Journal of Electronic Research and Application

Editor-in-Chief

Joselito A. Dolot

Lyceum of the Philippines University-Batangas, Philippines

BIO-BYWORD SCIENTIFIC PUBLISHING PTY LTD
(619 649 400)
Level 10
50 Clarence Street
SYDNEY NSW 2000
Copyright © 2025. Bio-Byword Scientific Publishing Pty Ltd.
Complimentary Copy

ISSN (ONLINE): 2208-3510
ISSN (PRINT): 2208-3502

Journal of Electronic Research and Application

Focus and Scope

Journal of Electronic Research and Application is an international, peer-reviewed and open access journal which publishes original articles, reviews, short communications, case studies and letters in the field of electronic research and application.

Topics covered but not limited to:

- Automation
- Circuit Analysis and Application
- Electric and Electronic Measurement Systems
- Electrical Engineering
- Electronic Materials
- Electronics and Communications Engineering
- Power Systems and Power Electronics
- Signal Processing
- Telecommunications Engineering
- Wireless and Mobile Communication

About Publisher

Bio-Byword Scientific Publishing is a fast-growing, peer-reviewed and open access journal publisher, which is located in Sydney, Australia. As a dependable and credible corporation, it promotes and serves a broad range of subject areas for the benefit of humanity. By informing and educating a global community of scholars, practitioners, researchers and students, it ndeavors to be the world's leading independent academic and professional publisher. To realize it, it keeps creative and innovative to meet the range of the authors' needs and publish the best of their work.

By cooperating with University of Sydney, University of New South Wales and other world-famous universities, Bio-Byword Scientific Publishing has established a huge publishing system based on hundreds of academic programs, and with a variety of journals in the subjects of medicine, construction, education and electronics.

Publisher Headquarter

BIO-BYWORD SCIENTIFIC PUBLISHING PTYLTD
Level 10
50 Clarence Street
Sydney NSW 2000
Website: www.bbwpublisher.com
Email: info@bbwpublisher.com

2025, Volume 9, Issue 5

Table of Contents

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Optimization of Laminating Angles for Skirt Panels of EMUs Front Using Composite Materials Based on the Cheetah Optimizer

Yuqing Ma, Chunge Nie, Siqun Ma

Collage of Zhan Tianyou, Dalian Jiaotong University, Dalian 116028, Liaoning, China

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: With the development of composite materials, their lightweight and high-strength characteristics have caused more widespread use from aerospace applications to automotive and rail transportation sectors, significantly reducing the energy consumption during the operation of EMUs (Electric Multiple Units). This study aims to explore the application of composite materials in the lightweight design of EMU front skirts and proposes a design method based on three-dimensional Hashin failure criteria and the Cheetah Optimizer (CO) to achieve maximum lightweight efficiency. The UMAT subroutine was developed based on the three-dimensional Hashin failure criteria to calculate failure parameters, which were used as design parameters in the CO. The model calculations and result extraction were implemented in MATLAB, and the Cheetah Optimizer iteratively determined the optimal laminating angle design that minimized the overall failure factor. After 100 iterations, ensuring structural integrity, the optimized design reduced the weight of the skirt panel by 60% compared to the original aluminum alloy structure, achieving significant lightweight benefits. This study provides foundational data for the lightweight design of EMUs.

Keywords: Composite; Cheetah Optimizer; EMU; FEA

Online publication: September 26, 2025

1. Introduction

Composite materials, as new materials, are characterized by their lightweight, high strength, high customizability, reliability in extreme environments, and closed-loop recyclability, making them increasingly popular in transportation applications. With advancements in composite material research, their applications have expanded from aerospace to automotive and rail transportation. For instance, the CRH EMU's energy consumption decreased by 17% after adopting composite materials [1].

