which indicate their positions in the graph structure and feature space, can be obtained by fusing the GCN-encoded vectors and the autoencoder-encoded vectors. After feeding the fusion vector into the second layer GCN, the matching output after using this encoding vector as a new feature of the node can be written as follows:

$$Z = \sigma(\hat{D}^{-\frac{1}{2}}\hat{A}\hat{D}^{-\frac{1}{2}}Z_2W_1 + b_1) \tag{26.7}$$

In Formula 26.7, $\widehat{A}$ represents the result of adjacency matrix A plus self ring, and $\widehat{D}$ is the degree matrix corresponding to $\widehat{A}$. W_1 and b_1 represent the weight matrix and intercept of the second layer GCN, respectively. For a multi-classification task with classes C, it minimizes the loss function and is defined as the cross-entropy loss function on the collection of labeled samples:

$$L_{gcn} = \sum_{m \in n_m} \sum_{c=1}^{C} N_{mc} \ln Z_{mc} \tag{26.8}$$

The collection of node indices with label N is represented by n_m in Eq. 26.8. This model's total loss function is

$$L = L_{gcn} + L_{ae} \tag{26.9}$$

(f) Classification: In the job of classifying network traffic, nodes can be categorized using the final encoded vectors. This procedure can be broken down into two steps: first, the encoding vector must be fed into a classifier, and then the classifier must be used to map nodes to various groups. To accomplish the goal of classification in network traffic, we can input the final encoding vector into a Softmax classifier and map the encoding vectors of nodes to various categories. As specifically stated

$$X = \mathrm{Softmax}(D^{-\frac{1}{2}}AD^{-\frac{1}{2}}Z_2W_1 + b_1) \tag{26.10}$$

The classification performance of the model can be improved by using a two-layer GCN model, which can improve the feature representation of nodes and prevent GCN from overemphasizing the association of nearby nodes while ignoring the characteristics of the nodes themselves.

26.3.2 Experimental Environment

The experiment's physical host configuration specifications include an Intel i5-12400F processor, an NVIDIA GeForce GTX 1080Ti graphics card, and 64G of RAM. The deep learning framework, TensorFlow GPU 1.8 and Keras 2.2.0, Python

3.6, and the programming environment, PyCharm 2021, are all running on the 64-bit Windows 10 operating system.

26.4 Experimental Results

The classification performance of the AT-GCN model is tested in this paper using the USTC-TFC2016 and ISCX 2016 VPN-nonVPN datasets, with the experimental findings shown in Figs. 26.3 and 26.4.

Figure 26.3 displays the index values for the AT-GCN model's classification results in the USTC-TFC2016 dataset. The AT-GCN model achieved a classification accuracy of 99.7% for Skype and Htbot types of data, with an average classification accuracy of 99.2%. In terms of accuracy, the classification accuracy of the AT-GCN model for Miuref-type data reached 99.7%, and the average classification accuracy was 98.6%. In terms of recall rate, the classification recall rate of the AT-GCN model for Virut and Zeus type data reached 99.7%, and the average classification recall rate was 98.7%. In terms of F1-score, the F1-score of the AT-GCN model for Geodo-type data classification reached 99.8% and the average F1-score was 98.7%.

Figure 26.4 displays the index values for the AT-GCN model's classification results in the ISCX 2016 VPN-nonVPN dataset. For data classification of the Netflix, Facebook, AIM, and Tor types, the AT-GCN model gets an accuracy rate of 99.7%, while the lowest accuracy rate for data classification of the ICQ type is 98.2%, with an average classification accuracy rate of 99.3%. The AT-GCN model has a precision of 98.6% on average, with a precision in classification of 99.8% for data of the type Email and a classification precision of 96.3% for data of the type Torrent. The

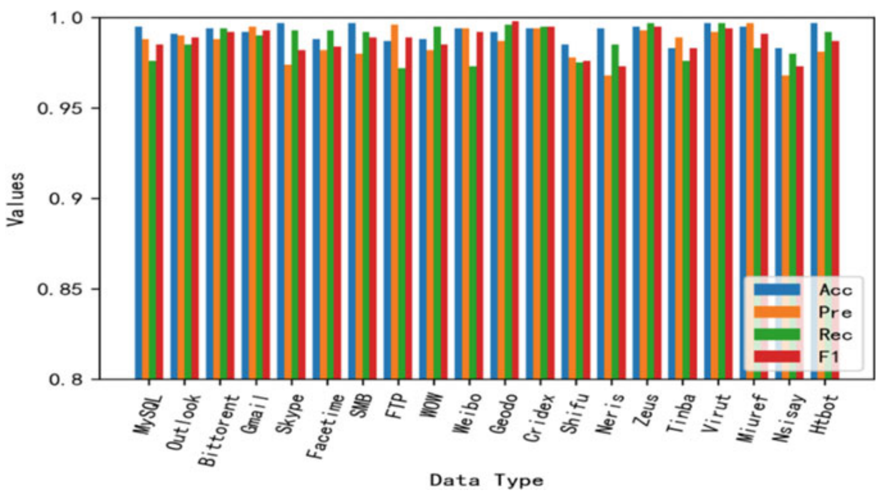


Fig. 26.3 Classification results of the USTC-TFC2016 dataset using the AE-GCN model

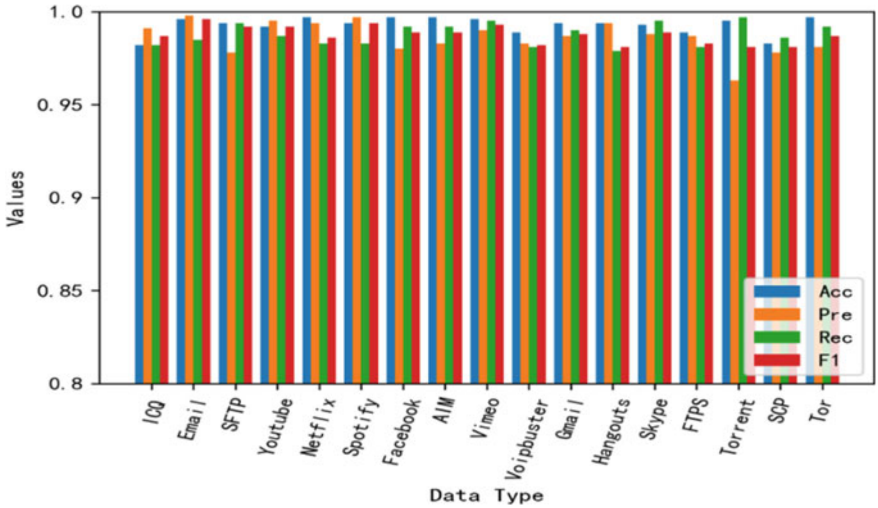


Fig. 26.4 Classification results of VPN-nonVPN datasets by AE-GCN model

AT-GCN model's classification recall rate for data of the torrent type reached 99.7%, while the recall rate for data of the Hangouts type was the lowest, achieving 97.9%, with an average recall rate of 98.8%. The AT-GCN model had an F1-score of 99.6% for the categorization of email-type data, and the lowest F1-score (98.1%) for the classification of Hangouts, Torrent, and SCP-type data, with an average F1-score of 98.8%.

26.5 Conclusion

The use of graph convolutional networks in traffic categorization is first discussed in this chapter. The fundamental ideas and features of graph convolutional networks, such as neighbor aggregation and feature propagation, are then explored. Then, a more effective traffic classification algorithm using a graph convolutional network is suggested, and a thorough explanation is provided from the model's architecture and mathematical analysis. The model can directly use the network traffic topology to completely mine the characteristics and relationships between traffic, increasing classification efficiency and accuracy.

Two datasets, USTC-TFC2016 and ISCX 2016 VPN-nonVPN, are used to evaluate the model in the experimental section. The experimental findings demonstrate that the model's categorization performance is largely reliable, with an average classification accuracy of over 96%. Traditional traffic classification techniques and feature-based classification techniques are outperformed by the convolutional network's traffic classification algorithm in terms of precision and efficiency. The algorithm can also produce good classification results when faced with a variety of network traffic types and sizes because it has excellent robustness and generalization capabilities.

References

1. Nye Jr, J. S.: Power in the global information age: from realism to globalization. London: Routledge, pp. 1–39 (2004)
2. Martin, W.J.: The global information society. London: Routledge, pp. 184–201 (2017)
3. Dainotti, A., Pescape, A., Claffy, K.C.: Issues and future directions in traffic classification. IEEE Network **26**(1), 35–40 (2012)
4. Dharmapurikar, S., Krishnamurthy, P., Sproull, T., Lockwood, J.: Deep packet inspection using parallel bloom filters. 11th Symposium on High Performance Interconnects, 2003. Proceedings. IEEE, Stanford, CA, USA, 44–51 (2003)
5. El-Maghraby, R.T., Abd Elazim, N.M., Bahaa-Eldin, A.M.: A survey on deep packet inspection. 2017 12th International Conference on Computer Engineering and Systems (ICCES). Cairo, Egypt, pp. 188–197 (2017)
6. Jordan, M.I., Mitchell, T.M.: Machine learning: trends, perspectives, and prospects. Science **349**(6245), 255–260 (2015)
7. Pacheco, F., Exposito, E., Gineste, M., Baudoin, C., Aguilar, J.: Towards the deployment of machine learning solutions in network traffic classification: a systematic survey. IEEE Commun. Surv. Tutor. **21**(2), 1988–2014 (2018)
8. Dong, S.: Multi class SVM algorithm with active learning for network traffic classification. Expert Syst. Appl. **176**, 114885 (2021)
9. Erman, J., Mahanti, A., Arlitt, M., Cohen, I., Williamson, C.: Offline/realtime traffic classification using semi-supervised learning. Perform. Eval. **64**(9–12), 1194–1213 (2007)
10. Aouedi, O., Piamrat, K., Bagadthey, D.: A semi-supervised stacked autoencoder approach for network traffic classification. In 2020 IEEE 28th International Conference on Network Protocols (ICNP) (pp. 1–6). IEEE (2020)
11. Wang, W., Huang, Y., Wang, Y., Wang, L.: Generalized autoencoder: a neural network framework for dimensionality reduction. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 490–497). IEEE (2014)
12. Wang, W., Zhu, M., Zeng, X., Ye, X., Sheng, Y.: Malware traffic classification using convolutional neural network for representation learning. 2017 International conference on information networking (ICOIN) (pp. 712–717). IEEE (2017)
13. Draper-Gil, G., Lashkari, A. H., Mamun, M. S. I., Ghorbani, A. A.: Characterization of encrypted and VPN traffic using time-related. Proceedings of the 2nd international conference on information systems security and privacy (ICISSP). Rome, Italy, (pp. 407–414) (2016)

Chapter 27
RVDNet: Rotated Vehicle Detection Network with Mixed Spatial Pyramid Pooling for Accurate Localization

Jianhong Zhou, Zhangzhao Liang, Zijun Tan, Wenba Li, Qing Li, Zilu Ying, Yikui Zhai, Yuanfeng He, and Zhanyu Shen

Abstract The objective of UAV-based vehicle detection is to accurately locate and classify vehicles in remote sensing images. It contributes to the improvement of intelligent city traffic management and enhances public safety monitoring. However, UAV offers more finely detailed remote sensing images than orbital satellites, resulting in more complex backgrounds, which poses a significant challenge. To tackle this challenge, we created a vehicle detection dataset named UAV Vehicle, which was collected by UAV flying at altitudes ranging from 250 to 400 m. The dataset comprises more than 200 remote sensing images, each with a size of 8192×8192, encompassing diverse complex scenes including urban, rural, and construction sites. However, the excessive complexity of the background conditions results in an imbalance between the number of foreground and background instances, consequently impacting the accuracy of final detection and localization. To address this issue, we proposed the RVDNet, which extracts precise positioning information from multiple scales, resulting in a substantial improvement in the model's localization performance. We developed an MSPP module, which employs random pooling methods to extract features at various scales, serving as a regularization technique to mitigate overfitting. Additionally, we introduced an attention mechanism to enhance the model's ability to represent and infer object positions more accurately. Extensive experiments conducted on the UAV Vehicle dataset demonstrate the efficacy of the proposed MSPP module and the introduced attention mechanism.

J. Zhou · Z. Liang · Z. Tan · W. Li · Q. Li · Z. Ying · Y. Zhai (✉) · Y. He · Z. Shen
Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen, China
e-mail: yikuizhai@163.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_27

27.1 Introduction

Vehicle detection involves locating and categorizing vehicles in images or videos using diverse detection techniques and algorithms. This task finds wide applications in fields like smart city traffic management [1], intelligent assisted driving, and security monitoring. Presently, the majority of remote sensing images are collected using orbital satellites. Although orbital satellites can capture high-resolution remote sensing images, the high information density poses challenges in identifying specific vehicle categories. Additionally, remote sensing images acquired from orbital satellites have long acquisition cycles and cannot be detected in real time. Moreover, the acquisition of these satellite-based remote sensing images entails professional processing and authorization procedures. Lastly, the usage conditions frequently deviate from the actual situation, resulting in issues like reduced target resolution, significant atmospheric interference, and complex processing.

With the rapid development of Unmanned Aerial Vehicle (UAV), they have found wide-ranging applications in various scenarios, including landscape photography, security monitoring, and smart city management, due to their flexibility. In comparison to orbital satellites, remote sensing images obtained through UAV offer the following advantages: (1) higher spatial resolution, enabling more detailed information in UAV remote sensing images; (2) real-time and flexible image acquisition, as UAV can be instantly deployed and controlled at any time and location for image capture tasks; (3) cost-effectiveness and less stringent acquisition regulations. Consequently, in recent years, numerous researchers in related fields have substituted orbital satellites with UAV for remote sensing image collection, facilitating noteworthy advancements in vehicle detection. Unfortunately, the majority of images captured by UAV are aerial images, which exhibit lower geometric accuracy than remote sensing images. Furthermore, the acquisition height is typically below 150 m, resulting in a lack of images at higher altitudes ranging from 250 to 400 m, thereby creating a gap. From a higher vantage point, the surveyed area expands, causing vehicle targets to appear smaller and more concentrated, thereby introducing greater challenges in target detection tasks. The three images are from the UAV Vehicle dataset (see Fig. 27.1). The detection results within the images were obtained using RVDNet and presented by merging the Car and Van categories into the "S_car" category while merging the Truck and Bus categories into the "L_car" category.

In response to the aforementioned issues, we acquired more than 200 remote sensing images using UAV. After applying appropriate processing, we annotated the images according to four categories: Car, Van, Truck, and Bus. This was done to address the lack of vehicle detection remote sensing images at heights between 250 and 400 m. To achieve more accurate identification of densely small vehicle targets, we proposed a Rotated Vehicle Detection Network (RVDNet). This network primarily comprises the mixed spatial pyramid pooling (MSPP) module and the convolutional block attention module (CBAM), aiming to improve the localization of vehicle targets and achieve better performance.

Fig. 27.1 Detection results of images from the UAV Vehicle dataset using the RVDNet

27.2 Related Work

In recent years, several bird's-eye view datasets have been proposed for unmanned aerial vehicle surveys. The DLR 3 K [2] dataset comprises 20 remote sensing images captured by the DLR 3 K camera system. It primarily focuses on two categories: car and truck. Considering that the images are acquired at an altitude of 1,000 m, the resolution of each image is set to 5616 × 3744 pixels. The VEDAI [3] dataset is specifically designed for evaluating small vehicle target detection in remote sensing images. It encompasses over 1,200 images with more than 3,700 instances of distinct vehicle categories. Additionally, it provides images in two modalities: color and infrared. On the other hand, the UAVDT [5] dataset includes raw videos with over 80,000 available frames collected over a span of 10 h. This dataset provides 14 available attributes, including flight height, vehicle category, and occlusion, along with annotated bounding boxes. The VisDrone [6] dataset comprises 10,209 images cropped from 263 videos. It offers comprehensive annotations such as object categories, occlusion, truncation, and bounding boxes. DOTA [4], another dataset, encompasses 15 different object categories with a total of 188,282 annotated instances. Lastly, the DroneVehicle [7] dataset contains 28,439 pairs of multimodal images (color and infrared) and covers five common vehicle types: Car, Truck, Bus, Van, and Freight. The data in this dataset is collected at heights ranging from 50 to 150 m.

Vehicle detection aims to localize vehicles and classify them according to their categories. Recent years have witnessed remarkable advancements in object detection, and bird's-eye view vehicle detection has emerged as a crucial subfield of rotated object detection in remote sensing, capitalizing on the progress in object detection. Nonetheless, commonly used evaluation metrics like Pascal VOC [8] metrics may not be appropriate for rotated object detection, particularly in scenarios with densely small objects. This is because even slight variations in the detection box angle can significantly impact the intersection over union (IOU) calculation. This problem becomes more pronounced when dealing with high-altitude remote sensing images captured by unmanned aerial vehicles. To tackle this challenge, we presented RVDNet, a method that integrates multi-scale mixed information and employs attention mechanisms to enhance the accuracy of detection localization outcomes.

27.3 UAV Vehicle Dataset

This section presents a comprehensive overview of data collection and preprocessing as well as information related to the UAV Vehicle dataset in the context of vehicle detection experiments.

27.3.1 Data Collection

The UAV Vehicle dataset comprises 15,313 RGB images, captured exclusively using the DJ Mavic2 UAV. To capture a wider range of remote sensing images, we obtained permission from the government for image capture at altitudes ranging from 250 to 400 m in various towns and urban areas within Shantou City and Jiangmen City, Guangdong Province, China. The UAV Vehicle dataset includes specific instance categories such as Car, Van, Bus, and Truck. It is important to note that the UAV Vehicle dataset was collected using the DJ Mavic2 UAV under various scenes and flight altitudes. The collection process included determining the flight altitude and GPS position of the UAV, capturing remote sensing images from multiple directions at designated locations, and synthesizing these images into panoramic images through stitching algorithms. The application of this method resulted in over 200 panoramic images. Moreover, the UAV Vehicle dataset was manually annotated with 152,280 bounding boxes.

27.3.2 Data Preprocessing

Due to the images being captured at flight altitudes ranging from 250 to 400 m, the data collection process was easily influenced by factors such as real scene lighting, camera angles, and wind speed. Accordingly, it was necessary to process the collected

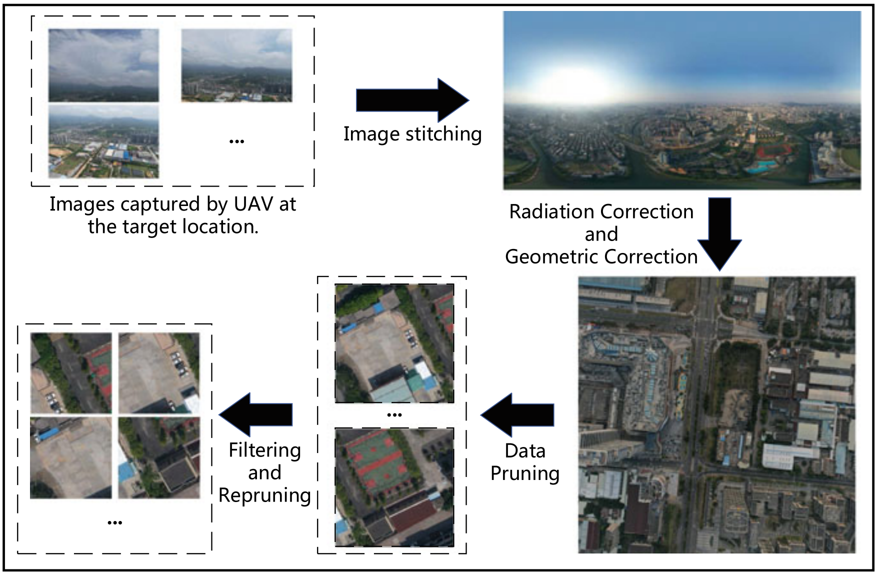


Fig. 27.2 The distribution of class samples varies between the training set and validation set of the UAV vehicle dataset

data in order to mitigate the interference of external factors on the images. The specific process of data preprocessing is shown in Fig. 27.2.

Radiation Correction.

To improve the quality of the images and eliminate variations in image pixel grayscale resulting from lighting and atmospheric conditions, it was necessary to undergo radiometric calibration. Common radiometric calibration methods include empirical calibration, physical model construction, horizontal plane projection, spectral matching, and radiometric calibration based on pseudo-invariant characteristics. These methods are commonly used for radiometric calibration. During the process of UAV image stitching, radiometric calibration differences arise as a result of varying shooting angles. Nevertheless, pseudo-invariant characteristics exhibit consistent spectral characteristics across different times and conditions. Thus, employing methods based on pseudo-invariant characteristics allows for the calibration of the images' radiometric values, enhancing their consistency in terms of brightness and color and facilitating seamless stitching. Moreover, radiometric calibration based on pseudo-invariant characteristics offers high practicality as it eliminates the need for additional radiation data collection devices. Consequently, we adopt the radiometric calibration method based on pseudo-invariant characteristics. Initially, representative buildings and roads were selected for feature matching. Subsequently, radiometric calibration coefficients were calculated using the matching results. Finally, the results were utilized for performing radiometric calibration on the UAV images. Prior to further analysis, this method was employed to conduct radiometric calibration on the stitched panoramic images.