In the design of composite material components, optimization methods can be categorized into structural optimization and parameter optimization. Structural optimization is mainly carried out by changing the geometric shape of the entire composite material component, etc., while parameter optimization is mainly carried out by changing the ply angle, etc., without changing the geometric shape. Huang *et al.* conducted a lightweight design for a space-borne optical remote sensor using composite materials through structural topology optimization and fiber orientation parameter optimization, achieving structural lightweighting while meeting stiffness and strength requirements [2]. Islam *et al.* experimentally tested and simulated the mechanical properties of JUCO fiber-reinforced composites under tensile and bending static strength conditions [3]. They found out that the experimental results are highly consistent with the data of the finite element simulation results and fiber orientation angles significantly affect the mechanical performance of fiber-reinforced composites. This indicates that optimizing the lamination angles in simulations, without changing the number of layers, can effectively improve the static strength mechanical properties of composite materials.

Sohouli *et al.* developed a design optimization framework for composite material lamination angles based on a novel Decoupled Discrete Material Optimization, using non-convex and convex sequential approximation optimization methods to optimize composite material thin plates with four discrete angles (0°, 45 °,-45°,90°) [4]. Zhang *et al.* improved the Multi-Objective Particle Swarm Optimization (MOPSO) algorithm with intelligent algorithms, implementing multiscale simulations via Python code to achieve lightweighting and lamination angle optimization for carbon-fiber-reinforced composite drive shafts, enhancing comprehensive performance while meeting reliability requirements [5]. Tran *et al*. proposed a meta-heuristic optimization algorithm combining the Shrimp and Goby Association Search Algorithm (SGA), Balancing Composite Motion Optimization (BCMO), and Differential Evolution (DE), optimizing lamination angles for composite laminate plates under vibration and buckling conditions without four discrete angles [6].

Wu *et al.* used finite element methods to calculate the vibration frequency of composite structures and conducted a supporting vector regression dataset for optimizing lamination angles [7]. They will develop a multi-objective optimization loop based on machine learning. Wang *et al*. trained a Multi-DNN model and introduced transfer learning for fine-tuning to predict failure factors, combining genetic algorithms and models to optimize composite pressure vessels, obtaining the optimal lamination angles [8].

This paper focuses on the front skirt panels of EMUs, employing the Cheetah Optimizer in MATLAB with auxiliary iterations using ABAQUS and Python, optimizing lamination angles using the factor of three-dimensional Hashin failure criteria as design parameters. After 100 iterations, the optimal composite material lamination angles were obtained with the minimum overall failure factor.

2. Intelligent algorithms and failure criteria

The Cheetah Optimizer (CO) is inspired by the hunting behavior of cheetahs [9]. Cheetahs often hunt in groups, each group containing several cheetahs. Each cheetah represents a solution containing all parameters to be optimized. In this study, five lamination angles will be the parameters to be optimized. Since cheetahs are dispersed relative to their prey, the cheetah that is the closest one to the prey will become the leader, representing the current best solution. However, the leader may lose its leading position if it moves away from the prey or is surpassed by other cheetahs. The basic logic of the CO involves initializing cheetah groups, calculating fitness to identify the leader, and iterating to obtain the global optimal solution. The CO consists of four stages: searching for prey, waiting strategy, attack strategy, and abandoning the hunt.

During the prey search stage, a cheetah explores its territory to locate prey. For the cheetah it can choose two ways to locate prey, which are observing the prey in place and moving around actively to search for the prey. After spotting its prey, if the conditions aren't favorable for an immediate hunt, the cheetah will choose to hide and wait patiently for a better opportunity—either for the prey to come closer or for another to create a situation that aids in the hunt. The reason for this behavior is that although cheetahs have strong explosive power, they cannot maintain such a high-speed pursuit of prey for a long time. Therefore, cheetahs need to ensure the efficiency of catching prey. This is also the advantage of the CO, which will not waste computing power on possibly wrong solution routes. The waiting strategy inspired by this behavior involves hiding and waiting for a better opportunity to hunt.

The attack strategy includes two steps: sprinting and capturing the prey. Sprinting is the cheetah sprinting towards prey at its fastest speed. Capturing is when a cheetah catches the prey by its advantage of high-speed running and flexibility. Abandoning the hunt is also a strategy that avoids wasting computational resources on possibly incorrect solutions. There are two situations for this strategy. One of them is that the cheetah fails to hunt prey and return to its territory or original position, preparing for the next hunt. The other is when the hunting takes too much time, but still fails to hunt prey, then the cheetah will return to the latest position it found the prey to search for a new target.

The Hashin failure criterion used in this research, widely used for predicting composite material failures, divides failures into fiber failure and matrix failure, further subdivided into fiber tensile failure, fiber compressive failure, matrix tensile failure, and matrix compressive failure [10]. The failure criteria are calculated by Equations (1) to (4).

(1) Fiber tensile failure:

$$F_{ft} = \left(\frac{\sigma_1}{X_T}\right)^2 + \alpha\left(\frac{\sigma_{12}}{S_{12}}\right)^2 + \alpha\left(\frac{\sigma_{13}}{S_{13}}\right)^2 \geqslant 1 \tag{1}$$

(2) Fiber compressive failure:

$$F_{fc} = \left(\frac{\sigma_1}{X_C}\right)^2 \geqslant 1 \tag{2}$$

(3) Matrix tensile failure:

$$F_{mt} = \left(\frac{\sigma_2 + \sigma_3}{Y_T}\right)^2 + \left(\frac{1}{S_{23}^2}\right)\left(\sigma_{23}^2 - \sigma_{22}\sigma_{33}\right) + \left(\frac{\sigma_{12}}{S_{12}}\right)^2 + \left(\frac{\sigma_{13}}{S_{13}}\right)^2 \geqslant 1 \tag{3}$$

(4) Matrix compressive failure:

$$F_{mc} = \left(\frac{\sigma_2 + \sigma_3}{2S_{23}}\right)^2 + \left(\frac{\sigma_2 + \sigma_3}{Y_C}\right)\left[\left(\frac{Y_C}{2S_{23}}\right)^2 - 1\right] + \frac{1}{S_{23}^2}\left(\sigma_{23}^2 - \sigma_2\sigma_3\right) + \left(\frac{\sigma_{12}}{S_{12}}\right)^2 + \left(\frac{\sigma_{13}}{S_{13}}\right)^2 \geqslant 1 \tag{4}$$

3. Model and conditions

This study focuses on the lightweight and optimization simulation design of the front skirt panel of a city rail vehicle. The front of the vehicle experiences direct wind pressure during operation. The geometric model of the front is shown in **Figure 1**. The skirt panel is the area with the highest stress concentration, hence the choice for simulation. The T800S-3900-2B quasi-isotropic composite material was selected for the lightweight design of the front skirt panel, with material parameters listed in **Table 1** [11]. The finite element mesh model was established using Hypermesh software, with solid elements C3D8R, thickness 0.2 mm, and mesh size 10 mm. The main loading condition considered is surface pressure due to wind pressure during train operation. The surface pressure load value was calculated using Bernoulli's equation, shown by Equation (5), considering a maximum speed of 143 km/h and a wind speed of 23.53 m/s, resulting in a total wind speed of 63.25 m/s and a wind pressure load of 2.5 KPa [12]. The finite element model with applied loads is shown in **Figure 2**.

Figure 1. Geometric model of the vehicle front

Table 1. Material parameters of T800S-3900-2B

E_1(GPa)	E_2(GPa)	G_{12}(GPa)	v_{12}	v_{23}
132.6	8.0	4.03	0.34	0.45
X_T(MPa)	X_C(MPa)	Y_T(MPa)	Y_C(MPa)	S_L(MPa)
3100	1242	56.5	307	90.3