Geometric Correction.

Due to various systematic or non-systematic factors such as camera azimuth parameters, UAV flight altitude, and terrain variations, the relative positions of targets often caused distortions and stretching of pixels relative to their true positions. Therefore, during the data collection process, the UAV recorded camera pose information when capturing images. Then, we matched the images with the UAV's pose information and performed necessary operations to achieve geometric correction, ultimately generating digital orthophoto map with a resolution of 8192 $\times$ 8192 pixels.

Data Pruning.

Directly using radiometrically calibrated and pseudo-orthorectified remote sensing images for annotation or model training is not feasible. Therefore, additional processing of the orthorectified images is required. Pruning is a key step in dataset creation. First, we cropped the orthorectified images with a resolution of 8192 $\times$ 8192 pixels into a 1024 $\times$ 1024 image patch. Then, we manually inspected each image block, discarded poor quality images, such as blurry images, and re-cropped images that exhibited severe boundary deformations into 640 $\times$ 640 image patch. After further selection, we obtained data without annotations.

27.3.3 Data Annotation

In the field of object detection, the prevalent approach for data annotation involves annotating image objects using rectangular bounding boxes. Three common methods are used for bounding box annotation: (x_c, y_c, w, h), (x_{left}, y_{left}, w, h), and (x_{left}, y_{left}, x_{right}, y_{right}). Here, (x_c, y_c) denotes the center position, (x_{left}, y_{left}) represents the top-left corner position, (x_{right}, y_{right}) corresponds to the bottom-right corner position, while w and h denote the width and height, respectively. These annotation methods are applicable in typical object detection scenarios such as autonomous driving and hinge monitoring. However, due to the uncertainty of target orientation in pseudo-orthorectified images, accurately calibrating the targets using conventional rectangular boxes becomes challenging. To address this problem, during the annotation process, we employed arbitrary quadrilaterals to precisely and concisely represent the object contours. Subsequently, we derived the final annotation information by computing the minimum bounding rectangle. The annotation information of rotated vehicle bounding boxes was saved by referring to the DOTA dataset.

27.3.4 Statistics and Attributes

We selected and annotated four categories for the UAV Vehicle dataset: cars, trucks, buses, and vans. The dataset includes 15,313 RGB images collected by UAV. Rich annotations were provided using arbitrary quadrilaterals for these categories. The images were captured during daylight, encompassing diverse scenes such as parking

Table 27.1 Comparison of state-of-the-art benchmark datasets incorporating vehicle categories is conducted

Vehicle detection datasets	Scenario	#Images	Categories	Avg. $\#\frac{labels}{categories \times images}$	Resolution	Oriented B B	Year
DLR 3 K [2]	aerial	20	2	147.3	5616 × 3744	√	2015
VEDAI [3]	aerial	600	9	0.685	1024 × 1024	√	2015
DOTA [4]	aerial	2806	15	4.45	800 ~ 4000	√	2018
UAVDT [5]	UAV	80,000	3	3.51	1080 × 540	×	2018
VisDrone [6]	UAV	10,209	10	5.31	2000 × 1500	×	2020
DroneVehicle [7]	UAV	28,439	5	6.70	840 × 712	√	2021
UAV Vehicle(ours)	UAV	15,313	4	2.02	640 ~ 1024	√	2023

lots, residential areas, urban roads and blocks, construction sites, and rural areas. Among the annotations, there were 133,490 instances of cars, 11,822 instances of trucks, 3,895 instances of buses, and 3,073 instances of vans, totaling 152,280 instances of vehicles. On average, each image contained 9.95 vehicles, with a maximum count of 126 vehicles and a minimum count of 0 vehicles. For further reference, we provided a comparison with other datasets and more detailed information about the.

UAV Vehicle dataset (see Table 27.1). It is important to note that the term “resolution” denotes the uniform resolution or range of minimum and maximum resolutions included in benchmark or the videos/images contained within the dataset. The abbreviation “BB” stands for bounding box, while “#Images” indicates the quantity of color images.

In conclusion, this dataset covers complex scenarios and higher flight altitudes, featuring a substantial number of meticulously annotated objects, thus addressing the gap in vehicle detection for remote sensing images above 150 m.

27.4 Rotated Vehicle Detection Network

To achieve precise localization of vehicle targets, we proposed RVDNet, which incorporates an MSPP module combining the Spatial Pyramid Pooling (SPP) [12] module and mixed pooling [17]. Moreover, to enhance the model’s performance during the inference stage, we incorporated an attention mechanism to provide the model with supplementary information. The overall architecture of our proposed

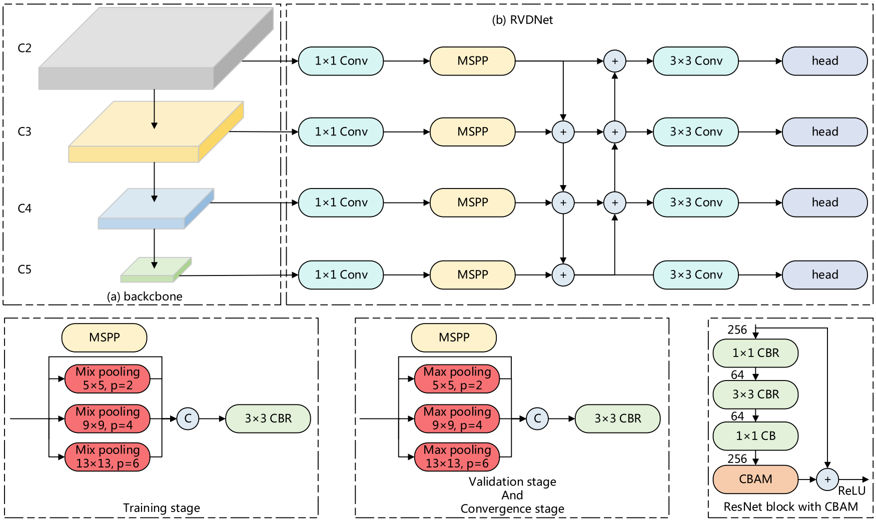


Fig. 27.3 Complete framework diagram of the RVDNet model

approach is illustrated (see Fig. 27.3). The model of RVDNet integrates the CBAM attention mechanism into the ResNet50 back-bone blocks to enhance the model's capability of extracting representations. The feature maps C2–C5 from the backbone network undergo 1 × 1 convolutional operations for dimension reduction before being inputted into the MSPP module. During the training stage, the MSPP module employed mixed pooling with three different scales (5 × 5, 9 × 9, and 13 × 13), randomly selecting either average pooling or max pooling to extract features at various scales. These features were then fused with the source inputs through concatenation operations, followed by a 3 × 3 convolutional operation to reduce feature dimensions. In the model convergence and inference stages, max pooling replaced the mixed pooling. Finally, the PANet was utilized to fuse position and semantic information from feature maps of varying sizes. The resulting feature maps were transmitted to individual detection heads for object prediction. Please note that "C" represents feature concatenation, and "add" denotes feature addition. For detailed information about "CBAM", please refer to [9].

27.4.1 Mixed Spatial Pyramid Pooling Module

Prior to introducing the MSPP Module, let's first review the SPP module and mixed pooling methods.

The SPP Module is a frequently employed pooling method in object detection and classification tasks. It is designed to perform pooling on feature maps at different scales, thus achieving scale invariance for input images. Consequently, this method effectively addresses images with multi-scale features. In specific applications, such as the YOLO series [13–16] detection models, the SPP module is introduced to capture a broader range of contextual information through pooling operations at various scales, without introducing additional computational complexity. This, in turn, improves the detection performance of the models for objects with varying scales.

Mixed pooling is a commonly used pooling method in deep learning. Its primary purpose is to enhance the model's robustness and boost its performance by obfuscating the features during training. Traditional pooling methods, such as max pooling or average pooling, may result in the loss of fine-grained information as they may prioritize prominent features while overlooking the details or emphasize general features at the expense of the prominent ones. Therefore, mixed pooling integrates and combines features by incorporating two distinct random pooling methods, thereby enhancing the representation of features and improving the model's representation and generalization capabilities.

We integrated the SPP module with mixed pooling to propose the MSPP module. The primary enhancement involves substituting max pooling with mixed pooling in the SPP module during the training stage while utilizing only max pooling during the inference stage. Given that entirely random pooling operations can impact model convergence and introduce inconsistencies between training and inference, we deviated from the approach described in the [17]. Instead, we substituted mixed pooling with max pooling during the model convergence phase, extending the application of mixed pooling to the SPP module through this training strategy, thereby enhancing the model's representational capacity.

27.4.2 Attention Mechanism

The attention mechanism is a method that enhances the representation capability of models in computer vision tasks. Common approaches to attention mechanism include the Squeeze-and-Excitation (SE) [18] module, CBAM [9], and self-attention in Transformer [19].

The SE module enhances representation capability by adaptively incorporating crucial information from channels through compression and excitation stages. During the compression stage, global average pooling is utilized to derive comprehensive channel-wise information. In the excitation stage, a multilayer perceptron (MLP) is employed to determine the significance of each channel and propagate the learned weights back to the original feature map. Moreover, CBAM integrates channel attention and spatial attention to more effectively capture essential information within the feature map. As for the Transformer, it utilizes a multi-head attention mechanism known as self-attention. This mechanism maps the input sequence into multiple projection spaces and performs self-attention calculations in each space. Consequently, it enables parallel computation of multiple self-attention heads, thereby enhancing the model's representation capability across diverse positions and semantic levels.

Given the versatility of attention mechanisms and their positive impact on the final detection performance, we incorporated the CBAM into the backbone network to augment the model's representation capability and optimize its overall performance.

27.5 Experiment

In this section, we will present our experimental setup as well as the obtained results, and provide a comprehensive analysis. Initially, we performed experiments on the UAV Vehicle dataset to assess the effectiveness of the proposed method. Subsequently, we examined the operations that influence the performance of this method.

27.5.1 Experimental Setting

27.6 Implementation Details.

ResNet50 [10] served as the backbone, with FPN [20] or PANet [11] used as the NeckNet. The model was initialized with a pre-trained ResNet50. Data augmentation techniques included random horizontal flipping and random 90–180-270 rotations, both applied with a probability of 0.5, to enhance the diversity of the dataset. The SGD + Momentum optimizer was employed for a single schedule of optimization consisting of 12 epochs. The initial learning rate was set to 0.0025, the batch size was 2, and the learning rate decay occurred at the 7th and 10th epochs, decreasing to 0.1. Weight decay was set to 0.0001, and the momentum was set to 0.9. The entire codebase was implemented using the Jittor framework, and the experiments were performed on a server equipped with 8 NVIDIA 3090 GPUs. Experiments were conducted to evaluate the replacement of PANet, MSPP module, SPP module, and the usage of CBAM attention mechanism. It should be noted that the model convergence phase refers to all epochs following the initial learning rate decay. The experimental results demonstrated that the model achieved the highest values for mAP50, mAP75, and mAP50:95 when employing PANet + MSPP + CBAM.

Partition Protocol.

The training set was split into two subsets: a training set and a validation set. The training set comprises 12,248 RGB images, whereas the validation set comprises 3,065 RGB images. Within the training set, there are 106,534 instances of cars, 2,465 instances of vans, 3,069 instances of buses, and 9,502 instances of trucks. Similarly, the validation set includes 26,956 instances of cars, 608 instances of vans, 826 instances of buses, and 2,320 instances of trucks. All experiments in this paper were performed on the training set for training purposes and evaluated using the validation set.

Evaluation Metric.

Mean Average Precision (mAP) serves as the standard metric for evaluating algorithmic detection accuracy on the UAV Vehicle dataset. mAP50 denotes the precision achieved when the predicted bounding boxes have an Intersection over Union (IOU) greater than 0.5 with the ground truth annotations. Similarly, mAP75 represents the precision obtained when the predicted bounding boxes have an IOU greater than 0.75 with the ground truth annotations. The mAP50:95 calculates the average precision across the IOU range of 0.50–0.95, with an interval of 0.05. Specifically, mAP50 represents the model's standard detection performance, mAP75 reflects the accuracy of detection localization, and mAP50:75 signifies the overall detection performance of the model.

27.6.1 Experimental Results and Analysis

To validate the benefits of PANet in integrating feature maps from different levels, we conducted relevant ablation experiments (see Table 27.2). PANet exhibited improvements compared to FPN, with a 0.89% increase in mAP50, 0.6% increase in mAP75, and 1.05% increase in mAP50:95. These results demonstrate the superior detection performance of PANet over FPN. Consequently, PANet was selected as the NeckNet for subsequent experiments.

To assess the effectiveness of MSPP, we performed corresponding ablation experiments using the SPP module. Initially, we trained separate instances of PANet, one with the embedded MSPP module and the other with the embedded SPP module. Additionally, to introduce attention mechanisms and enhance the model's representational capacity, we incorporated attention mechanisms into the blocks of ResNet50. The detection head utilized Oriented RCNN, while ResNet50 was employed as the backbone network. CBAM indicates the presence of the CBAM attention mechanism in ResNet50. NeckNet represents the employed NeckNet architecture, which includes FPN and PANet. SPP indicates the utilization of the SPP module, which can be further categorized as either employing only max pooling or mixed pooling (see Table 27.2). The experimental findings revealed that MSPP outperformed SPP in terms of localization accuracy, achieving a 1.16% increase in mAP75 and a 0.6% increase in mAP50:95. Furthermore, there was a slight improvement of 0.05% in standard detection performance. However, solely embedding the CBAM attention mechanism in the model resulted in negative optimization due to the additional attention parameters, leading to overfitting in this task. In contrast, when the CBAM attention mechanism was embedded in the model with the MSPP module, the detection performance was further improved over the original model. This enhancement translated to a 0.72% increase in mAP50, a 0.6% increase in mAP75, and a 0.69% increase in mAP50:95. These results indicate that the MSPP module also serves as a regularization technique, enhancing the model's detection performance.

Table 27.2 Ablation experiments were performed on the UAV Vehicle dataset

Backbone	CBAM	NeckNet	SPP		AP50(%)				mAP50 (%)	mAP75 (%)	mAP50:95 (%)
			Max	Mix	Car	Bus	Truck	Van			
ResNet50		FPN			90.70	90.61	84.02	55.96	80.32	67.51	55.50
ResNet50		PANet			**90.71**	**90.70**	84.08	59.36	81.21	68.11	56.55
ResNet50	√	PANet			90.69	90.43	83.19	57.98	80.57	67.43	55.75
ResNet50		PANet	√		90.70	90.56	83.43	60.17	81.21	67.31	55.85
ResNet50		PANet		√	90.70	90.64	**84.61**	59.09	81.26	68.47	56.45
ResNet50	√	PANet		√	90.70	90.50	83.93	**62.81**	**81.98**	**69.07**	**57.14**

27.7 Conclusion

This paper introduced a UAV Vehicle dataset that addressed the gap in vehicle detection using UAV remote sensing imagery at flight altitudes ranging from 250 to 400 m. Furthermore, extensive experimentation and quantitative analysis were conducted on the proposed method using the provided dataset. Through the evaluation of different modules, we obtained experimental outcomes for the model under various conditions, thereby substantiating the dependability and efficacy of the proposed approach. Consequently, this methodology offers a novel perspective for advancing high-precision and accurate detection in remote sensing vehicle detection tasks.

Acknowledgements This work was supported by Guangdong Basic and Applied Basic Research Foundation (No. 2021A1515011576), Guangdong, Hong Kong, Macao, and the Greater Bay Area International Science and Technology Innovation Cooperation Project (No. 2021A050530080, No. 2021A0505060011), Key Research Projects for the Universities of Guangdong Provincial Education Department (No. 2020ZDZX3031, No. 2022ZDZX1032), Jiangmen Basic and Applied Basic Research Key Project (2021030103230006670), Jiangmen Science and Technology Plan Project (2220002000246), and Key Laboratory of Public Big Data in Guizhou Province (No. 2019BDKFJJ015), Development of a Container Intelligent Panoramic Trademark Quality Inspection System Based on Machine Vision(HX22105).

References

1. Wang, S., Jiang, F., Zhang, B., et al.: Development of UAV-based target tracking and recognition systems. IEEE Trans. Intell. Transp. Syst. **21**(8), 3409–3422 (2019)
2. Liu, K., Mattyus, G.: Fast multiclass vehicle detection on aerial images. IEEE Geosci. Remote Sens. Lett. **12**(9), 1938–1942 (2015)
3. Razakarivony, S., Jurie, F.: Vehicle detection in aerial imagery: a small target detection benchmark. J. Vis. Commun. Image Represent. **34**(1), 187–203 (2016)
4. Xia, G.S., Bai, X., Ding, J., et al.: DOTA: A large-scale dataset for object detection in aerial images. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3974–3983. IEEE/CVF, Salt Lake City, UT, USA (2018)
5. Du, D., Qi, Y., Yu, H., et al.: The unmanned aerial vehicle benchmark: object detection and tracking. In: Proceedings of the European conference on computer vision, pp. 370–386. Springer, Munich, Germany (2018)
6. Zhu, P., Wen, L., Bian, X., et al.: Vision meets drones: a challenge (2018). arXiv preprint arXiv:1804.07437
7. Sun, Y., Cao, B., Zhu, P., et al.: Drone-based RGB-infrared cross-modality vehicle detection via uncertainty-aware learning. IEEE Trans. Circuits Syst. Video Technol. **32**(10), 6700–6713 (2022)
8. Hoiem, D., Divvala, S.K., Hays, J.H.: Pascal VOC 2008 challenge. World Liter. Today **24**(1) (2019)
9. Woo, S., Park, J., Lee, J.Y., et al.: Cbam: Convolutional block attention module. In: Proceedings of the European conference on computer vision, pp. 3–19. Springer, Munich, Germany (2018)
10. He, K., Zhang, X., Ren, S., et al.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778. IEEE/CVF, Las Vegas, NV, USA (2016)

11. Liu, S., Qi, L., Qin, H., et al.: Path aggregation network for instance segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp: 8759–8768. IEEE/CVF, Salt Lake City, UT, USA (2018)
12. Purkait, P., Zhao, C., Zach, C.: SPP-Net: Deep absolute pose regression with synthetic views (2017). arXiv preprint arXiv:1712.03452
13. Redmon, J., Farhadi, A.: Yolov3: An incremental improvement (2018). arXiv preprint arXiv:1804.02767
14. Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M.: Yolov4: Optimal speed and accuracy of object detection (2020). arXiv preprint arXiv:2004.10934
15. Li, C., Li, L., Jiang, H., et al.: YOLOv6: A single-stage object detection framework for industrial applications (2022). arXiv preprint arXiv:2209.02976
16. Wang, C.Y., Bochkovskiy, A., Liao, H.Y.M.: YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7464–7475. IEEE/CVF, New Orleans, LA, USA (2023)
17. Yu, D., Wang, H., Chen, P., et al.: Mixed pooling for convolutional neural networks. In: Rough Sets and Knowledge Technology: 9th International Conference, pp. 364–375. Springer, Shanghai, China (2014)
18. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7132–7141. IEEE/CVF, Salt Lake City, UT, USA (2018)
19. Vaswani, A., Shazeer, N., Parmar, N., et al.: Attention is all you need. Adv. Neural Inf. Proc. Syst. **30**(1) (2017)
20. Lin, T.Y., Dollár, P., Girshick, R., et al.: Feature pyramid networks for object detection. Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2117–2125. IEEE/CVF, Honolulu, HI, USA (2017)

Chapter 28
Research on the Application of QAR Decoding Technology in Civil Aviation

Yongliang Chen and Lijun Wang

Abstract The flight data and system parameters recorded in the quick access recorder of civil aircraft have become the basic basis for operation control and flight technology management, and also have a profound impact on the maintenance work. QAR (Quick Access Recorder) is an airborne flight parameter recorder for civil aviation aircraft, which has the advantages of fast transmission, strong timeliness, and convenient extraction. This paper analyzes the experience and methods in actual troubleshooting, extracts the effective information carried by the QAR data, and decodes and analyzes the recorded parameters, which can achieve accurate troubleshooting for aircraft maintenance, while reducing unnecessary maintenance work and component errors. The replacement rate provides strong support and guarantees for aviation safety.

28.1 Introduction

There are two main types of flight data storage recorders: FDR (Flight Data Recorder) and QAR/WQAR (Wireless Quick Access Recorder). The "black box" FDR that people often hear is the earliest flight data recording equipment. The outer shell of the FDR is made of steel plates and multiple layers of thermal insulation and shockproof protective materials [1]. The data can be saved, but the flight data in FDR is difficult to obtain, so the Quick Access Recorder (QAR) came into being. With the development of science and technology, QAR data can be directly transmitted wirelessly after the plane lands, making data acquisition more convenient [2].