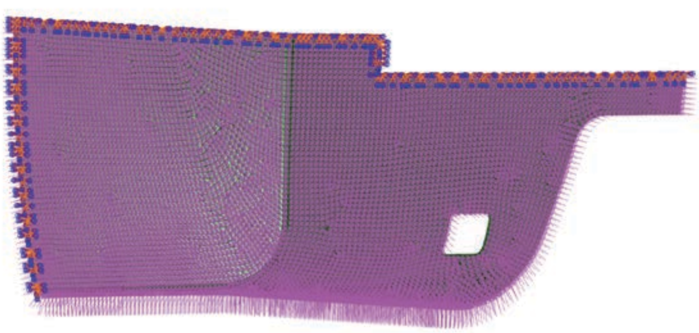


Figure 2. FE model of the skirt panel with conditions

$$wp = 0.5\rho v^2 \tag{5}$$

4. Lightweight optimization simulation

The Cheetah Optimizer was executed on the MATLAB platform, with six cheetah groups, each containing two cheetahs. Since the two failure factors of composite skirt panels at the current layer count were similar, the average of the two failure factors was used as the evaluation criterion, the overall failure factor. MATLAB commands were used to call ABAQUS for failure factors calculations. Since MATLAB cannot directly read ABAQUS OBD files, Python scripts were written for initializing the cheetah groups, performing the search phase, and abandoning the hunting phase, ultimately outputting failure factors to MATLAB variables. After 100 iterations, the optimal lamination angles were [-14, -18, -31, -3, 24]s with an overall failure factor of 0.72. The entire optimization result is shown in **Figure 3**. Some lamination angle designs had overall failure factors greater than 1, indicating that at least one failure mode will occur, confirming that the layer count design was not overly redundant, thus achieving effective lightweighting.

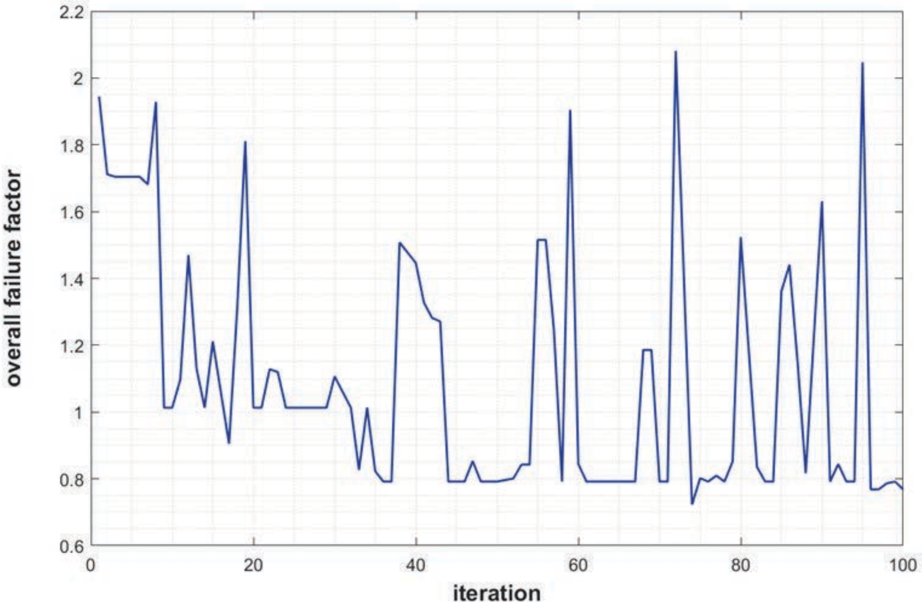


Figure 3. Optimization result

5. Conclusion

This study used T800S-3900-2B quasi-isotropic composite material to replace the original aluminum alloy material for the front skirt panel of a city rail vehicle. The surface pressure was calculated using Bernoulli's equation. Using MATLAB and Python in conjunction with ABAQUS, the Cheetah Optimizer was applied for 100 iterations, yielding the optimal lamination angles under common operating conditions. Compared to the original aluminum structure, the skirt panel's weight was reduced by 60%, providing reference data for applying composite materials to other force-bearing components of rail vehicles.

Disclosure Statement

The authors declare no conflict of interest.

References

[1] Ying L, Qian W, Lei L, et al., 2025, Structural Integrity Issues of Composite Materials and Structures in Future Transportation Equipment. Composite Structures, 358: 118943.

[2] Huang B, Li ZH, Tian XZ, et al., 2022, Concurrent Topology and Fiber Orientation Optimization of CFRP Structures in Space-Borne Optical Remote Sensor. Optik, 267: 169652.

[3] Islam MR, Hossain MF, Rana MS, et al., 2025, Effect of Fiber Orientation on Mechanical Properties of JUCO Fiber Reinforced Epoxy Composites. Hybrid Advances, 8: 100386

[4] Sohouli A, Yildiz M, Suleman A, 2017, Design Optimization of Thin-Walled Composite Structures Based on Material and Fiber Orientation. Composite Structures, 176: 1081–1095.

[5] Zhang H, Li S, Zh P, et al., 2024, A Multiscale Reliability-Based Design Optimization Method for Carbon-Fiber-Reinforced Composite Drive Shafts. CMES–Computer Modeling in Engineering & Sciences, 140(2): 1975–1996.

[6] Tran VT, Nguyen TK, Nguyen Xuan H, et al., 2024, Meta-Heuristic Optimization Algorithms for Vibration and Buckling Analysis of Laminated Composite Plates. Engineering Analysis with Boundary Elements, 169(PA): 105974.

[7] Wu C, Zhang R, Tang F, et al., 2023, Vibration Optimization of Cantilevered Bistable Composite Shells Based on Machine Learning. Engineering Applications of Artificial Intelligence, 126(5): 107158.

[8] Wang Q, Qin H, Jia L, et al., 2024, Failure Prediction and Optimization for Composite Pressure Vessel Combining FEM Simulation and Machine Learning Approach. Composite Structures, 337: 118099.

[9] Akbari MA, Zare M, Azizipanah-Abarghooee R, et al., 2022, The Cheetah Optimizer: A Nature-Inspired Metaheuristic Algorithm for Large-Scale Optimization Problems. Scientific Reports, 12(1): 10953.

[10] Zhang C, Duodu EA, Gu J, 2017, Finite Element Modeling of Damage Development in Cross-Ply Composite Laminates Subjected to Low Velocity Impact. Composite Structures, 173: 219–227.

[11] Divse V, Marla D, Joshi SS, 2023, Progressive Damage Analysis in an Open Hole Compression of FRP Laminates Including Fiber Kinking and Pre-Existing Damage. Composites Part A: Applied Science and Manufacturing, 169: 107523.

[12] Singh D, 2018, Comparison of Carbon Steel and Composite Side Wall of Light Rail Vehicle by Finite Element Analysis, thesis, The University of Texas at Arlington.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Application and Prospects of SDN Technology in Modern Network Management

Aoyu Li*, Yingjie Yang

Information Engineering University, Zhengzhou 450001, China

**Author to whom correspondence should be addressed.*

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: With the rapid development of information technology, the scale of the network is expanding, and the complexity is increasing day by day. The traditional network management is facing great challenges. The emergence of software-defined network (SDN) technology has brought revolutionary changes to modern network management. This paper aims to discuss the application and prospects of SDN technology in modern network management. Firstly, the basic principle and architecture of SDN are introduced, including the separation of control plane and data plane, centralized control and open programmable interface. Then, it analyzes the advantages of SDN technology in network management, such as simplifying network configuration, improving network flexibility, optimizing network resource utilization, and realizing fast fault recovery. The application examples of SDN in data center networks and WAN optimization management are analyzed. This paper also discusses the development status and trend of SDN in enterprise networks, including the integration of technologies such as cloud computing, big data, and artificial intelligence, the construction of an intelligent and automated network management platform, the improvement of network management efficiency and quality, and the openness and interoperability of network equipment. Finally, the advantages and challenges of SDN technology are summarized, and its future development direction is provided.
Keywords: Software-defined network; Network management; Data centers; Wide area network; Cloud computing