Y. Chen (✉)
China Southern Airlines Technology Branch Haikou Maintenance Baseline, China Southern Airlines Group Co.Ltd.Line, Haikou, China
e-mail: cafuccyl@126.com

L. Wang
Hainan Health Vocational College, Haikou, China

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_28

28.2 Brief Introduction of QAR

The aircraft communication addressing and reporting system ACARS transmits the monitoring messages in the data management unit (Data Management Unit, DMU) to the ground system through very high frequency (Very High Frequency, VHF) to become the original message. However, the ACARS message needs to pay a high usage fee when using it. If the message is extracted through the QAR data, the expenses of the airline in this part can be saved. Modern aircraft are equipped with QAR, which records 128 MB and can record continuously for 600 h. At present, the QAR equipment used by most of the Airbus aircraft of domestic airlines has been converted into a wireless type, that is, WQAR. The transmission method of flight parameter data has been changed from manually reading the PC card to wirelessly sending it to the ground base station through the SIM card. At the end of each flight segment, the WQAR device on the aircraft will send the recorded data of the flight segment to a dedicated server. After receiving the data, the ground base station server automatically imports the original data into the decoding software (such as AGS (Air Generation System), AirFase, etc.) for decoding, and obtains flight information, flight parameter data, charts, etc. [3]. In the decoding process, the QAR overrun event is obtained by comparing the flight parameter data with the monitoring standard. For the fault information or warning information of the aircraft, the maintenance department can download the QAR data from the server through the decoding software and decode it. QAR can collect more than 1,000 kinds of data at the same time, and continuously record up to 600 h of original flight data. Therefore, it can provide an important reference for aircraft status monitoring, assisting aircraft maintenance and troubleshooting, etc., in an all-round way [4].

28.3 Application of QAR in Aircraft Fault Analysis and Diagnosis

In the process of fault analysis and diagnosis of aircraft systems, the fault phenomenon cannot be reproduced, and troubleshooting can only be carried out by relying on the information of the onboard maintenance system and the pilot's dictation. In the process of flight driving, the amount of information that pilots need to monitor and analyze is very large, and they can only focus on the parameters to ensure flight safety. Therefore, the after-the-fact oral presentation cannot provide sufficient support for maintenance [5].

The QAR data of the aircraft can accurately reflect the operation status of various system components during the use of the aircraft. Airlines and maintenance organizations can monitor and analyze the data to understand the working status of the aircraft more accurately and in a timely manner. Repair [6]. Only take the A320s aircraft airspeed inconsistent fault monitoring application as an example. The airspeed indication system includes eight Air Data Modules, each ADM (Air

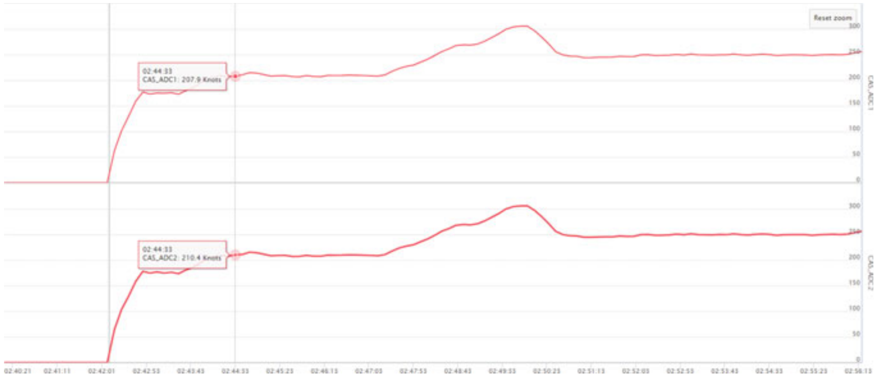


Fig. 28.1 B-30EE aircraft airspeed data change curve

Data Module) collects atmospheric data through pitot tubes and static pressure holes, converts it into airspeed data, and displays it on the cockpit indicating instrument. Every time the airspeed is inconsistent, the ADM accuracy test needs to be done when troubleshooting, which can easily lead to test errors, thus mistakenly replacing the ADM.

As can be seen from Fig. 28.1, the difference between the left and right airspeeds reached 3kts, and the number of consecutive occurrences exceeded 10. The ADM accuracy parameters were added to the QAR data, and the decoding found that the total pressure ADM on the captain's side was more accurate than other positions. The accuracy data of 2 is smaller than that of 2, so it can be preliminarily judged that the total pressure ADM on the captain's side is abnormal. Through the QAR decoding analysis, the airspeed inconsistency fault can be effectively eliminated, which not only reduces the delay risk of flight operation but also improves the safety level of flight operation.

QAR's whole-course record of aircraft flight parameters and system operation parameters can make up for the insufficiency of the airborne maintenance system and pilot's oral description, provide maintenance personnel with objective and comprehensive aircraft operation parameters, and reproduce the faults during the flight with data. Combined with information such as the fault code of the airborne maintenance system, and theoretical analysis based on the system principle, the source of the fault can be located and the right remedy can be achieved.

28.4 Application of QAR in Auxiliary Aircraft Maintenance and Troubleshooting

The structure of the aircraft system is complex. Using the results of QAR data analysis can effectively determine the true and false of the fault and improve the efficiency of troubleshooting. The use of QAR data to make fault judgments or conclusions should be cautious and objective, and should be verified by combining flight operations and fault descriptions in different flight stages, and must not be subjective [7]. Therefore, it is required to put the system operation situation reflected by each QAR parameter into a specific situation, and cannot regard the normal parameters under some special conditions as a fault condition.

FCU1 (Flight Control Unit 1) FAULT failures are frequent in the A320 fleet, some of which belong to the false FCU1 FAULT mentioned by Airbus in ISI (In-Service Information) 22.81.0009R0. Refer to Fig. 28.2 by analyzing the QAR/Digital Flight Data Recorder decoded data, it can effectively identify whether the false FCU1 FAULT fault. Because the false AUTO FLT FCU1 FAULT (Auto Flight Flight Control Unit 1 Fault) is generated with the push-in or pull-out action of the ALT SELECTOR (Altitude Selector) switch, it is possible to identify whether a certain FCU 1 FAULT is mentioned in ISI22.81.0009R0 by observing the operation of the switch at the time of the fault. false faults.

In the QAR/DFDR data, find the moment when the fault occurs, and observe whether the related Autopilot /Flight Director vertical guidance mode parameters change before and after this moment. We can judge whether there is a push-in or pull-out action of the ALT SELECTOR switch by observing whether the AP/FD (Autopilot/Fight Director) vertical guidance mode parameters change.

Fig. 28.2 Fault display of FCU on ECAM (Electronic Centralized Aircraft Monitor)

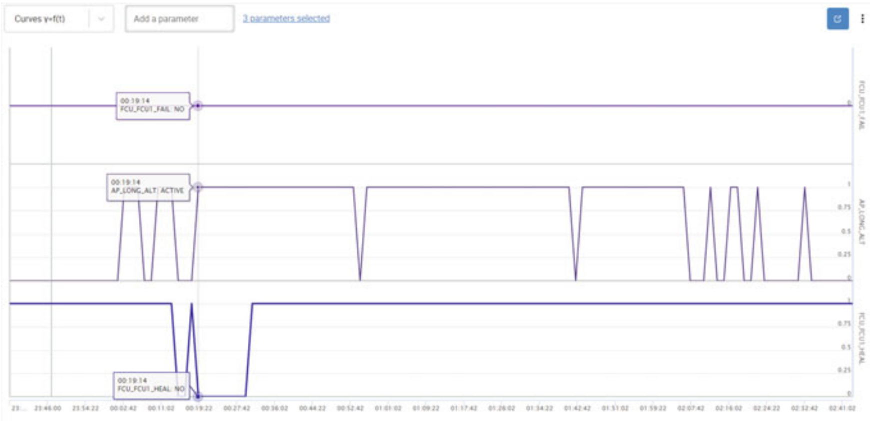


Fig. 28.3 B-8547 aircraft QAR data analysis

1. When the AP/FD vertical guidance mode is Climb or DES (Descent), when the parameter changes from 0 to 1, it means that the ALT selector switch is pressed at this time.
2. The AP/FD vertical guidance mode is OPEN CLB or OPEN DES. When the parameter changes from 0 to 1, it means that the ALT selector switch is pulled out.

Because the false AUTO FLT FCU1 FAULT is generated with the push-in or pull-out action of the ALT SELECTOR switch, it is possible to identify whether a certain FCU 1 FAULT is mentioned by ISI22.81.0009R0 by observing the operation of the switch at the time of the fault. false faults. In the QAR/DFDR data, find the moment when the fault occurs, and observe whether the related AP/FD vertical guidance mode parameters change before and after this moment. For example:

i. When the AP/FD vertical guidance mode is CLB or DES, when the parameter changes from 0 to 1, it means that the ALT selector switch is pressed in at this time;

ii. The AP/FD vertical guidance mode is OPEN CLB or OPEN DES. When the parameter changes from 0 to 1, it means that the ALT selector switch is pulled out at this time.

Note: When the target altitude and the actual flight altitude of the aircraft are not much different, the aircraft may maintain the current vertical guidance mode, resulting in no change in the parameters of the vertical guidance mode.

Note: Refer to AMM22-12–00 for AP/FD vertical guidance mode.

We can judge whether there is a push-in or pull-out action of the ALT SELECTOR switch by observing whether the AP/FD vertical guidance mode parameters change. When the system monitors an AUTO FLT FCU 1 FAULT warning, or the crew reports an AUTO FLT FCU 1 FAULT warning in the air, it is necessary to identify whether it is a false fault caused by the crew operating the ALT selector:

#	RawId	ACNO	FlightNO	TimeStamp	Route	MsgType	Warning DateTime	Warning ATA NO	Warning Flight Phase	Warning Desc	FR DateTime	FR ATA Number	FR Flight Phase	FR MSG	FR Source ID	Operation
1	358081845	B-8547	CZ6675	2306300240	DAZJHK/DSZLLL	MPF	230630001400	220000	06	AUTO FLT FCU 1 FAULT	230629234300	341234	02	[illegible]	RADAR 1	RawDetail PFR
2	358081845	B-8547	CZ6675	2306300240	DAZJHK/DSZLLL	MPF					230630001400	228112	06	AFS:FCU 20	AFS 1	RawDetail PFR
3	358067744	B-8547	CZ6675	2306300015	-------/-------	FLR					230630001400	228112	06	AFS:FCU 20	AFS 1	RawDetail PFR
4	358067647	B-8547	CZ6675	2306300014	-------/-------	WRN	230630001400	220000	06	AUTO FLT FCU 1 FAULT						RawDetail PFR

Fig. 28.4 B-8547 aircraft FCU monitoring failure

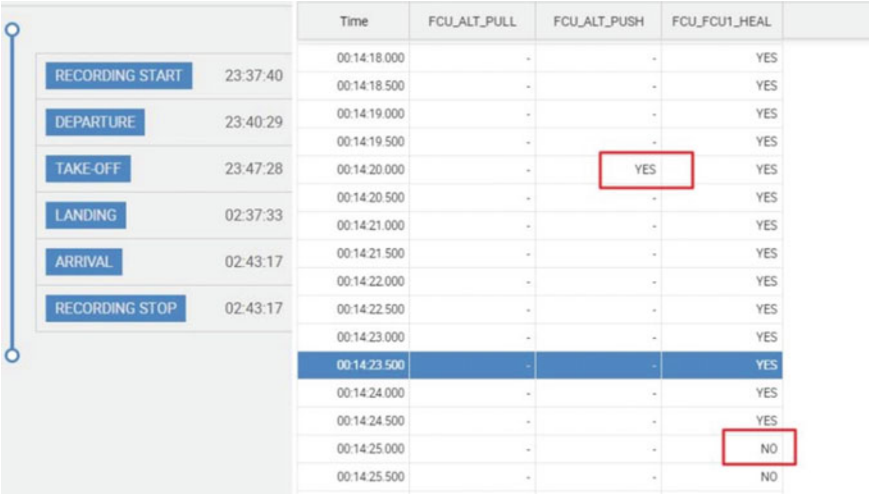
1. Check with the crew the timing of the warning—that is, whether the warning appears after the crew actuates the altitude selector. If it is confirmed that the ALT selector switch operates, it can be judged as a false fault.
2. Print the post-flight report and record the UTC (Universal Time Coordinated) time when the fault occurred. (You can also find the time when the warning is triggered through the FIADA (Feeling Information And Diagnosing Aircraft) system).
3. In the QAR/DFDR decoding data, find the moment when the fault occurs, and observe whether the relevant AP/FD vertical guidance mode parameters change before and after this moment.

The fault occurs in the cruise phase. When the fault occurs, the AP/FD vertical guidance mode parameters change, indicating that the ALT SELECTOR switch is operated, so it can be judged that this AUTO FLT FCU1 FAULT is a false warning.

It can be seen from Figs. 28.4 and 28.5, for the troubleshooting of some difficult faults, using QAR data for analysis can accurately point to the source of the fault, faithfully reproduce the fault situation, and confirm the fault condition and environment. Of course, this also puts forward higher requirements for troubleshooting personnel, requiring them to be familiar with the principle of the system, understand the operation of the system during the entire flight process, and the support that QAR data can provide, so that they can effectively use QAR to find the source of the fault and get Troubleshoot.

28.5 The Role of QAR in Fault Prevention

QAR also plays an important role in the early monitoring of faults. The failure of inconsistent airspeed is one of the failures that each maintenance base attaches great importance. It not only brings great pressure to flight operation but also has a serious impact on flight safety. Therefore, when the failure of inconsistent airspeed occurs, it is very easy to cause flight delay. The airspeed indication system includes eight ADMs. Each ADM converts the atmospheric data collected through the pitot tube and static pressure hole into airspeed data and displays it on the cockpit console. For each failure of inconsistent airspeed, an ADM accuracy test is required for troubleshooting, and this test is sensitive to equipment and instruments. If the equipment accuracy is not high, it is easy to lead to test errors, thus replacing ADM by mistake, resulting in an increase in the cost of air materials.



Time	FCU_ALT_PULL	FCU_ALT_PUSH	FCU_FCU1_HEAL
00:14:18.000	-	-	YES
00:14:18.500	-	-	YES
00:14:19.000	-	-	YES
00:14:19.500	-	-	YES
00:14:20.000	-	YES	YES
00:14:20.500	-	-	YES
00:14:21.000	-	-	YES
00:14:21.500	-	-	YES
00:14:22.000	-	-	YES
00:14:22.500	-	-	YES
00:14:23.000	-	-	YES
00:14:23.500	-	-	YES
00:14:24.000	-	-	YES
00:14:24.500	-	-	YES
00:14:25.000	-	-	NO
00:14:25.500	-	-	NO

Fig. 28.5 B-8547 aircraft QAR data analysis

In August 2020, flight CZ6773 (Haikou Guangzhou) will be executed, and the crew will report that the airspeed of the right seat is lower than that of the F/O and ISIS (integrated standby instrument) up to 40 knots when the runway of Haikou Airport runs at a high speed (the airspeed is more than 100 knots). After the AP (autopilot) is switched on, ECAM has an airspeed inconsistency warning, and the crew presses the ECAM prompt to set the right seat atmospheric data conversion switch to position F/O 3 (that is, the data of ADR3 used by the copilot) until landing. During this period, there is no other abnormal indication or information in the cockpit, and the altitude indication is normal throughout the whole process.

The Fig. 28.6 decoding counter checks the previous flight. On August 13, the difference between the left and right airspeed of the aircraft was large, about 5 kts (standard 6 kts). Therefore, the fault can be detected in advance.

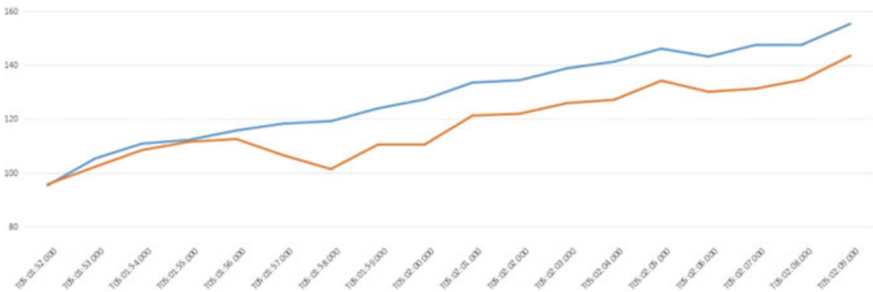


Fig. 28.6 B-8638 Left and right airspeed data change curves

time	Flight_Pha	CZSBTP1	CZSBSP1	CZFOTP1	CZFOSP2	CZFOSP1	CZCATP1	CZCASP2	CZCASP1	CO
06:43:01	1.00						[illegible]		1003.22	
06:43:02	1.00			1003.19		1003.20				
06:43:03	1.00		1003.34					1003.37		
06:43:04	1.00	1003.13			1003.33					
06:43:05	1.00						1001.82		1003.20	
06:43:06	1.00			1003.17		1003.18				
06:43:07	1.00		1003.28					1003.31		
06:43:08	1.00	1003.13			1003.31					
06:43:09	1.00						1001.80		1003.20	
06:43:10	1.00			1003.18		1003.18				
06:43:11	1.00		1003.30					1003.33		
06:43:12	1.00	1003.16			1003.33					
06:43:13	1.00						1001.80		1003.18	
06:43:14	1.00			1003.14		1003.16				
06:43:15	1.00		1003.30					1003.33		
06:43:16	1.00	1003.12			1003.30					
06:43:17	1.00						1001.82		1003.19	
06:43:18	1.00			1003.15		1003.16				
06:43:19	1.00		1003.27					1003.30		
06:43:20	1.00	1003.12			1003.29					
06:43:21	1.00						1001.79		1003.16	
06:43:22	1.00			1003.05		1003.16				
06:43:23	1.00		1003.27					1003.28		
06:43:24	1.00	1003.10			1003.27					
06:43:25	1.00						1001.80		1003.12	
06:43:26	1.00			1003.10		1003.15				
06:43:27	1.00		1003.28					1003.28		
06:43:28	1.00	1003.12			1003.27					

Fig. 28.7 DAR decoding B-8639 aircraft ADM accuracy data

With the support of a large number of QAR data and the analysis of flight speed data before the failure, most of the inconsistent airspeed failures can be warned in advance. Through the statistical analysis of big data, we can determine the threshold value of the average value of the difference between each two airspeed values and the threshold value of the frequency of continuous occurrence of the maximum difference value. Finally, the airspeed monitoring threshold is formed.

By determining the monitoring logic of each speed interval and airspeed difference, it is possible to give early warning to the faults of aircraft with inconsistent airspeed; By determining the ADM accuracy difference threshold, it can be accurately determined which ADM caused the failure of inconsistent airspeed, thus reducing the false removal rate of ADM.

Through the above monitoring logic, we have carried out a logical inspection of the air speed inconsistency faults since 2018. The results show that about 90% of the faults can be found in advance, indicating that the reliability of this set of monitoring logic is good.

The threshold set by QAR monitoring data can be used to give early warning through monitoring logic, accurately determine the fault source, prevent the occurrence of faults in advance, and reduce the occurrence of Aircraft On Ground events (see **Error! Reference source not found.**).

Airspeed is a very important flight parameter. Abnormal airspeed of A320s aircraft may cause serious flight accidents, and it is difficult to completely eliminate abnormal airspeed faults through ground inspection Aiming at the problems of high frequency of airspeed anomalies, multiple fault types, great impact on flight safety, and heavy burden on pilots to deal with after airspeed anomalies, the project adopts big data analysis method to develop the difference monitoring logic of airspeed in different speed intervals, which can find 90% of inconsistent airspeed faults in advance. By adding ADM accuracy parameters and developing the ADM accuracy difference threshold, the fault source can be effectively determined, the false removal rate of ADM can be reduced, and the cost of air materials can be saved. Later, this project will continue to track more fault cases, improve the monitoring logic, and finally promote it to China Southern Airlines A320s fleet, to improve the safe height of aircraft operation.

28.6 The Application of QAR in Ensuring Aviation Safety

The GPS (Global Positioning System) system needs to connect 4 satellites for accurate positioning [8]. After the connection, the GPS system enters the normal navigation mode. GPS signals are low-power signals. It is equivalent to the energy emitted by a 60W light bulb located more than 20,000 km from the earth's surface. GPS signals are susceptible to interference from any terrestrial signal sources emitting in the GPS L1 band (1575.42 MHz $\pm$ 10 MHz) located near the aircraft, resulting in loss of GPS data and a significant risk to flight safety. Using QAR data to carry out GPS signal loss analysis and accurately obtain the time, latitude, longitude, altitude, and other data of GPS signal loss and recovery, it can effectively improve the efficiency and timeliness of manual investigation of GPS signal interference sources, thereby reducing the hidden dangers of civil aviation flight safety.