Online publication: September 26, 2025

1. Introduction

With the rapid development of information technology, the network has become one of the infrastructures of modern society. The traditional network architecture and management methods have been unable to meet the growing network demand, so software-defined network (SDN) technology came into being, bringing new solutions and unlimited possibilities for modern network management. By separating the control plane and the data plane of the network, SDN realizes the programmability, controllability, and flexibility of the network [1], so that the network manager can configure and manage the network more conveniently, which brings new opportunities and challenges for modern network management. At the same time, SDN technology can optimize the utilization of network resources, improve network performance, and realize rapid fault recovery, which brings many advantages to modern network management. In modern network management, the application prospects of SDN technology are very broad. It can be applied to network traffic control to improve network bandwidth utilization and user experience through intelligent scheduling and optimization of network traffic. In addition, SDN technology can also be used to implement security policies through centralized management and control of security policies, enhancing the security protection capability of the network [2]. At the same time, it can also be applied to quality of service management to achieve fine quality of service control and improve the reliability and stability of network services. In addition, SDN technology can also support multi-tenant network management, meet the network requirements of different tenants, and realize the sharing and optimal utilization of network resources. Looking forward to the future, SDN technology will continue to be deeply integrated with modern network technology to promote the intelligent and automated development of network management. With the continuous progress of cloud computing [3], big data, artificial intelligence, and other technologies, SDN technology will play a more important role in network management, laying a solid foundation for building an efficient, secure, and intelligent modern network management system. Therefore, in-depth discussion of the application prospects of SDN technology in modern network management will not only help promote the innovation and development of network management technology, but also provide more reliable and efficient network services for modern society, and help the sustainable and healthy development of the information society. This paper will discuss the basic principle, application examples, and development status and trend of SDN, in order to provide references for promoting the further development of SDN technology [4].

2. Basic principle and architecture of SDN

The core principle of SDN is to separate the control plane and the data plane of the network, which makes the network more programmable, controllable, and flexible. In the traditional network architecture, the network device is usually a whole, with limited control and management functions, and mainly depends on the management panel of the device. By introducing the SDN controller, SDN technology centralizes the abstract and logical control functions of network resources and realizes the centralized control and management of the network. Specifically, the separation of SDN is achieved through the SDN controller and the SDN data plane. SDN switches are divided into data planes and control planes. The data plane is responsible for managing the data flow of the switch, including packet forwarding and stream operation, etc., while the control plane is the core part of configuring, managing, and monitoring the switch [5], which is realized by the SDN controller in SDN. As the core part of the control surface, the SDN controller abstracts the network resources into the logical layer and realizes the core functions such as centralized control, policy definition, and control policy implementation. In the SDN architecture, the SDN controller layer communicates with the SDN application layer through the SDN northbound interface, and with the SDN infrastructure layer through the SDN southbound interface. The SDN infrastructure layer consists of SDN-enabled switches that support the SDN southbound protocol and can be either physical switches or virtual switches. SDN southbound interfaces, such as the OpenFlow protocol, provide communication implementations between the SDN controller layer and the SDN infrastructure layer. In addition, the typical architecture of SDN can be divided into three layers: the upper layer is the application layer, including a variety of different services and applications; the middle control layer is mainly responsible for arranging data plane resources, maintaining network topology, state information, etc. The lowest infrastructure layer is responsible for data processing, forwarding, and state collection. This three-tier architecture makes SDN networks cleaner, more modular, and easier to manage. In general, SDN realizes centralized control, flexible programming, and efficient management of the network through its unique separation principle and three-layer architecture, bringing new solutions and unlimited possibilities to modern network management.

3. SDN application examples

3.1. Data center network

Data center network is an important application scenario of SDN technology. Traditional data center networks usually adopt complex and static routing protocols, which are difficult to adapt to the rapidly changing network requirements [6]. The application of SDN technology can realize flexible network topology configuration, dynamic adjustment of network traffic, and rapid deployment and migration of VMS, thereby improving the utilization and flexibility of data centers.

3.2. Wide Area Network optimization and management

Wide Area Network (WAN) optimization and management is also an important application area of SDN technology. Traditional WAN management often requires manual configuration of network devices, management of complex VPN connections, etc., which is inefficient and error-prone. SDN technology can realize centralized and automated network management, automatically adjust network topology according to the real-time demand of network traffic, and optimize network performance [7]. At the same time, SDN can be combined with traffic engineering technology to achieve load balancing, link fault recovery, and other functions to improve the reliability and quality of network services.