The base station uses the AGS system to decode and analyze the raw QAR data received every day and uses the AGS to write the GPS signal loss monitoring event. When the judgment logic is satisfied during the decoding process, the GPS signal loss event will be triggered and the signal will be lost. Key parameters such as latitude, longitude, altitude, and speed at the moment are reserved and stored. By analyzing the performance of GPS signal loss in QAR data, and using GPS-related parameters recorded in QAR, a GPS signal loss monitoring program has been developed. Frequency analysis, the research results provide accurate positioning information for later elimination of interference sources, thereby significantly improving the efficiency of investigation and reducing flight safety hazards (see Fig. 28.8).

28.7 Problems Faced by QAR Data Application

The decoding and analysis of QAR data plays an increasingly important role in troubleshooting, and it provides strong technical support for maintenance personnel. Maintenance engineers should not only have a deep understanding of the aircraft operating system but also focus on summarizing and exchanging troubleshooting experience. They also need to master the parameters and characteristics provided by QAR, analyze the causes of failures behind the data, and cultivate the awareness and skills of using QAR for troubleshooting [9]. The analysis results of QAR data can not only improve the accuracy of fault location but also fundamentally improve the efficiency of aircraft maintenance work and reduce the waste of aviation materials

19/04/2020 Time	TIMR_R	FLIGHT_PHASE	HEIGHT (feet)	GPS_HIL_F0 (NM)	GPS_HOL_CA (NM)	GPS_HFCM_FO (NM)	GPS_HFCM_CA (NM)
22:42:48	22:42:44	DESCENT	21272	0.0313	0.0547	0.0078	0.0117
22:42:49	22:42:44	DESCENT	21260	0.0313	0.0547	0.0078	0.0117
22:42:50	22:42:44	DESCENT	21244	0.0313	0.0547	0.0078	0.0117
22:42:51	22:42:48	DESCENT	21220	0.0352	15.9961	0.0078	15.9961
22:42:52	22:42:48	DESCENT	21212	0.0352	15.9961	0.0078	15.9961
22:42:53	22:42:48	DESCENT	21168	0.0352	15.9961	0.0078	15.9961
22:42:54	22:42:48	DESCENT	21168	0.0352	15.9961	0.0078	15.9961
22:42:55	22:42:52	DESCENT	21152	0.0352	15.9961	0.0078	15.9961
22:42:56	22:42:52	DESCENT	21132	0.0352	15.9961	0.0078	15.9961
22:42:57	22:42:52	DESCENT	21116	0.0352	15.9961	0.0078	15.9961
22:42:58	22:42:52	DESCENT	21100	0.0352	15.9961	0.0078	15.9961
22:42:59	22:42:56	DESCENT	21054	0.0352	15.9961	0.0078	0.0117
22:43:00	22:42:56	DESCENT	21054	0.0352	15.9961	0.0078	0.0117
22:43:01	22:42:56	DESCENT	21044	0.0352	0.0547	0.0078	0.0117
22:43:02	22:42:56	DESCENT	21032	0.0313	0.0547	0.0078	0.0117
22:43:03	22:43:00	DESCENT	21016	0.0352	0.0547	0.0078	0.0117
22:43:04	22:43:00	DESCENT	20996	0.0664	15.9961	0.0117	15.9961
22:43:05	22:43:00	DESCENT	20976	0.0781	15.9961	0.0117	15.9961
22:43:06	22:43:00	DESCENT	20964	0.0781	15.9961	0.0117	15.9961
22:43:07	22:43:04	DESCENT	20940	0.0781	15.9961	0.0117	15.9961
22:43:08	22:43:04	DESCENT	20924	0.0781	15.9961	0.0117	15.9961
22:43:09	22:43:04	DESCENT	20912	0.0781	15.9961	0.0117	15.9961
22:43:10	22:43:04	DESCENT	20888	0.0352	15.9961	0.0078	15.9961
22:43:11	22:43:08	DESCENT	20876	0.0352	15.9961	0.0078	15.9961
22:43:12	22:43:08	DESCENT	20860	0.0352	15.9961	0.0078	15.9961
22:43:13	22:43:08	DESCENT	20840	0.0352	15.9961	0.0078	0.0117
22:43:14	22:43:08	DESCENT	20828	0.0352	0.4570	0.0078	0.0117

Fig. 28.8 A320 GPS signal interfered data graph

and human resources; the use of QAR data can effectively evaluate flight safety: a certain The analysis and evaluation of typical unsafe events, such as heavy landing, low-altitude wind shear, and overrunning the runway, are of great significance to the realization of flight quality monitoring and the improvement of airline flight quality. Ensuring the safety level of flight operations enables airlines to target Measures to improve the level of flight safety should be formulated accordingly [10]. Improve the safety and reliability of civil aviation flight and help airlines obtain better benefits.

References

1. Gao, Y.Y.: Exploring the application of airline QAR data. Civil Aviation of China, 2015(9) (2015)
2. Feng, X.J., Feng, X.R., Wang, Y.H.: Application of kernel function-based PCA in QAR data analysis. Comput. Eng. Appl. **045**(014), 207–209 (2009)
3. Lin, B.: Research on data release and location monitoring technology of sensor network based on internet of things. J. Web Eng **2021**(3), 689–712 (2021)
4. Cao, L.: Analysis of QAR principle and its application in aircraft maintenance. Jiangsu Aviation, 2011(1) (2011)

5. Gao, H., Li, B.: Establishment of motion model for wave capture buoy and research on hydrodynamic performance of floating-type wave energy converter. Polish Maritime Res. **22**(s1), 106–111 (2015)
6. Cao, H.L., Zhou, B.Z.: Application of QAR data in aircraft engine monitoring. J Civil Aviat. Univ. China 28(3) 5(2010)
7. Mao, X.X.: Reflections on the application of QAR technology in China's civil aviation. Civil Aviat. China **2010**(12), 2 (2010)
8. Businger, S., Chiswell, S.R., Bevis, M., Duan, J.P., Anthes, R.A., Rocken, C., Ware, R.H., Exner, M., VanHove, T., Solheim, F.S.: The promise of GPS in atmospheric monitoring. Bull. Am. Meteor. Soc. **77**(1), 5–18 (1996)
9. Dong, J.T., Wang, C.H., Sun, H.B., et al.: A study on the estimation of China civil aviation flight emission inventory based on QAR data. J. Environ. Sci. 42(12), 10 (2022)
10. Wang, L., Zhang, J., Dong, C., Sun, H., Ren, Y.: A method of applying flight data to evaluate landing operation performance. Ergonomics **62**(2), 1–20 (2018)

Chapter 29
Design of Business Application System Based on WeChat Mini-Program

Wenbin Yuan, Siyu Du, Tao Jiang, Yifeng Zhang, and Zhibin Li

Abstract With the growth of traditional e-commerce platforms sluggish, flexible, compact, and convenient diversified demands are rising day by day. This project uses the mini-program launched by the WeChat platform as a carrier to realize a "Silk Road Shop" business service platform under the "new retail" model. While providing users with a shopping system that is convenient, fast, and low in memory consumption, it also provides a new channel for the transformation of offline merchants. By adopting the WeChat MINA framework for customer construction, JavaScript language is used to realize functional logic. The back-end development adopts the MVC development mode, MySQL is selected as the project database technology, and the online simulation debugging shows that the system meets the requirements of the design goal.

29.1 Introduction

With the continuous improvement of network facilities, the promotion of third means of payment and the rapid development of the logistics industry, traditional offline stores urgently need to transform to meet the shopping needs of users under the

W. Yuan (✉) · S. Du · T. Jiang
Shaanxi Tobacco Company Xi'an Company, Xi'an, China
e-mail: 123709518@qq.com

S. Du
e-mail: 1217336526@qq.com

T. Jiang
e-mail: 1213396440@qq.com

Y. Zhang · Z. Li
Chang'an University, Xi'an, China
e-mail: 470978436@qq.com

Z. Li
e-mail: 2805632707@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_29

“new retail” model. At the beginning of the birth of e-commerce, it emerged in the form of web pages. Although it satisfied users’ online shopping experience, due to the majority of operating platforms being PC platforms, its portability experience was reduced. Later, with the continuous development of smartphones, e-commerce began to shift to apps on mobile platforms, further improving convenience. With the increasing competition in e-commerce, there are various types of shopping apps and unique services. For most consumers, there are few opportunities to use some of these features, but sometimes they have to use them, which leads to excessive waste of mobile phone resources. Developing and maintaining a practical and stable app is too expensive for most offline physical stores. In addition, it is extremely difficult to attract and promote a new type of app during this traffic saturation stage. At present, the shopping platform developed with APP as the carrier occupies a dominant position in the field of e-commerce, and its business model is still the traditional B2C model. The biggest problem with this model is that it bypasses offline physical stores and does not solve the problem of the gradual separation of users and merchants under the “new retail” model. With the launch of the WeChat Mini-Program, there are new solutions to the above problems. Compared to apps, WeChat mini-programs are a brand new way to provide services to users. His efficiency and convenience have been further improved. This is mainly reflected in the fact that users do not need to download and install, and do not need to occupy too much memory to obtain and use mini-programs. The entire process can almost be considered as out of the box, and the stability and security of the mini-program have not been reduced as a result.

The “Silk Road Shop” project is a business service platform developed on the platform of WeChat Mini-Program. The project made full use of the WeChat Mini-program’s large number of users, convenient operation, and low memory consumption. It successfully attracted the attention of consumers and quickly opened up the market. While supporting online shopping, the system will also recommend surrounding offline physical stores to users, so as to shorten the distance between the two. It not only meets the online shopping needs of modern people but also provides a solution for the transformation of traditional shops.

29.2 Development Platform

Since its launch in 2017, the WeChat Mini-Program has been favored by a large number of users for its features such as simple use, low memory consumption, and good stability. All kinds of applications have moved to the WeChat Mini-program market. In order to promote the further development of the mini-program ecology, the WeChat team launched a “developer tool” specifically for developing mini-programs [1]. As the development platform of this project, “Developer Tools” provides a set of MINA (Multipurpose Infrastructure for Network Applications) frameworks including various API interfaces, which is the only optional architectural method for the development of mini-programs. The MINA framework is a framework with a three-layer structure, which is composed of view layer, logic layer, and system layer. Its essence is a data binding mode, that is, the MVVM (Model–View–View-Model) mode. The data binding mode is optimized from the classic MVC (Model-ViewControl) mode, using this method can make the data and view consistent easily [2]. The frame diagram of MINA and the example diagram of MVVM mode are shown in Figs. 29.1 and 29.2, respectively.

As shown in Fig. 29.1, the MINA framework mainly includes the view layer, logical layer, and system layer.

The development of the view layer utilizes the WXML language optimized by the HTML language and a series of basic components. Developers use WXML files to construct the basic view structure of the page and use WXSS files to control the presentation style of the page. As shown on the visual layer, the basic units used are a series of basic components. The general steps of the entire display process are as follows: first, the data is processed by the logical layer; then sent the processed data that needs to be displayed to the viewing layer; and finally, basic components are utilized to achieve data display.

The development of the logic layer uses JavaScript language. Based on JavaScript, the WeChat team has made some improvements to improve the efficiency of developing mini-programs. The main modifications include using application and web

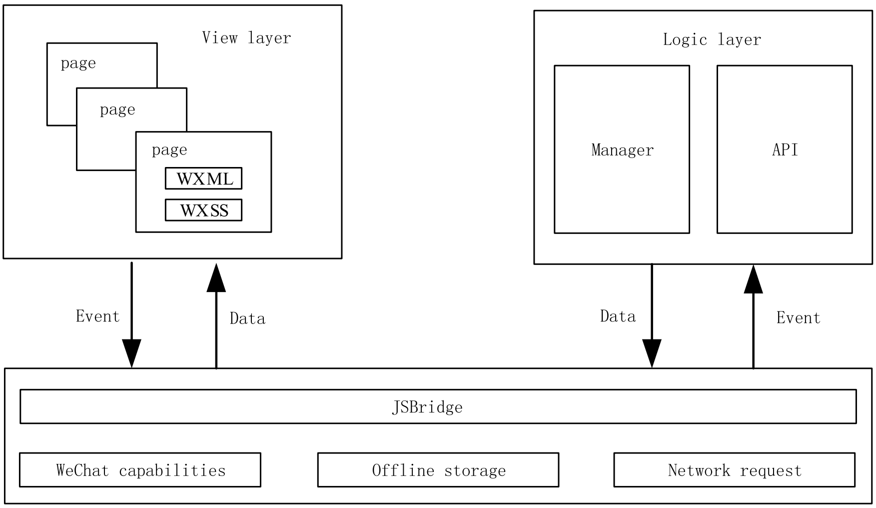


Fig. 29.1 MINA framework

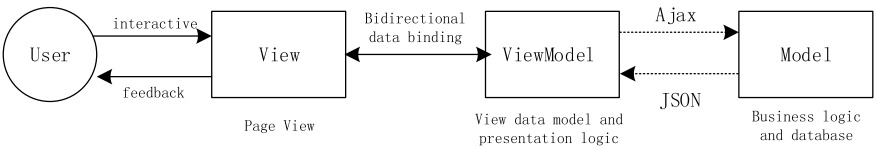


Fig. 29.2 MVVM pattern

page methods to register programs and web pages; provide rich APIs, such as WeChat Pay and other specific functions; each page has an independent scope and provides modular functionality.

The system layer is a relatively vague concept, different from the view layer and logic layer. The system layer has strong functionality, and its functions are diverse and complex. Overall, the system layer is equivalent to a data cache, similar to the design of a computer CPU. Due to the fast data processing speed of the logical layer and server, there is no direct data transmission between the view layer and the logical layer, and they are cached and retrieved through the JSBridge in the system layer. The functions of the system layer are not singular, including data storage, network request initiation, and some WeChat-specific functions. In addition, app. wxss, app. js, and app. json are global system files that form the main part of the mini-program and are all configured under the root directory. Js and json are files responsible for the logic and public configuration of the mini-program, and are essential files for the mini-program. Wxss is responsible for the public stylesheet of the mini-program, and is not a required file.

The MVVM mode is shown in Fig. 29.2. MVVM, also known as the Model–View–ViewModel, is essentially an improved version of MVC. View refers to the page you see. The view model is the core of theMVVM pattern and serves as a bridge connecting views and models. It has two directions: one is to convert the model into a view, that is, to convert the data transmitted by the backend into the page it sees. The implementation method is data binding. The second is to convert the view into a model, that is, to convert the page you see into backend data. The implementation method is: DOM event listening. Both directions are implemented, which we call bidirectional binding of data. The MVVM pattern helps to clearly separate the business and presentation logic of an application from the user interface (UI). Maintaining a clear separation between application logic and UI can help solve many development problems and make applications easier to test, maintain, and evolve. It can also significantly improve code reuse opportunities and allow developers and UI designers to collaborate more easily when developing their respective parts of the application.

29.3 Technical Solution

29.3.1 Client Program

The main content of client development is the page layout of the view layer and the writing of data processing methods in the logic layer. In the process of client development, we adopt an object-oriented program development scheme. In the context of this scheme, we regard each display interface required in the applet as the smallest unit. In this way, in addition to the theme files required for global configuration, the root directory of the applet also contains subfolders responsible for the display functions of each page [3]. The MINA framework used by the applet is mainly composed of four files, namely app.js, app.json, app.wxml, and app.wxss. Among them, the first two are respectively responsible for the definition of the applet logic and the configuration of public parameters. These two belong to the main file of the entire project and exist in various folders in the project. The latter two only exist in the subfolders of each page, and the wxml file plays a major role in it, which determines the display structure of the page to which it belongs. The basic unit of wxml file to realize the function is "functional component". The file binds the components with data and transmits the bound information to the logic layer through the system layer. The app.js file in the logic layer processes relevant information and returns the processing results to the view layer through the system layer. The wxml file will call related components to refresh the page display according to the processing information. The wxss in the view layer is an additional file used to assist wxml to implement functions and is not a required file in the project [4].

29.3.2 Database Scheme

The design of this project includes functional modules related to data management, such as user information management, order generation, and product information storage. In order to ensure the security and stability of information storage and usage, a reasonable database solution must be adopted during backend development. At present, there are two main database technologies, which can be divided into relational database and non-relational database according to the different ways of data storage. The most typical structure of relational database technology is a table, which is usually a data set composed of two related quantities [5]. Its advantages mainly include consistent format; simple operation and universal SQL language; can support complex queries. However, it only applies to situations with low data volume.

Strictly speaking, non-relational database technology cannot be regarded as a kind of database, because this kind of technology only applies a data-structured storage method. This has the advantages of flexible storage format and fast read/write speed. Compared with relational database technology NoSQL, it can not only use hard disk as a carrier but also use random access memory (RAM) for high scalability. The disadvantage is that the cost is high; the data structure is complex and does not support complex queries.

By analyzing two database technologies and combining the characteristics and functional requirements of the mini-program itself. It is more appropriate to select relational database. This project selects MySQL database in relational database. Because the data required by the mini-program is not large, MySQL can well meet this requirement. And, its lower maintenance cost and convenient usage can further reduce the cost of mini-programs and increase the user and merchant experience [6].

29.3.3 Payment Scheme

Due to the consideration of the entire transaction process, this project has selected WeChat payment technology provided by the WeChat team as the payment method. The use of WeChat payment technology can easily enable users to complete a closed-loop transaction and achieve the monetization of social e-commerce.

Unlike our understanding of the types of WeChat payments in daily life, the classification of WeChat payments during the development process is no longer as simple as actively scanning or being scanned. During development, there are mainly four channels: APP, applet, WeChat internal browser, namely, WeChat official account, and WeChat external browser. When developing mini-programs, if you need to use the WeChat payment function, you can directly call the wx.requestPayment interface.

29.4 Systems Design

29.4.1 System Function Requirements Analysis

By analyzing the business process of the system, we can conclude that the core objects in the entire business process mainly include users, system front-end, system server, WeChat server, and local cache. Combined with the data flow of running the mini-program, we can divide users into tourists and loggers. When using the mini-program, tourists can open the mini-program and perform basic functions such as page browsing. When users want to modify their personal information and place orders for services such as purchasing, they are required to log in. Thus, the function Use Case Diagram of the applet client user's system can be classified and listed, as shown in Fig. 29.3.

The front-end design of this project adopts an object-oriented programming concept. Firstly, an overall analysis of the system functions is conducted to clarify the

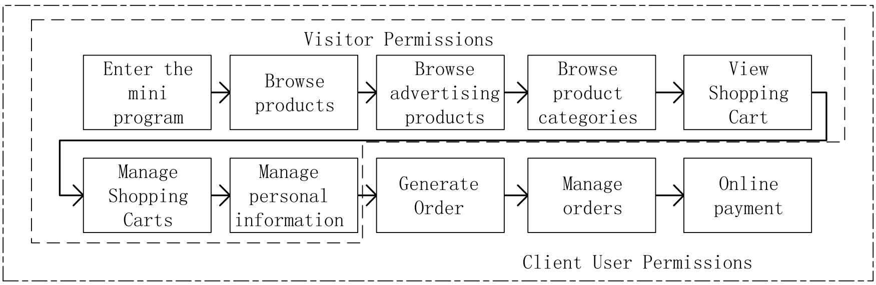


Fig. 29.3 Address scrolling selector

functional requirements that the system needs to achieve. On this basis, development is carried out in modules, making full use of typical object-oriented development methods such as encapsulation and inheritance, so that each module can meet the technical specifications. During the design process, it is required that there should be no repetitive functions between functional blocks, and the functional blocks should be refined as much as possible to ensure that the entire project runs on each functional block as a basic unit, avoiding repetitive development and greatly improving the reliability and usability of the program. Finally, organize and integrate various modules to make the entire system structure clear and easy to manage.

29.4.2 Overall System Architecture

The "Silk Road Small Shop" business enterprise service platform is divided into three layers in terms of system architecture, including the view layer, logic layer, and system layer. The view layer, as the front-end part of the client, is a direct place to achieve page rendering and page interaction. The reasonable design of the view layer directly affects the user experience. The logical layer serves as the service center of the entire project architecture, and the logical processing and data calls required for page interaction and page display are all implemented in this layer [7]. The system layer is a relatively abstract concept, and there is no direct data exchange between the view layer and the logic layer. All data exchange requires the system layer as a transit station.

29.4.3 System Functional Design

Before conducting system functional design, my team conducted thorough market research and feasibility analysis while also taking into account the customer needs under the current "new retail" model in China. The core functional modules of the system are shown in Fig. 29.4. The core functions of this project were determined to include product information display, order generation, user information management, order payment, and nearby recommendation functions. The most distinctive feature among them is the recommendation function of nearby shops. Through the built-in GPS positioning function of WeChat, "Silk Road Shop" can recommend relevant information of nearby franchise shops to users [8]. This can effectively solve the problem of gradually separating users from offline physical stores in the "new retail" environment. After the functional design is completed, functional modules need to be divided, which is the basic idea of object-oriented programming methods. By subdividing each function into multiple distinct functional units, not only can resource waste caused by repeated development be avoided but also program availability can be increased [9].

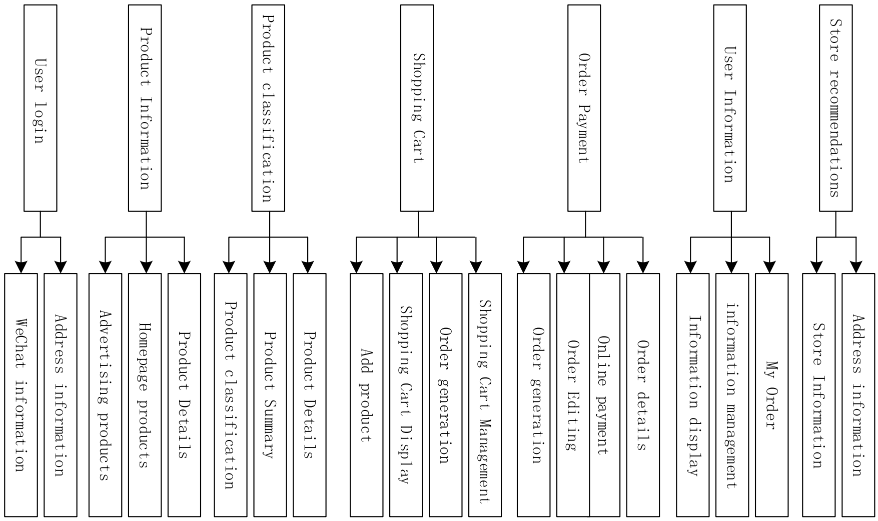


Fig. 29.4 System function module

29.5 System Simulation and Operation

Due to the WeChat developer tool's built-in real machine simulator, the simulation process can be directly conducted on the open platform without the need to purchase different models for related operations. Through a real machine simulator, we can simulate the relevant operations performed by users during use. Users can access the mini-program through WeChat's built-in scanning or search function. After entering the mini-program, they will receive a temporary account without real name authentication [10]. By clicking the "My" button, you can enter the user information management interface for real name authentication. Only authenticated accounts can be used for placing orders, making payments, and other operations. By clicking on the menu bar directly below the page, users can view product categories, manage shopping carts, obtain information about nearby stores, and manage shipping addresses. After all the information is filled in, the JavaScript file will use the API to retrieve the user's token information. If the user is not logged in, or if the token information has expired. The JS file will once again use the ShowModel function to pop up a prompt box indicating that the user's password is invalid. Details are shown in Figs. 29.5 and 29.6.

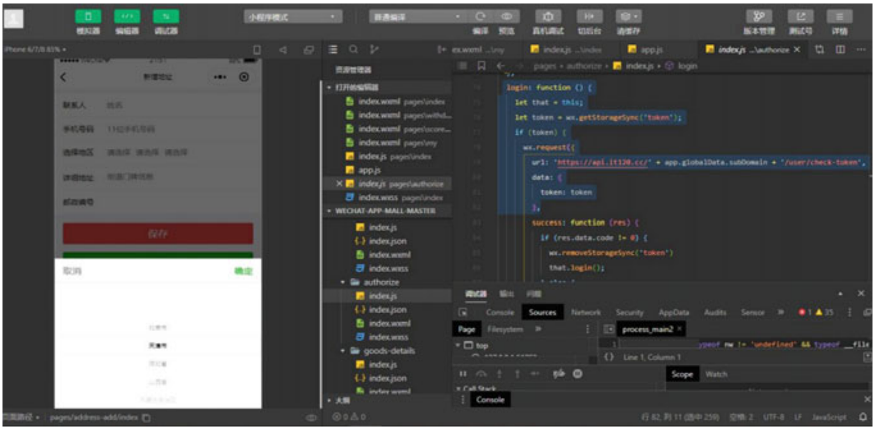


Fig. 29.5 Address scrolling selector

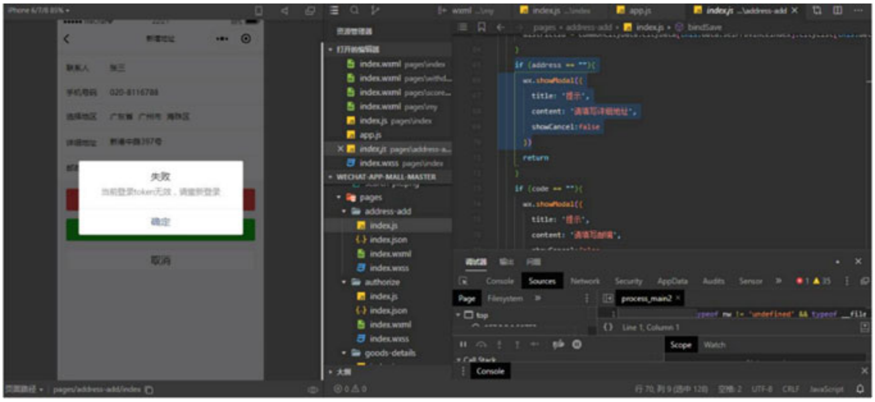


Fig. 29.6 Password failure display

29.6 Conclusion

The "Silk Road Shop" business enterprise service platform based on WeChat Mini-Programs fully utilizes the advantage of a large number of WeChat client users. This reduces the high costs required for early drainage and facilitates the promotion of enterprises. At the same time, mini-programs quickly gained popularity among users due to their convenient use and low memory consumption [11]. The "Silk Road Shop" not only provides the necessary online shopping needs of the current consumer group but also enhances the connection between consumers and offline stores. Thus, providing new channels for the transformation of enterprises under the "new retail" model. In the development of the system, using WeChat supporting developer tools not only enhances project compatibility but also further reduces the difficulty of development. This enables the system to achieve the expected goals at a lower cost.

References

1. Sung, H., Ferlay, J., Siegel, R.L.: Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. **71**(3), 209–249 (2021)
2. Díaz-Verdejo, J.E.: A methodology for conducting efficient sanitization of http training datasets. Futur. Gener. Comput. Syst. Gener. Comput. Syst. **109**, 67–82 (2020)
3. Wu, X.L.: Design and implementation of college online learning system based on WeChat mini program. In: 2018 9th International Conference on Information Technology in Medicine and Education (ITME). IEEE (2018)
4. Bjeladinovic, S., Marjanovic, Z., Babarogic, S.: A proposal of architecture for integration and uniform use of hybrid SQL/NoSQL database components. J. Syst. Softw.Softw. **168**, 110633 (2020)
5. Petković, D.: SQL/JSON standard: properties and deficiencies. Datenbank-Spektrum **17**(3), 277–287 (2017)
6. Nagpal, B., Chauhan, N., Singh, N.: SECSIX: security engine for CSRF, SQL injection and XSS attacks. Int. J. Syst. Assur. Eng. Manag.Manag. **8**, 631–644 (2017)
7. Ivanovich, B.V., et al.: Using MVC pattern in the software development to simulate production of high cylindrical steel ingots. J. Cryst. Growth **526**, 125240 (2019)
8. He, G.M., et al.: Design and implementation of WeChat mini program reservation system. In: ICETIS 2022; 7th International Conference on Electronic Technology and Information Science. VDE (2022)
9. Jin, F.: Development of Household Internet of Things System based on WeChat Mini Programe. Zhejiang University (2019)
10. Nichols, S., Rainbird, S.: The mall, the library and the church: inquiring into the resourcing of early learning through new spaces and networks. Int. J. Qual. Stud. Educ. **26**(2), 198–215 (2013)
11. Chen, Z.Z.: The Design and Implementation of Online Shopping System Based on WeChat Mini Program. Capital University of Economics and Business (2018)

Chapter 30
Influence on QoE in Different Physical Environments for PUBG Game

Sohail Basheer, Asif Ali Laghari, Muhammad Ibrar, Zuhaib Phul, and Mansoor Ahmed Khuhro

Abstract The gaming industry is growing daily, and gaming traffic is now a significant part of internet traffic. Gaming is continuously moving from desktop to portable devices through cloud environments. Game development organizations are generating vast amounts of revenue through these games. PUBG is a world-renowned online game popular among all age groups, especially teenagers. People play games anywhere as per their choice and availability of time. In this research, we invite PUBG game players to play games in different physical environments, like while traveling by Car or through Public Transport and while staying at home or during a visit to the park. We assess Quality of Experience (QoE) through the International Telecommunication Union (ITU) scale. Data was collected through a survey form, and collected data was analyzed through SPSS. Two variables are the Game's Speed and the Satisfaction Level with given physical environments. Results show that the QoE is high in the home physical environment over a WIFI internet connection for both variables. Researchers assess the user QoE in different environments, such as on Desktops, Mobile phones, or different machines of different configurations. Still, researchers have yet to work on the QoE of a game's players who play the game in the physical environment. Users rate the game on Google Play Store. Currently, no method exists to check the user's physical background. This research will help Game development organizations to look into the reason for a low rating due to players' comfortable physical zone.

S. Basheer · A. A. Laghari (✉) · Z. Phul
Department of Computer Science, Sindh Madressatul Islam University, Karachi, Pakistan
e-mail: asiflaghari@synu.edu.cn

M. Ibrar
Software College, Shenyang Normal University, Shenyang, China

M. A. Khuhro
Department of Artificial Intelligence, Sindh Madressatul Islam University, Karachi, Pakistan

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_30

30.1 Introduction

Online gaming is going popular nowadays among all age groups peoples, especially teenagers. PUBG (Player Unknown's Battleground) is one of them [1]. It is a mission-based game where a team has to fight until they remain alone on the battlefield [2]. PUBG was officially developed and launched by PUBG Corporation in December 2017. The idea of this game is based on a movie named "Battle Royale" released in Japan. This game provides a virtual battle environment where a team of people fights against each other at the end remaining person or team will declare as winner team. It was PC-based game. Later, its mobile version was introduced, which increased its popularity and rating. Its rank is greater among world-renewed games [3]. It has been downloaded over 500 million times from Google Play Store [4].

The average daily user of PUBG games is 1,57,230 [5]. PUBG generated 24 million dollars in revenue in December 2021 [6]. The trend of online gaming is increasing day by day. There are all types of traffic moving through the internet. Now gaming traffic has a significant part of internet traffic [7]. There are many ways of generating income through gaming; one way of earning money is by selling in-app features of Google Play Store, which is also done by PUBG corporation [8]. It is estimated that mobile game application revenue is about $69.7 billion through in-game advertising; it may rise to more than $188.9 billion. Compared to conventional or digital purchase systems, the in-game purchase is more relevant to game players' commitment to the game [9].

Speed of internet connection is the key term for online gaming. Due to the portability property of mobile phones, gamers can play games everywhere. Game players use different internet connectivity options, like 4G depending upon the local environment [10]. Many types of high-speed networks, like 3G, 4G (LTE), and 5G, are available for online gaming and video streaming to be used anywhere and anytime through wireless communication devices [11].

Existing solutions only cover QoE of gaming in different environments like desktop, mobile, cloud, and on different devices, but they do not cover QoE in different physical environments. Here our main objective is to assess PUBG game players' QoE in different physical environments concerning the Game's Speed and Level of Satisfaction. This research is limited to four physical environments for playing the PUBG game, like traveling in a car or public transport and staying at home or in public parks. We provide a way to assess user QoE in these physical environments. This method will be beneficial to assess the actual QoE due to the user's existing physical environment.

30.2 Literature Review

Several researchers have suggested that the architecture of games can improve the QoE [12–14]. It also improves the level of optimization [15]. Different researchers work on the quality of experience of games like Subway Surfers, Need for Speed, and Underground II in different environments. They found different results on Desktop PC and mobile devices [16]. In-home game streaming and online gaming are two modes of game playing. In which QoE was conducted, results show that game players preferred online gaming over in-home streaming [17]. Online game players show high expectations about the quality of services from the game operator. Game operators take care of those things, which can affect QoE. These studies suggest that gaming is moving from traditional PC-based gaming toward mobile gaming, so there is a need to explore the QoE of the quality of the game and user behavior [18].

QoE is the key term to determine the quality of the game. Online gaming requires real-time interaction between players; unfortunately, players are not satisfied with the quality of services. Twelve million sessions of online games were recorded, of which 7.12% of sessions were terminated not in the normal way, and 13% of users suffered from resynchronization issues at least in one location [19].

With the popularity of mobile phones, users of online games have increased. Some online game players prefer massively multiplayer online role-playing games on their mobile phones. To explore user behavior, a tree model was presented. Game players have some pre-consideration of the game. Their proposed decision model shows 85.62% accuracy of predictive rate about the game player decision [20]. The PUBG game is heavy; it requires more storage area, RAM, and processing power. It is tested and found that Samsung S9 Plus is better than iPhone X by user experience [21]. Online gaming and live streaming are open research areas where QoE is a measure for continual improvement. QoE suggests that the PUBG mobile platform is an excellent choice for playing games [22].

PUBG is considered a best-selling gaming application, continuously working to meet the demand of its user. Factors that affect user experience are player interest, motivation, addiction, and behavior [23–25]. Minor changes in elements of the game have an impact on players' experience and emotions [26]. After getting easy access to a high-speed internet connection, Massively Multiplayer Online Role Playing Games (MMORPG) like PUBG became more popular among game players. People who play these types of games from different locations have different cultural values, impacting different game players, like speaking a foreign language, tolerating differences, accepting defeat, etc. [27]. PUBG is a third-person shooter game. Its QoE was accessed based on three subscales GUESS-24, ENJOY, and UEQ-s [28].

Psychological effects on the game player are common questions raised by the scientific community and the public. To overcome this conception, a study was conducted to suggest that game players are critical thinkers. Results show that the game players who design games prefer to play games on personal computers, and the game players, which not design games prefer to play games on a mobile device [29].

Gaming development industries include many things in the game, like the dissemination of games and cultural influence [30]. Many random factors, such as player skills, can affect a player's ability, locations, play zone, airdrop, and accuracy of the weapon [31].

30.3 Experimental Methodology

We study related QoE-based research papers to access the QoE in different environments. After the study of related data, a survey form was prepared. We gathered data from thirty PUBG players who were willing to experiment. We offer different physical environments (Car, Public Transport, Players' own Homes, and Parks/Public places) to play the PUBG game. A survey form fills by the game player. QoE as feedback from the game player was recorded according to the provided physical environment given in Figs. 30.1, 30.2, 30.3, and 30.4, respectively, for two variables: the Speed of the Game and Level of Satisfaction. Users also tick the type of internet connection that they use for the experiment, as shown in Table 30.4 and Fig. 30.9.

Later we assign a rating scale to the Level of Satisfaction and Speed of the Game according to user experience. ITU scale is used to assess the QoE of users is given in Table 30.1 [32].

SPSS is a widely used statistical analysis software for data entry, calculation, and analysis of data. SPSS used to analyze and finalize the data. SPSS was used to keep survey form data in tabular form. We calculate the score of each experience, i.e., *Speed of the Game* and *Level of Satisfaction* in different physical environments. We calculate the mean score to obtain the single value of experiences of each group. For a

Fig. 30.1 Playing PUBG game while traveling in vehicle by using 3G/4G internet connection

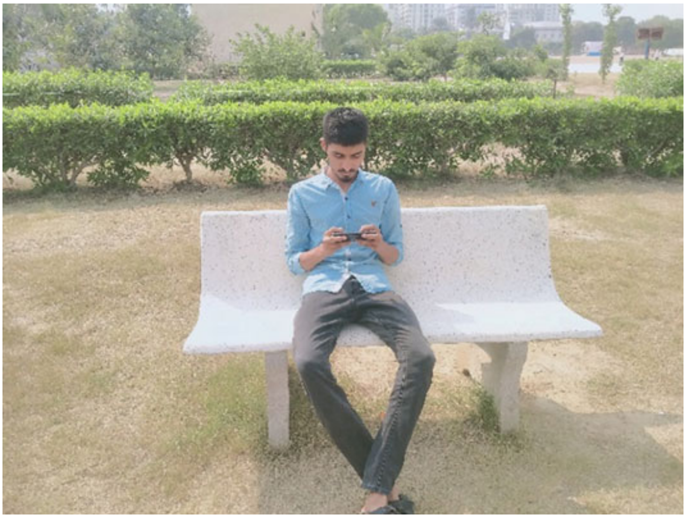


Fig. 30.2 Playing PUBG game in park/public place by using 3G/4G internet connection

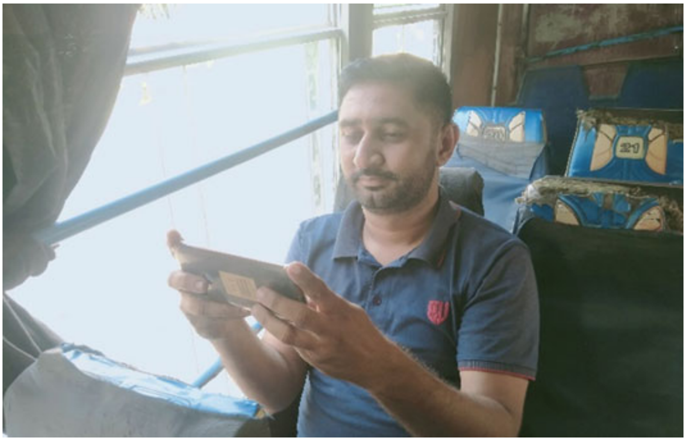


Fig. 30.3 Playing PUBG game in public transport by using 3G/4G internet connection

better understanding graph of each variable is also generated, as shown in Figs. 30.5, 30.6, 30.7, and 30.8.

Fig. 30.4 Playing PUBG game at home by using 3G/4G or WiFi internet connection

Table 30.1 ITU scale

Quality	Excellent	Good	Fair	Poor	Worst
Score	5	4	3	2	1

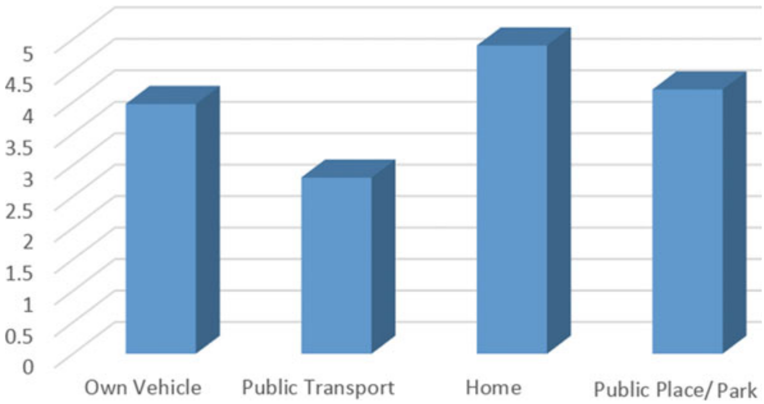


Fig. 30.5 Speed of game (different physical environments)

30.3.1 Data Collection

The data is collected from PUBG game players after playing PUBG games in different physical environments—collected through a survey form of willing game players.

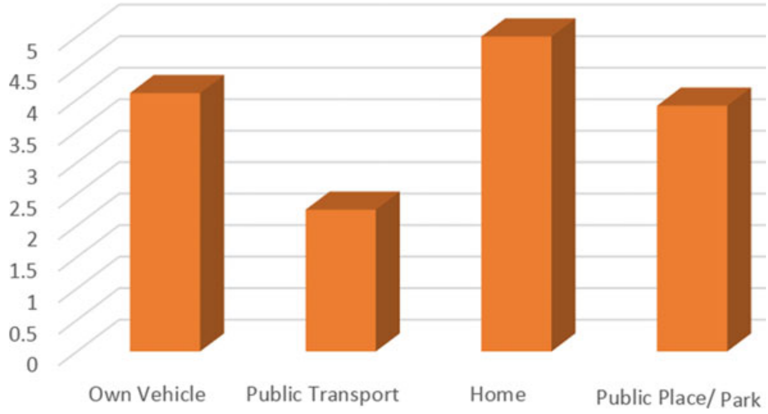


Fig. 30.6 Level of satisfaction (different physical environments)

Fig. 30.7 Comparison of speed of game and level of satisfaction (different physical environments)

30.4 Results and Discussion

Here we have four physical environments Own Vehicle, Public Transport, Home, and Public Place or Park, and two variables, *Speed of Game* and *Level of Satisfaction*, with thirty sample sizes.

Each variable has 30 values for each environment taken from the survey form. SPSS is considered a good tool for data analysis. Collected data put to SPSS sheet for calculation of mean score and standard deviation for data analysis and interpretation. We have calculated the mean score and standard deviation of each physical environment, also given in Tables 30.2 and 30.3. Table 30.4 shows the type of internet connection of the sample size.

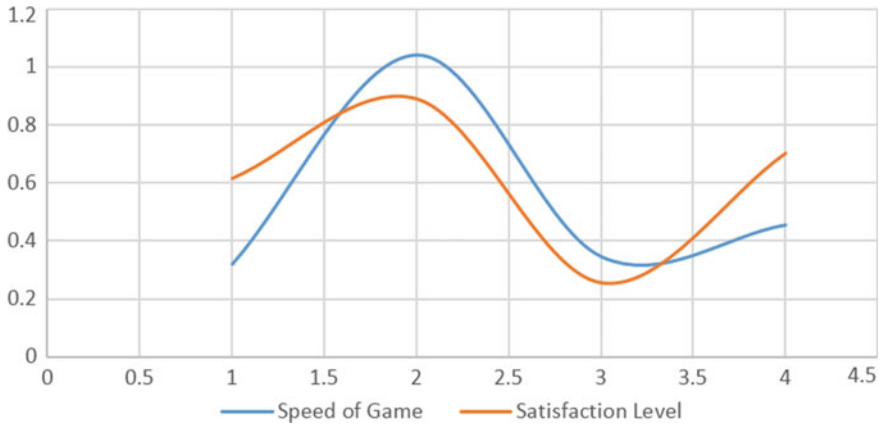


Fig. 30.8 Standard deviation graph of both variables (speed of game and level of satisfaction)

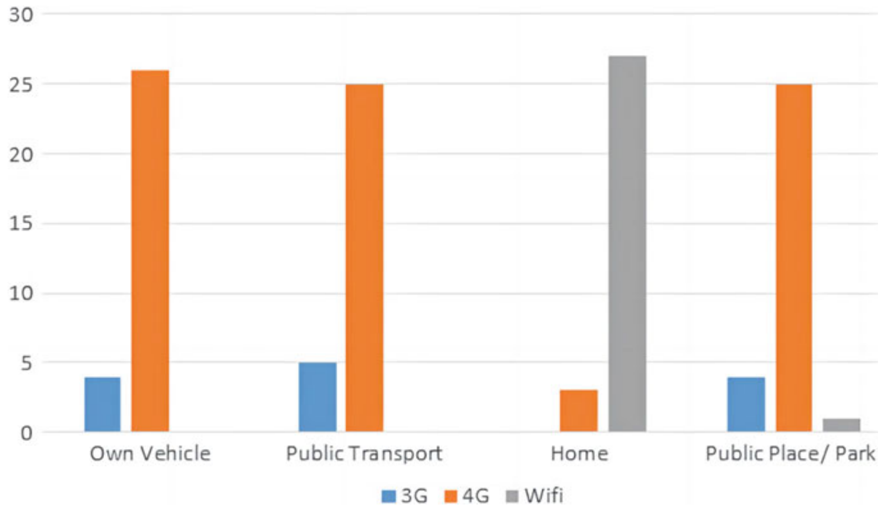


Fig. 30.9 Type of internet connection of sample size (in different physical environments)

Table 30.2 Speed of game (in different physical environments)

	Own vehicle	Public transport	Home	Public place/park
Mean	3.97	2.57	4.87	4
Std. Deviation	0.320	1.040	0.346	0.455

Table 30.3 Level of satisfaction (in different physical environments)

	Own vehicle	Public transport	Home	Public place/park
Mean	3.97	2.03	4.93	3.70
Std. Deviation	0.615	0.890	0.254	0.702

Table 30.4 Type of internet connection (in different physical environments)

	Own vehicle	Public transport	Home	Public place/park
3G	04	05	0	04
4G	26	25	03	25
Wifi	0	0	27	01

With the help of mean score and standard deviation values, graphs are designed as Fig. 30.5 *Speed of game* and Fig. 30.6 *Level of Satisfaction*, Fig. 30.7 Comparison of both variables, and Fig. 30.8 standard deviation score. Figure 30.9 shows the type of internet connection of the sample size.

After calculating mean score as shown in Tables 30.2 and Table 30.3, which is based on the player's QoE of different physical environments (Car, Public Transport, Home, and Parks) and type of internet connection, we develop a graphical representation of results for better understanding.

Figure 30.5 indicates that the *Speed of the Game* is better in the *Home* environment than in other physical environments. Figure 30.6 shows that *the Level of Satisfaction* is better in the *Home* environment than in other physical environments. Figure 30.7 shows the comparative scheme of *the Speed of the Game* and *the Level of Satisfaction* of game players, which shows that *the Speed of the Game* and *Level of Satisfaction* of both QoEs are high concerning other physical environments. Figure 30.8 shows variation in QoE in different physical environments for both variables, i.e., *Speed of Game* and *Level of* Satisfaction. Figure 30.9 shows the type of internet connection which game players used during the experiment. Mobile data 4G seems common outside of the home environment, and the ratio of 3G is less than 4G. WiFi commonly used in home environments. The game player's QoE is high over a WiFi internet connection. Many factors affect results. Speed of internet connection, 3G/4G coverage area, and variation in signal strength due to traveling, commonly known as handoff in telecommunication. The Hardware configuration of the cellular device and signal strength also plays an important role in QoE.

Survey form data is kept as records for authenticity and reliability, and photographs of willing game players have been taken. Collected data is processed, elaborated, and analyzed. If game players have a good configuration phone, then the speed of the game and level of satisfaction of game players increase. At the same time, if users have good configuration devices but are present in a crowded area or uncomfortable zone like traveling on public transport, then QoE will get worse. Hence, the above facts and figures show that user QoE differs for all environments. Here, we found that Game Player's QoE in the home environment is higher than in other environments for both variables: The Speed of the Game and Level of Satisfaction.

30.5 Conclusion

Google Play Store is considered the most reliable source to install phone applications; gaming applications are one of them. Game players download games from *Google Play* Store. *Play Store* asks users to assign a rating and give comments. Based on rating *Play Store* assigns a rating score to the application, which is publicly visible to all users. By seeing this rating, users decide whether to download this application.

Presently, there is no method to check game players' physical environment. This research work will help the Game development organizations to bifurcate their rating on the Google Play Store due to Game players' comfortable and uncomfortable zone. Types of internet connection like 3G, 4G, and WiFi can be noted down, which also affects the player's QoE. This research will also help to find the reason for the low rating. Hence the objective of user QoE in the different physical environments for the PUBG game has been achieved. Results show that players were highly satisfied with both variables while playing the PUBG game at Home over a WiFi internet connection.

References

1. Ballabio, M., Griffiths, M.D., Urbán, R., et al.: Do gaming motives mediate between psychiatric symptoms and problematic gaming? An empirical survey study. Addict. Res. Theory **25**, 397–408 (2017). https://doi.org/10.1080/16066359.2017.1305360
2. Ding, Y.: Research on operational model of PUBG. MATEC Web Conf. **173**, 03062 (2018). https://doi.org/10.1051/matecconf/201817303062
3. Mamun, M.A., Griffiths, M.D.: The psychosocial impact of extreme gaming on Indian pubg gamers: the case of PUBG (Player Unknown's Battlegrounds). Int. J. Ment. Heal. Addict. **19**, 2170–2174 (2019). https://doi.org/10.1007/s11469-019-00102-4
4. PUBG Mobile-Apps on Google Play. In: Google. https://play.google.com/store/apps/details?id=com.tencent.ig&hl=en&gl=US. Last accessed 13 Jan 2022
5. PUBG: Battlegrounds. In: Steam Charts. https://steamcharts.com/app/578080. Last accessed 13 Jan 2022
6. In: Tencent Mobile International Revenue & App Download Estimates from Sensor Tower-Apple App Store. https://app.sensortower.com/ios/publisher/publisher/504515421?user_app_id=61e1bef6180097014a509002&country=US&adaptive_colors=off&device=phone&start_date=2021-12-16&end_date=2022-01-14&tab_id. Last accessed 15 Jan 2022
7. Moll, P., Lux, M., Theuermann, S., Hellwagner, H.: A network traffic and player movement model to improve networking for competitive online games. In: 2018 16th Annual Workshop on Network and Systems Support for Games (NetGames) (2018). https://doi.org/10.1109/netgames.2018.8463390
8. Widodo, N.G., Balqiah, T.E.: The effect of addiction, playfulness and good price on purchase intention of in-app features. In: Proceedings of the International Conference on Business and Management Research (ICBMR 2020) (2020). https://doi.org/10.2991/aebmr.k.201222.002
9. Balakrishnan, J., Griffiths, M.D.: Loyalty towards online games, gaming addiction, and purchase intention towards online mobile in-game features. Comput. Hum. Behav. **87**, 238–246 (2018). https://doi.org/10.1016/j.chb.2018.06.002
10. Septian, M.R., Susilowati, M.H.D., Zulkarnain, F.: Spatial patterns of telecommunication signal quality, internet speed, and online game latency in East Jakarta. J. Geogr. Trop. Environ. **5**, 53–66 (2021). https://scholarhub.ui.ac.id/jglitrop/vol5/iss1/3

11. Lahby, M.: Enhancing modeling for network selection using graph theory in beyond 4G networks. Int. J. Bus. Data Commun. Netw. **16**, 48–69 (2020). https://doi.org/10.4018/ijbdcn.2020010104
12. Madiha, H., Lei, L., Laghari, A.A., Karim, S.: Quality of experience and quality of service of gaming services in fog computing. In: Proceedings of the 2020 4th International Conference on Management Engineering, Software Engineering and Service Sciences (2020). https://doi.org/10.1145/3380625.3380644
13. Laghari, A.A., He, H., Memon, K.A., et al.: Quality of experience (QoE) in cloud gaming models: a review. Multiagent Grid Syst. **15**, 289–304 (2019). https://doi.org/10.3233/mgs-190313
14. Laghari, A., He, H., Khan, A., et al.: Crowdsourcing platform for QoE evaluation for cloud multimedia services. Comput. Sci. Inf. Syst. **19**, 1305–1328 (2022). https://doi.org/10.2298/csis220322038l
15. Yang, K., Zhang, X., Zhao, Y., et al.: Looking into online gaming from measurement perspective. In: 2019 IEEE International Conference on Service-Oriented System Engineering (SOSE) (2019). https://doi.org/10.1109/sose.2019.00035
16. Laghari, A.A., Laghari, K.-R., Memon, K.A., et al.: Quality of experience (QoE) assessment of games on workstations and mobile. Entertain. Comput. **34**, 100362 (2020). https://doi.org/10.1016/j.entcom.2020.100362
17. Slivar, I., Suznjevic, M., Skorin-Kapov, L., Matijasevic, M.: Empirical QoE study of in-home streaming of online games. In: 2014 13th Annual Workshop on Network and Systems Support for Games (2014). https://doi.org/10.1109/netgames.2014.7010133
18. Mo, C., Zhu, G., Wang, Z., Zhu, W.: Understanding gaming experience in mobile multiplayer online battle arena games. In: Proceedings of the 28th ACM SIGMM Workshop on Network and Operating Systems Support for Digital Audio and Video (2018). https://doi.org/10.1145/3210445.3210450
19. Meng, Y., Zhang, S., Ye, Z., et al.: Causal analysis of the unsatisfying experience in realtime mobile multiplayer games in the wild. In: 2019 IEEE International Conference on Multimedia and Expo (ICME) (2019). https://doi.org/10.1109/icme.2019.00321
20. Chuang, T.-L.: A mixed-methods approach to explore users' behavior in playing mobile games. J. Comput. **32**, 164–173 (2021). https://doi.org/10.3966/199115992021023201014
21. Mkpojiogu, E.O.C., Sherriff, G., Hussain, A., Hashim, W.: Comparing the performance of Player Unknown's Battlegrounds (PUBG) mobile app on iPhone X and Samsung S9 Plus smartphones. Int. J. Adv. Sci. Technol. **29**, 320–325 (2020). https://repo.uum.edu.my/id/eprint/27156/
22. Al-Mansour, J.: The success behind the PUBG era: a case study perspective. Acad. Strat. Manag. J. (2019). https://www.abacademies.org/articles/the-success-behind-the-pubg-era-a-case-study-perspective-8825.html
23. Moniaga, J., Tansil, B.Y., Sasmoko, et al.: The influence of user experience Player Unknown's Battlegrounds game toward adaptive learning. In: 2019 International Conference of Artificial Intelligence and Information Technology (ICAIIT) (2019). https://doi.org/10.1109/icaiit.2019.8834669
24. Baloch, J.A., Jumani, A.K., Laghari, A.A., et al.: A preliminary study on quality of experience assessment of compressed audio file format. In: 2021 IEEE URUCON (2021). https://doi.org/10.1109/urucon53396.2021.9647114
25. Laghari, A.A., He, H., Shafiq, M., Khan, A.: Application of quality of experience in networked services: review, trend & perspectives. Syst. Pract. Action Res. **32**, 501–519 (2018). https://doi.org/10.1007/s11213-018-9471-x
26. Hu, J., Xi, T.: The relationship between game elements and player emotions by comparing game frameworks. In: HCI International 2019–Late Breaking Papers, pp. 320–329 (2019). https://doi.org/10.1007/978-3-030-30033-3_25
27. Khairunisa, A.A.: Computer-mediated communication: online gaming communication culture. In: Proceedings of the 2nd Jogjakarta Communication Conference (JCC 2020) (2020). https://doi.org/10.2991/assehr.k.200818.039

28. Shelstad, W.J., Chaparro, B.S., Keebler, J.R.: Assessing the user experience of video games: relationships between three scales. Proc. Hum. Factors Ergon. Soc. Annu. Meet. **63**, 1488–1492 (2019). https://doi.org/10.1177/1071181319631300
29. Setya Murti, H.A., Dicky Hastjarjo, T., Ferdiana, R.: Platform and genre identification for designing serious games. In: 2019 5th International Conference on Science and Technology (ICST) (2019). https://doi.org/10.1109/icst47872.2019.9166177
30. Vervoort, J.M.: New frontiers in futures games: leveraging game sector developments. Futures **105**, 174–186 (2019). https://doi.org/10.1016/j.futures.2018.10.005
31. Gałka, P., Strzelecki, A.: How randomness affects player ability to predict the chance to win at PlayerUnknown's Battlegrounds (PUBG). Comput. Games J. **10**, 1–18 (2021). https://doi.org/10.1007/s40869-020-00117-1
32. ITU- T Recommendation P.910: subjective video quality assessment methods for multimedia applications. International Telecommunication Union, Geneva, Switzerland (1996)

Chapter 31
An Attachment Processing Module for a Website Implemented Using Two Solutions

Guoliang Dong, Hao Li, Huimin Li, Yucheng Du, Shaobo Bai, and Guofang Wu

Abstract Regular business websites generally need to interact with users about attachments. Malicious attackers will also use attachments as an attack mode to launch information monitoring, scanning, or attacks on the website server. According to the different security management requirements of the website and the actual management needs of the business system, different development solutions can be adopted to realize the attachment processing function of the website. In this paper, two solutions are used to realize the attachment processing function of an active website. Both solutions can meet user business requirements with different security requirements and development costs.

31.1 Introduction

Since the birth of the Internet, cyberattack technology based on network technology has also changed with the progress of network technology. At present, the basic protection technology of some websites can block most malicious attacks. However, malicious actors are constantly looking for new ways to break through these security barriers in hopes of gaining maximum control of the website [1–3].

For regular portals, information is generally provided to users in one direction. In addition to the basic information interaction between the regular business system and the user, some business systems also require the user to submit or provide additional attachment materials to the user. These attachments are generally in JPG or PDF format. Malicious attackers will embed malicious code in JPG or PDF files, listen to the website's server information when the attachment is uploaded to the website, or steal information. Therefore, when working on a website with attachments, you should consider security in advance.

G. Dong (✉) · H. Li · H. Li · Y. Du · S. Bai · G. Wu
Research Institute of Highway Ministry of Transport, Beijing 100088, China
e-mail: gl.dong@rioh.cn

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_31

A business website system in use, due to the expansion of business functions, needs to add a new business module with an attachment processing function. This paper discusses two different solutions to implement the attachment handling function of websites from the aspects of developers' technical requirements, security strength requirements, and module extensibility.

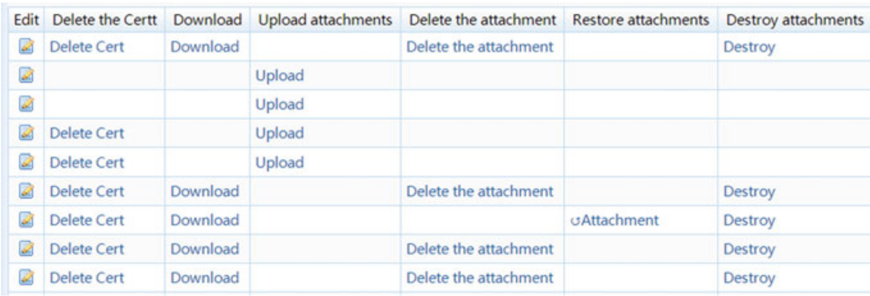
31.2 Implementation Based on Rapid Development Platform

The rapid development platform has been developed rapidly due to its advantages of low code or no code, low technical threshold, and high development efficiency [4–6]. The general rapid development platform provides common modules for regular websites, such as login management, business query, user management, system management, etc. These rapid development platforms provide great convenience to developers, greatly reduce the work intensity of developers, save development time, and provide work efficiency [7–9]. In addition, the rapid development platform model is solidified and the security of the project completely depends on the development platform, which is determined by its own structure.

An authorized rapid development platform was used to realize the attachment upload function of the business website. The trial version of the rapid development platform has certain restrictions on the number of end customers and the trial period and is generally used for user trials, experiences, and demonstrations. The licensed version requires a fee and can be used commercially for a long time without a limit on the number of customers. This rapid development platform is an enterprise-level development tool that adopts the mainstream MVC (Model–View–-Controller) design pattern. After years of product iteration, it has become a relatively mature development platform. The rapid development platform has the following features:

- Cross-platform;
- Online development, maintenance, and release of business modules;
- More than 80% of modules do not use code;
- Low technical requirements for developers;
- The built-in MVC control engine implements most of the business logic;
- The built-in WEB report engine can realize business functions such as queries, statistics, and reports;
- Providing external application interfaces can implement more complex business logic;
- Use a custom form engine to customize special pages.

Figures 31.1 and 31.2 show the user interface and development interface of accessory processing implemented using the rapid development platform.



Edit	Delete the Certt	Download	Upload attachments	Delete the attachment	Restore attachments	Destroy attachments
	Delete Cert	Download		Delete the attachment		Destroy
			Upload			
			Upload			
	Delete Cert		Upload			
	Delete Cert		Upload			
	Delete Cert	Download		Delete the attachment		Destroy
	Delete Cert	Download			↺Attachment	Destroy
	Delete Cert	Download		Delete the attachment		Destroy
	Delete Cert	Download		Delete the attachment		Destroy

Fig. 31.1 User interface for attachment processing implemented by a rapid development platform

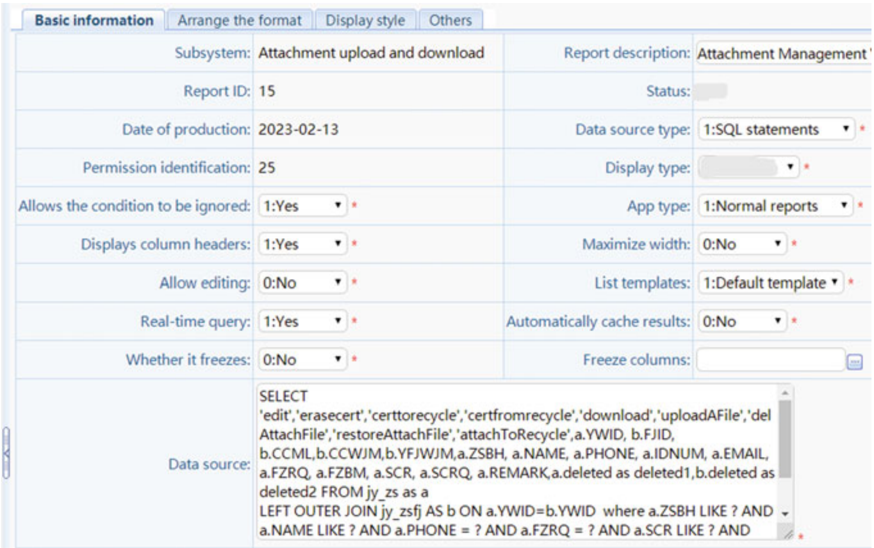


Fig. 31.2 A development interface for attachment processing implemented by a rapid development platform

31.2.1 Attachment Upload

General data is typically saved as fields in tables in the back-end database. In order to improve the operational efficiency of the database and ensure the security of the database, the attachment file itself is not directly saved in the database, but stored in the form of a similar link. Attachment files are stored in a separate folder, and information such as attachment file name, type, storage location, and size are stored in the database.

In the rapid development platform, attachment upload is mainly implemented using form resources and MVC actions provided by the platform. Form resources

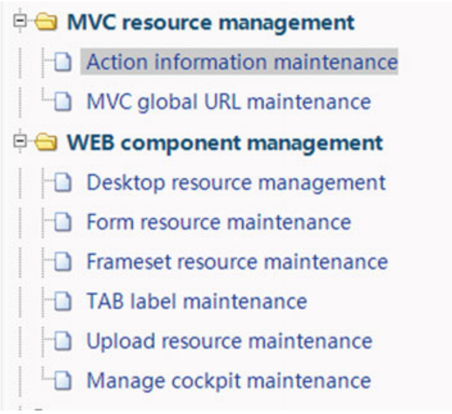


Fig. 31.3 Develop interface for attachment upload (form resources and actions)

mainly provide a place for data to display resources. MVC Action uses the pre-packaged commands of the rapid development platform to implement the attachment upload function. Figure 31.3 shows the development interface for uploading attachments.

To implement the attachment upload function, the Form Type should be set to Upload Form when designing the form. Fill in the MVC resources and parameters passed in the form to execute the desired MVC Action. Figure 31.4 shows the form resources of the attachment upload module.

Form resources provide a container for manipulating attachments, and attachment upload is done by MVC Action. Figure 31.5 shows the MVC resource designer that implements attachment upload and implements the basic operation of MVC action.

Call the built-in script in the Business Action Resource in MVC Resource Designer to upload attachments. Figure 31.6 shows the development interface of "Business Operation Resources". Fill in the "Script Content" with the built-in operation statements of the rapid development platform, and fill in the parameters corresponding to the script operation statements in the "Parameter Expression Set". The script and parameters cooperate to realize the function of attachment upload.

Follow the above steps to complete the module development of attachment upload on the rapid development platform.

31.2.2 Attachment Download

Create intelligent reports with the rapid development platform and generate a management interface for downloading and uploading attachments. Figure 31.7

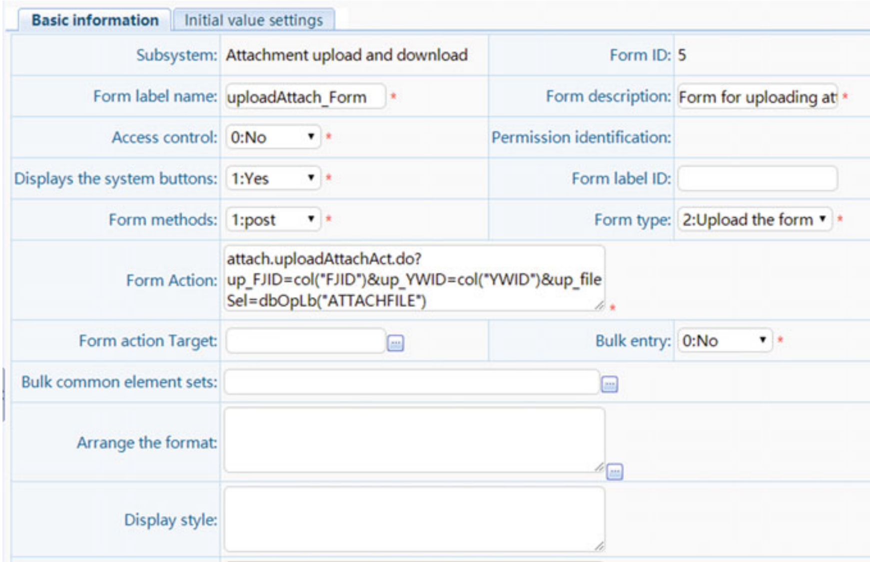


Fig. 31.4 Form resource for the attachment upload module

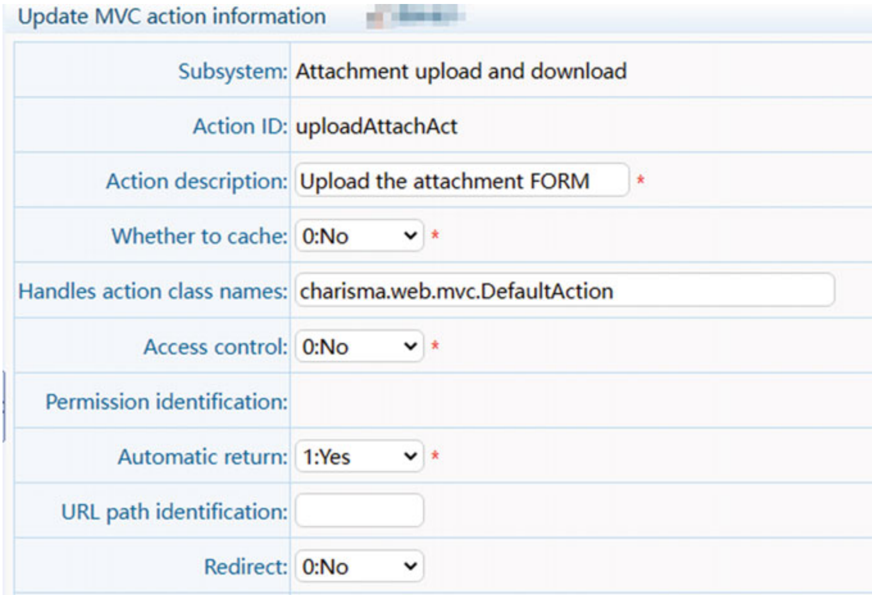


Fig. 31.5 MVC Resource Designer for attachment upload

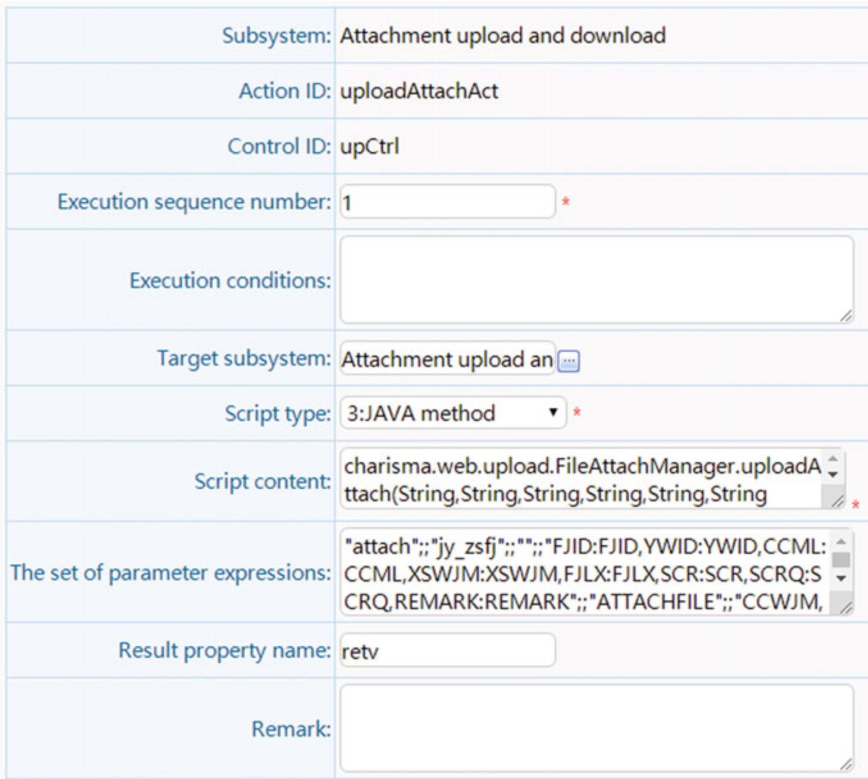


Fig. 31.6 The development interface of Business Action Resources

shows the development interface of the shortcut button associated with the attachment download, which is the "Normal display" display mode and the displayed text is "Download".

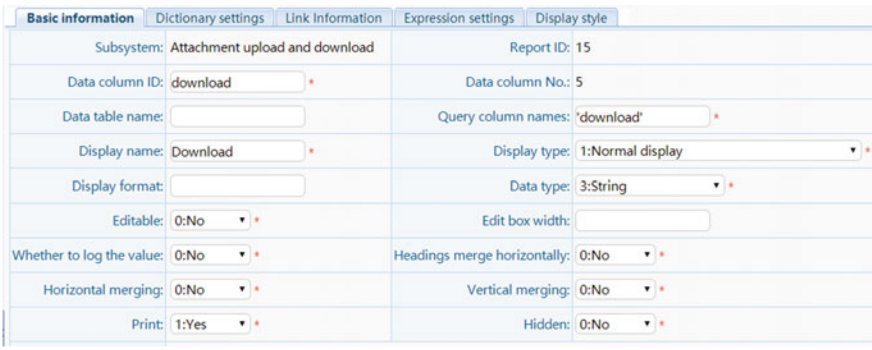


Fig. 31.7 The development interface for shortcut buttons associated with attachment downloads

In the “Link Information” field, fill in the “Linking URL” and parameter columns. Figure 31.8 shows the design interface of “Link Information”, which calls MVC resources to download attachments.

Figure 31.9 shows the development interface for downloading attachments. MVC resources call “system.fileAttachDownload.do” in the URL path to implement the attachment download function.

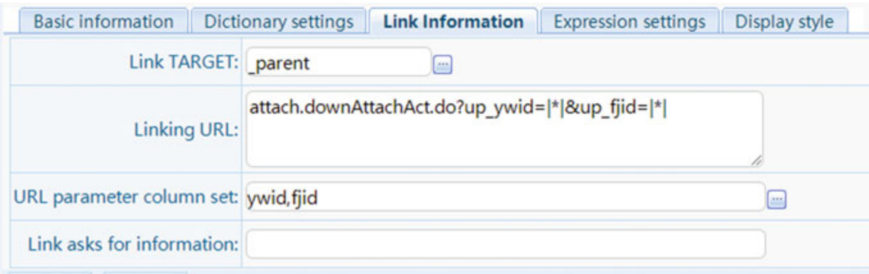


Fig. 31.8 Link Information

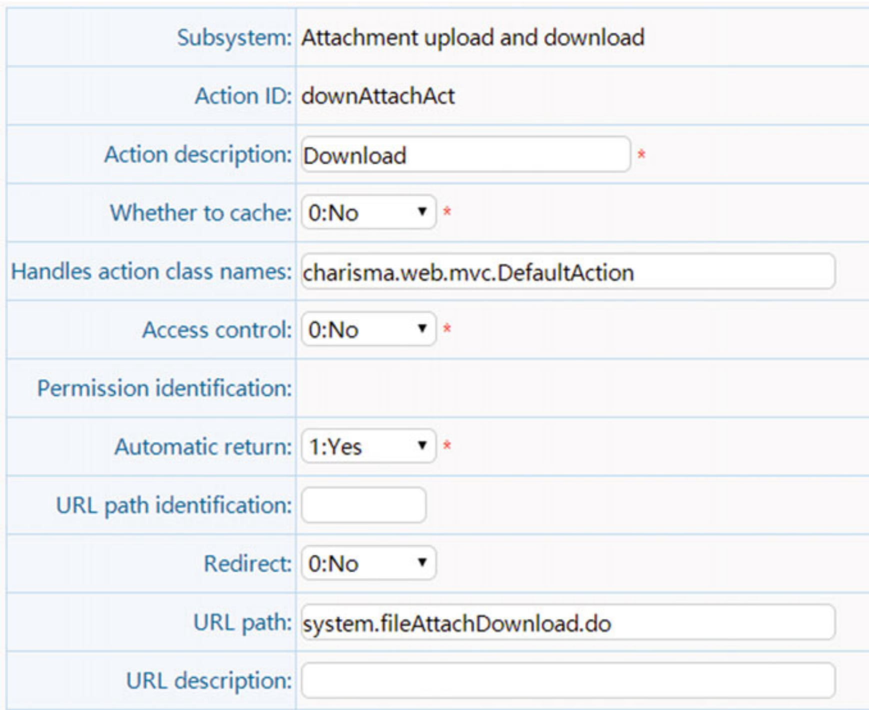


Fig. 31.9 Development interface for attachment download

Follow the above steps to complete the module development of the attachment download on the rapid development platform.

31.3 Java-Based Implementation

Using Java as the development language and using MVC to develop the design pattern is a more mainstream website development mode [10–12]. Developers should have experience in website development and need to be familiar with Java development language, MVC design patterns, and website development debugging environment. Developers use Java as the development language, freely control the security strength of the module, and perform customized business processing and security verification.

Figure 31.10 shows the implementation of the MVC design pattern in project development.

The control logic of each layer of project development and the relationship with MVC are as follows.

- The user uses a browser to access the JSP page, and the JSP page corresponds to the View layer.
- All operations on the JSP page will be sent to the Controller layer for processing. The Controller layer is implemented by servlets in project development.
- In project development, the service layer and the Dao layer work together to implement the functions of the model layer. The service layer receives the parameters passed by the control layer and calls the Dao layer to access the database.
- The Dao layer encapsulates the data read from the database and returns it to the Service layer, Controller layer, and View layer step by step.

Figure 31.11 shows the tree structure of each layer in the project development.

In the treemap, “dao”, “model”, and “util” correspond to the Model layer. “web” corresponds to the Controller layer, which is the core part of implementing attachment upload and attachment download. The View layer is not shown here.

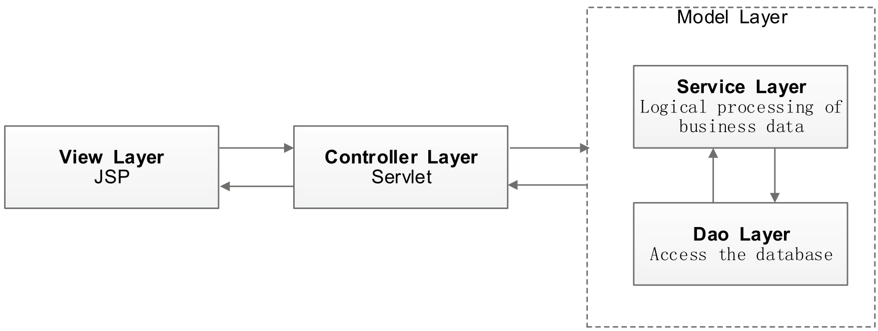


Fig. 31.10 Implementation of MVC design patterns in project development

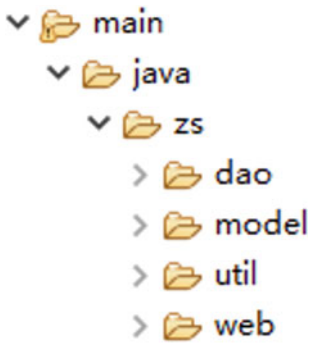


Fig. 31.11 Treemap in project development

31.3.1 Attachment Download

Attachment downloads are mainly implemented in the Controller layer, and in the project by inheriting the HttpServlet class. The following is part of the implementation code for the HttpServlet class that implements the attachment download.

```
ServletContext servletContext = getServletContext();
String mimeType = servletContext.getMimeType("/upload/"+
downloadFileName);
resp.setContentType(mimeType);
String
str="attachment;fileName="+URLEncoder.encode("cert."+ext,
"UTF-8");
resp.setHeader("Content-Disposition",str);
Files.copy(Paths.get(realPath,downloadFileName),resp.getO
utputStream());
```

31.3.2 Attachment Upload

Similarly, attachment upload is implemented in the project by inheriting the HttpServlet class. The following is part of the implementation code for the HttpServlet class to upload attachments.

```
String realName =upl.getSubmittedFileName();
String fileExt = real-
Name.substring(realName.lastIndexOf("."));
String newName = UUID.randomUUID().toString()+fileExt;
String fileRealPath =
req.getServletContext().getRealPath("/hereupload") +"/"+
newName;
upl.write(fileRealPath);
```

The following shows the main code to implement the View layer.

```
<form action="uploadFile" method="post"
enctype="multipart/form-data">
<input type="text" name="fileDesc"/>
<input type="fileName" name="fileName"/>
<input type="submit" value="upload"/>
</form>
```

31.4 Conclusion

This paper describes two solutions for implementing an attachment management module for a website. Both solutions achieve the intended purpose, and the characteristics of both solutions are summarized below.

Features of the rapid development platform solution:

- Low-code development, basically no code written in module implementation;
- Low technical requirements for developers and a certain understanding of the database can basically complete this project;
- Before development, developers need to learn the development mode of the rapid development platform and understand and master the syntax rules, functions, and expressions built into the rapid development platform.
- Under the condition of familiarity with the development environment, the development speed is fast;
- Low commissioning effort;
- The syntax rules built into the rapid development platform are only applicable to this platform and are not universal;
- The trial version of the rapid development platform has a trial time limit, which limits the number of concurrent requests from end users, while the authorized version does not have the above restrictions;

- The security of business systems depends entirely on rapid development platforms, which are suitable for conventional business systems without special requirements.

Features of the self-written code solution:

- High technical requirements for developers, basic knowledge of website development, familiarity with HTML, JAVA, CSS, database, and other website development tools;
- Developers need to have the ability to develop, debug and deploy, and code debugging will take up more time;
- The security strength of business logic and code can be freely controlled, which is suitable for business systems with high-security requirements.

From the above comparison, it can be seen that the rapid development platform solution can be used for business systems with general security requirements, enterprise-level development environments, and subsequent business expansion. If a business system with high security and complex business logic needs to be developed, a solution that writes code should be used.

Acknowledgements This paper is funded by the Science and Technology Innovation Special Fund Project (QG2021-3-15-2, 2021-C414).

References

1. Xie, B., Fu, C.: Inventory of the characteristics and differences of common network attack types. Netw. Secur. Informatiz. **1**, 124–127 (2022)
2. Zheng, Z., Cheng, Y., Yao, J.: Website security cloud protection platform practice. Inf. Technol. Stand. **000**(005), 235–239 (2022)
3. Jundong, H.: Website security risk and solution analysis. Inn. Mong. Radio Telev. Technol. **037**(011), 69–73 (2020)
4. Wang, M.: Design and Implementation of Enterprise Software Rapid Development Platform, pp. 50–51. China Informatization (2021)
5. Tan, S.: Design and Implementation of Java Web Management Information System Rapid Development Platform. University of Electronic Science and Technology of China (2018)
6. Guanqun, W., Haibo, T.: Design and research of lightweight shared rapid development platform based on cloud technology. J. Chengde Pet. Coll. **23**(2), 38–41 (2021)
7. Song, M., Xianli, D., Li, J., Li, P., Xiao, Y.: The practice of low-code rapid development platform in the digital transformation of Baiyun airport. Electron. Technol. Softw. Eng. **4**(21), 86–89 (2022)
8. Niu, Y., Qu, C., Guan, F., Yuan, Y., Gao, H., Zhang, L.: Rapid development platform for information system based on excel server. Comput. Program. Ski. Maint. **2**(11), 63–64 (2015)
9. Zuo, G., Xianghong, W., Huaidong, Z., Min, W.: Design of special project management system for geology and mineral survey and evaluation based on workflow and rapid development platform technology. Geol. Chin. Bull. **35**(6), 1025–1033 (2016)
10. Wang, G.: Web-based substation monitoring system design. Microcomput. Appl. **38**(10), 206–208 (2022)

11. Yanchi, J.: Design research of science and technology management information system based on MVC model. Electron. Technol. Softw. Eng. **016**, 192–195 (2022)
12. Hu, S.: Design of enterprise field management system based on MVC framework. Software **043**(007), 19–23 (2022)

Chapter 32
Research and Application of an Integrated Budget Management System for Provincial Financial Management

Ting Cao

Abstract Financial management is an important link in government management work, and the exercise of functions and powers by government agencies at all levels both domestically and externally relies on financial management work. If financial management is chaotic or fails to reflect the demands of government management, it will have a huge negative impact on the operation and function of the government, and even result in the lack of functionality of government institutions. Budget management is the primary link of financial management, which determines the success or failure of government financial expenditure and its guiding role. After decades of continuous practice and exploration, the methods and practical conditions of budget management are gradually developing towards standardization and information. Budget management needs to consider numerous factors in social management. If not fully considered, it will seriously affect and mislead the accuracy of the budget. This article proposes an integrated budget management method for provincial-level financial management and studies the implementation method of the system, which is applied and verified in financial management practice.

32.1 Introduction

Good financial management is a necessary condition for the stable operation of society, and with the continuous deepening and improvement of social development, financial management is also continuously optimized to adapt to new social development trends. Since the implementation of budget reforms in the financial departments of governments around the world, after more than ten years of continuous exploration, practice, and improvement, it has played a positive role in deepening

T. Cao (✉)
Financial Information Management Center of Jiangsu Provincial Department of Finance Nanjing, Jiangsu 210024, China
e-mail: caotiiing@163.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_32

budget management concepts, refining departmental budget preparation, strengthening budget constraint rigidity, and ensuring the performance of institutional responsibilities [1]. However, with the continuous development of the economy and society in various countries around the world and the deepening of the construction of public finance systems, especially in the context of the current requirement to disclose budget, final accounts, and progress data of public funding implementation, the current situation of fiscal budget management is due to insufficient accuracy in predicting at the beginning of the year. There are still issues such as imbalanced budget execution progress, cross-use of project funds and public funds, large year-end balance, unreasonable spending by individual government direct management units, and lack of standardized reference basis and system for evaluating work performance [2]. These issues to some extent hinder the scientific rationality of government funding arrangements and the further improvement of social resource allocation efficiency [3]. In order to further enhance the seriousness and authority of the financial budget and enhance the scientific rationality of the budget allocation at the beginning of the year, it is necessary to prepare and report the beginning of the year budget based on a complete and accurate basic information database and a cross-year rolling financial project database.

The overall process of budgeting includes four steps. Firstly, the financial expenditure institution proposes and prepares the annual fund budget and detailed fund usage reports of the institution or unit, and outputs all the above information to the government financial management unit. At the same time, relevant financial budget reports that meet the funding needs of each unit must be submitted based on the various tasks assigned by the government [4]. After receiving the budget requirements from the government's direct management unit, the financial management unit shall verify the budget fund items item by item in accordance with the white paper on financial fund plans issued by the government and the legal provisions on the use of financial funds, to verify the necessity and legality of fund use. The final municipal government will organize a review meeting to discuss and evaluate the completed financial budget documents and provide the evaluation results. The integrated budget management system studied in this article digitizes the above processes and develops corresponding intelligent management systems, providing efficient digital management tools for government financial budget management [5]. The difficulty in the research process of this article lies in integrating various data distributed in different financial systems, designing and developing a reasonable fusion model that can not only support the normal development of financial management work but also make the system operation and maintenance simple.

32.2 Related Work

The focus of financial system development and reform in countries around the world is on fiscal revenue management, while the reform of fiscal expenditure management is relatively lagging behind. The budget management system basically follows the old practices of the past. The budget classification is not standardized, with cross-classification based on economic nature and government functions, and the budget subjects fail to accurately and comprehensively reflect the overall situation of government power and revenue and expenditure. Budgeting is relatively rough, with revenue budgeted by category and expenditure budgeted by function, resulting in low transparency in budget allocation and use. The budget approval is not timely, and there are occasional cases of additional or reduced budget execution, resulting in weakened budget constraints [6]. After 2000, the financial management of major European countries was refined as much as possible on the basis of existing revenue and expenditure classifications, and on the basis of cooperation and trial budgeting among various government departments, four departments, namely, the Ministry of Education, the Ministry of Agriculture, the Ministry of Science and Technology, and the Ministry of Labor and Social Security, were selected as pilot units to submit departmental budgets to government units at all levels, taking the first step in the reform of the budget system [7]. Compiling government budgets is a common practice in market economy countries and a direction for budget reform in European countries [8]. The focus of budget reform in the future is to promote the departmental budget system as a starting point, while also promoting reforms such as the zero-based budget system, comprehensive budget system, national treasury system, and government procurement system. In order for these reforms to be proactive, stable, and effective, all government departments should promptly approve unit budgets. When applying for the use of financial funds, each unit must follow the budget procedures. It is necessary to ensure that there is a budget first, followed by expenditure, and strictly implement the budget expenditure management measures. To effectively enhance the binding force of the budget and maintain the standardization of the government's financial budget.

In 2023, the Ministry of Finance of China issued a standardized document on the integration of budget management, accelerating the expansion of the scope of integrated budget management business. The document proposes to integrate government debt management, asset management, performance management, and other businesses, covering all aspects of the budget management process [9]. Financial departments at all levels should accelerate the expansion of the business scope covered by the integrated budget management system on the basis of its comprehensive operation, and achieve integrated management of the entire budget management process as soon as possible. Effectively implement unified business standards into the integrated system. The budget integration system requires further refinement and improvement of the main workflow, basic control rules, and core management elements of government budget management at all levels [10].

32.3 Research on Centralized Analysis Model and System of Provincial Financial Management Budget

The implementation of the provincial financial management budget requires unified management of financial management elements, basic data, users, permissions, and other business data. At the same time, it is also necessary to integrate and embed various processes, control rules, logical database tables, database table elements, etc., of financial management into the integrated budget management system. Continuing the integrated technical architecture, fully utilizing more cutting-edge technologies such as cloud computing and big data to reconstruct the system. By using microservice invocation design to solve the problems of low performance, poor stability, and high difficulty in development and maintenance, we can achieve a "centralized" deployment of provincial-level applications in financial business. Using the achievements of integrated budget management construction, comprehensively review the current business status and problem gaps of the existing system to be integrated. Through comparison and analysis with the existing integrated budget management system, propose different integration methods for different software systems. In the process of integrating relevant systems, it is necessary to effectively ensure the integration of budget management before and after integration, as well as the safe and stable operation of the system to be integrated.

32.3.1 Centralized Analysis Model of Provincial Financial Management Budget

The budget for the beginning of the year in the management of budget review includes the basic expenditure budget and the project expenditure budget. The basic expenditure budget is used to ensure the normal operation of the unit. The basic expenditure budget is generally calculated based on the unit's assets such as personnel and vehicles according to quota standards, mainly including personnel salaries and office expenses, information on assets such as personnel and vehicles, and quota standards from the basic information database. Therefore, the project expenditure budget in this model is based on the performance requirements of the unit and the financial resources that can be met. The project declared by the unit is matched with funds according to the priority, and the final expenditure budget report is output. According to regulatory expenditure requirements, the preparation and submission of the beginning of the year budget must go through a two-up and two-down review process, based on a complete and accurate basic information database and a cross-year rolling project database.

As shown in Fig. 32.1, firstly, based on the work focus and business model of the budget preparation and review at the beginning of the year, the following business models are abstracted and described in detail. The budget review management product at the beginning of the year is closely related to external systems. The data

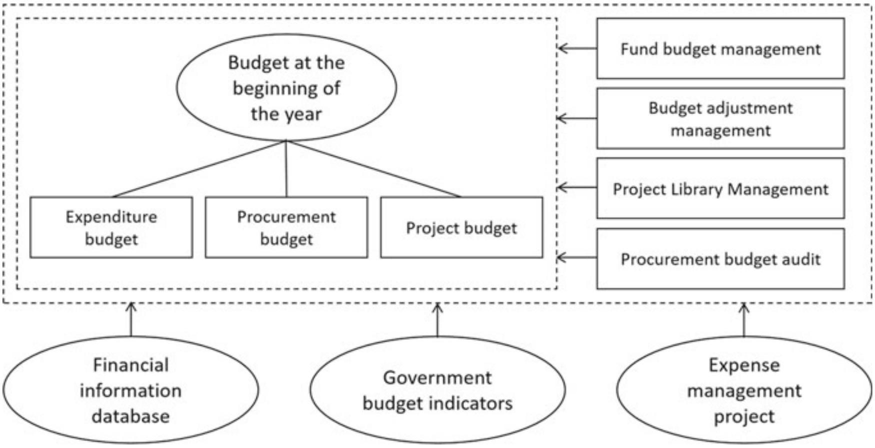


Fig. 32.1 Centralized analysis model of provincial financial management budget

that the budget review interacts with is mainly result data. Overall, the result data of other products/systems is usually input from the budget review management product at the beginning of the year, while the result data of the budget review management product at the beginning of the year is input from other products/systems. The budget review at the beginning of the year is initiated in June and July each year to prepare the expenditure budget for the next year. The core of the budget review management at the beginning of the year is the expenditure budget used for the new year's funding arrangement, which is divided into basic expenditure budget and project expenditure budget. The basic expenditure budget is generated based on the time points of assets such as personnel and vehicles in the basic information database, and calculated according to quota standards. The project expenditure budget is extracted from the execution projects and pending projects in the project library, and the pending budget is matched with the funding budget through the budget preparation and review process at the beginning of the year. The total amount of basic expenses and project expenditure budget at the beginning of the year should be within the financial permission range.

32.3.2 Design of Provincial Financial Management Budget Integration System

The system first needs to obtain basic information about financial expenditure units, which can be used as the basis for expenditure budget preparation and review. Users are required to input unit code name, function, internal organization, subordinate units, and annual work plan, as well as information such as unit leader, financial leader, phone number, postal code, address, etc. In addition to the annual work plan,

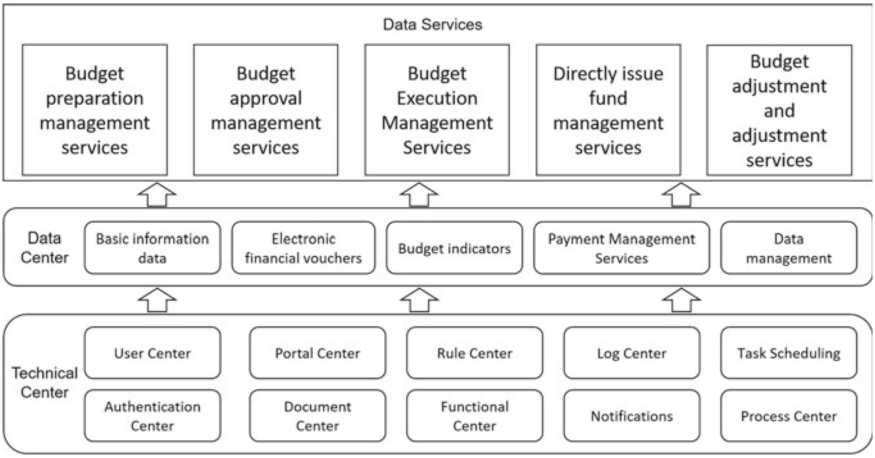


Fig. 32.2 Design of provincial financial management budget integration system

other information can be extracted from the basic information database or directly entered in this table. Then, data can be extracted from the basic information database, edited, imported from text files, or exported. The final output and submission of the basic unit information table to the competent department for review. The system also supports viewing the basic information of the reviewing unit as the basis for expenditure budget review.

As shown in Fig. 32.2, the system functions are divided into three layers, namely, the technical platform layer, the data platform layer, and the data service layer. The Ministry of Finance uniformly issues transfer payment indicators and reports task parameters through the data center of the summary system. The budget adjustment and adjustment services module of various financial departments in the integrated system receives transfer payment indicators and report task parameters. The reported data is generated by the integrated system of subordinate units, and financial departments in various regions submit reconciliation results, reports, and reports through the summary system.

32.4 Application and Improvement of Provincial Financial Management Budget Integration System

Firstly, deploy and implement the system in the pilot zone, and support its trial operation. Collect various suggestions and issues during the trial operation process, adopt and summarize them in a timely manner, and expand the promotion scope based on further improving the system functions. This plan is extended under the centralized management mode of provincial finance, and its common needs have been taken into consideration. In principle, there will be no new demand development, but unified management of the demand should be done. For requirements with strong commonality, a plan can be formulated and resources can be arranged for implementation after being agreed upon by the demand decision-making group. To ensure smooth promotion, short-term personalized requirements will not be considered temporarily.

32.4.1 Implementation of Provincial Financial Management Budget Integration System

The method of preparation and review by the financial expenditure unit, unified linkage of indicators by the finance, generation of payments, and distribution from the zero balance account of the finance to the unified salary distribution account. Firstly, through workflow settings, the process, permissions, and corresponding interfaces can be set according to the workflow. When enabling financial expenditure units or batches, the unified distribution mode can be selected. Secondly, for the overall configuration, verify the process, data, accounts, and generation methods based on business requirements. The method of preparing, reviewing, linking indicators, generating payments by the unit, and distributing them from the unit's zero balance account or designated account to the unified salary distribution account. Firstly, through the workflow settings in the system, the process, permissions, and corresponding interfaces can be set according to their processes. When the unit or batch is enabled, it can be selected in spontaneous mode. Secondly, for the overall configuration, verify the process, data, accounts, and generation methods based on business requirements. Regardless of whether it is centralized or spontaneous, after completing the corresponding configuration, verify according to business requirements. If there are special needs in the business, negotiate and handle them uniformly. Some divisions not only support unified or spontaneous mode settings at the unit level but also require different batches to support unified or spontaneous mode settings.

As shown in Fig. 32.3, the system has implemented three functional modules, namely, budget management module, payment management module, and report management module. The personnel and salary information in the preparation process can be maintained in the system, imported from external Excel methods, or accessed according to the provided standard interface. The audit process can be set up according to the needs of the division, with centralized or spontaneous processes. The synchronous or asynchronous review of human resources and salary can be achieved by configuring parameters. The distribution process unifies the method of sending bank pending details. In order to make the transmission process of salary data more secure, convenient, and efficient, it is necessary to develop and implement a green channel for data transmission at the data transmission bureau. Firstly, use configuration files to automate the configuration of data interfaces, and then add data labels such as personnel nature and salary categories to achieve data maintenance and matching. The acquisition of personnel-related data for salary distribution personnel relies on multiple data systems, for example, obtaining the information of the current

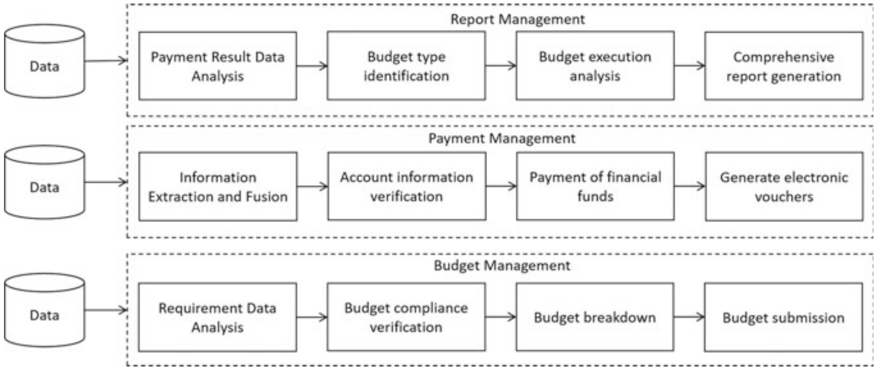


Fig. 32.3 Implementation of provincial financial management budget integration system

salary distribution personnel from the editor, and obtaining the salary data of the salary distribution personnel from external sources; obtaining salary information for this payment from the Human Resources and Social Security and Organization Department; obtaining plan information, generating payment information, and withdrawing the generated payment information; obtaining amount information during quota linking, generating payment information for wages, and withdrawing payment information generated; and pushing the bank's pending details. At the same time as generating payment information, synchronize the generation of bank payment details.

32.4.2 Application and Improvement of Provincial Financial Management Budget Integration System

Due to the abundance of microservices on integrated platforms, the logical relationship of content is implemented through separate statistics. Both integration and electronic payment adopt a centralized deployment method, which no longer reflects a hierarchical structure on the server side. According to the current situation, some integrated modules are not required to be launched on New Year's Day. Instead, an interface method is used to directly read payment information from the original system. Develop a unified and standardized salary data access interface which can support its configurability. The salary distribution module obtains salary data from external sources. For example, obtaining salary information from the Organization Department and Human Resources and Social Security, regularly synchronizing the salary distribution module, and maintaining salary information. External access to information is not allowed to be modified. If errors are found in the obtained data, the organization department and the human resources and social security system can be notified to adjust the data. After the adjustment, salary data can be synchronously obtained. According to previous research, there are also some districts that have implemented multi-batch management. Although the existing system supports multi-batch management, different distribution modes and payment methods need to be supported for different batches mentioned in some regions. It is necessary to further strengthen the development of multi-batch expansion requirements to meet user business processing needs.

As shown in Fig. 32.4, the integrated budget management system adopts a bypass deployment method. Provincial financial management units review the financial budget of financial expenditure units through this system, and the electronic vouchers generated in the process are managed by the electronic voucher management center. Some districts hope to automatically generate a payment plan application while matching the amount linkage, and the amount matching business will move forward. At the end of budget preparation and submission stage, the amount matching will be carried out to avoid the problem of insufficient amount and tight time for adding the amount in the final stage of payment generation. Secondly, for the withholding and payment part, it is also necessary to generate an authorized payment plan limit, which is convenient for the unit to handle the withholding and payment business.

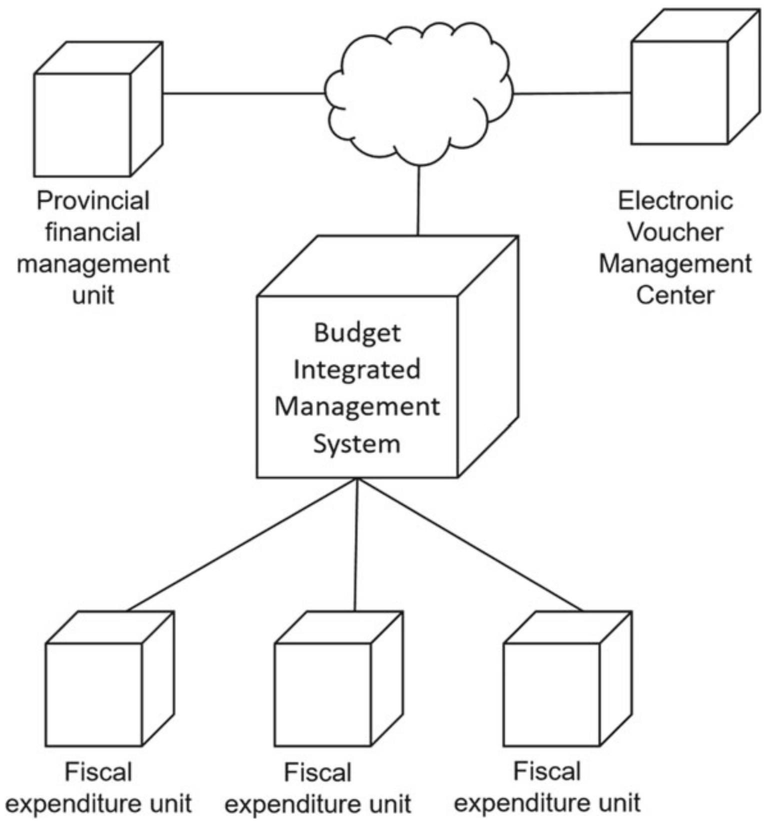


Fig. 32.4 Deployment of provincial financial management budget integration system

32.5 Application Results

The integration of budget management is an inevitable requirement for accelerating the establishment of a sound modern budget system and promoting the modernization of the national governance system and governance capabilities. Promoting the modernization of financial management and governance capabilities through information systems helps to enhance government governance capabilities. Therefore, budget integration in financial management is the main supporting means for deepening the reform of the budget system. At present, the reform of the budget system has entered a critical period, and relying solely on the introduction of individual systems and rules is no longer able to meet the needs of rapid social development today. To deepen the reform of the budget system, it is necessary to closely focus on the requirements of scientific standards, transparent norms, and strong constraints, and systematically formulate practical and feasible plans. Based on practical experience in financial management, establish scientific expenditure standards and comprehensive budget management norms. It is necessary to achieve effective linkage and control between budgets of various levels of government, budget management processes, government budgets, departmental budgets, and unit budgets. It is necessary to strengthen the overall management of various funds and assets of departments and units. Budget integration is a common challenge and demand faced by governments around the world in carrying out modernization and digital reforms. The budget integration system proposed in this article is an exploration of deepening the reform of the budget system, providing ideas and plan references for countries around the world to implement and improve the financial budget integration management system.

References

1. Chen, J.: Integrated information construction of financial management with budget as the core. Charming China (2019). (In Chinese)
2. Huang, A.: Research on university project performance management under the integration of budget management. Economics **5**(3), 139–141 (2022). (In Chinese)
3. Bao, G., Zhang, R.: From “Budget performance” to “Performance budgeting”-review and prospect of budget performance evaluation in China. J. Lanzhou Univ. Soc. Sci. Ed. **47**(5), 9 (2019). (In Chinese)
4. Wang, G., Zhang, X.: The construction path of university budget management system based on budget integration. Budg. Manag. Account. **5**(4), (2021). (In Chinese)
5. Liu, Y.: Promote efficient financial management of administrative institutions based on the integrated budget financial management system. Chin. Sci. Technol. J. Database (Full Text Ed.) Econ. Manag. **11**(4), (2022). (In Chinese)
6. Wang, L., Wang, C., Zhu, Hi., et al.: Research on the strategy of integrated management of financial budgeting in universities under the background of integration of industry and finance. China’s Collect. Econ. **3**(3), (2023). (In Chinese)
7. Bulgurcu, B.: Application of TOPSIS technique for financial performance evaluation of technology firms in Istanbul stock exchange market. Procedia-Soc. Behav. Sci. **62**, 1033–1040 (2017)

8. Moslem, A., Abbas, B.: A comparative analysis of dynamic and cross-sectional approaches for financial performance analysis. Am. J. Financ. Account. **5**(3), (2018)
9. Wu, P.: State owned asset management in public institutions under the background of budget integration. Chin. Sci. Technol. J. Database (Full Text Ed.) Econ. Manag. **12**(3), (2022). (In Chinese)
10. Song, M.: The impact of budget integration on the financial work of administrative institutions. Wealth Manag. Collect. **4**, 3 (2023). (In Chinese)

Author Index

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2