4. SDN development status and trend

4.1. Application and progress in the enterprise network

In the enterprise network, the application of SDN technology has achieved remarkable results. By using SDN technology, enterprises can realize more flexible network policy management [8] and security protection, and improve network reliability and manageability. At the same time, SDN can also be deeply integrated with cloud computing, big data, artificial intelligence, and other technologies to provide enterprises with more intelligent and efficient network services.

4.2. Open standards and interoperability

With the continuous development of SDN technology, open standards and interoperability have become an important direction of its development. By embracing open standards, SDN can prevent vendor lock-in, reduce equipment costs, and accelerate the deployment of new capabilities. At the same time, SDN will also promote the openness and interoperability of network equipment, so that devices from different vendors can seamlessly integrate and work together.

5. Results and discussion

5.1. Results

As a key component of modern network management, SDN technology has a broad and promising application prospect. With the continuous maturity of technology and the continuous expansion of the market, the application of SDN in modern network management has achieved remarkable results. Firstly, SDN technology realizes unified management and scheduling of network resources through its centralized control plane. This enables network administrators to configure, monitor, and optimize the network more conveniently, and improves the efficiency and accuracy of network management. At the same time, the openness and programmability of SDN also provide more possibilities for network innovation and promote the continuous development of network technology. Secondly, SDN technology helps to improve the security performance of the network. Through centralized security policy management and real-time security incident response [9], SDN can effectively prevent network attacks and threats and protect network security and stability. In addition, SDN can also achieve fine control of network traffic, effectively preventing network congestion and malicious traffic spread. Finally, the application of SDN technology in emerging fields such as cloud computing, big data, and the Internet of Things has also made remarkable progress. Through the deep integration with these technologies, SDN provides more efficient, flexible, and secure network services for various industries, and promotes the digital transformation and upgrading of the industry.

5.2. Discussion

Although the application of SDN technology in modern network management has broad prospects, there are also some challenges and problems that need to be further discussed and solved. First of all, the promotion and application of SDN technology needs to overcome the problem of inconsistent technical standards and protocols. Currently, there are many different technologies and standards in the SDN market, which leads to compatibility issues between devices and complexity in network management. Therefore, we need to strengthen the standardization work and promote the unified and standardized development of SDN technology. Secondly, the security and reliability of SDN technology are also important issues that we need to pay attention to. With the increasing number of network attacks and threats [10], the security of the SDN network is faced with severe challenges. We need to strengthen the security protection and monitoring of the SDN network and improve the anti-attack capability of the network. At the same time, it is also necessary to pay attention to the reliability and stability of the SDN network to ensure network continuity and availability. Finally, the application of SDN technology also needs to consider the problem of cost and benefit. Although SDN technology brings many advantages, its construction and maintenance costs are also relatively high. Therefore, we need to comprehensively consider the costs and benefits, and formulate reasonable investment and operation strategies to ensure that the application of SDN technology can bring actual economic and social benefits.

6. Conclusion

To sum up, SDN technology has a wide range of application prospects and great development potential in modern network management. Through continuous technological innovation and application practice, we are expected to build SDN technology into a more intelligent, efficient, and secure network management solution, providing strong support and guarantee for the development of modern society.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Zhao Y, Iannone L, Riguidel M, 2015, On the Performance of SDN Controllers: A Reality Check, Proc. IEEE Conf. Netw. Function Virtualization Softw. Defined Netw., 79–85.

[2] Xia W, Wen Y, Foh CH, et al., 2015, A Survey on Software-Defined Networking. IEEE Commun. Survey Tuts., 17(1): 27–51.

[3] Yu FR, Leung VMC, 2015, Advances in Mobile Cloud Computing Systems, CRC Press, New York.

[4] Lin YD, Pitt D, Hausheer D, et al., 2014, Software-Defined Networking: Standardization for Cloud Computing's Second Wave. Computer, 47(11): 19–21.

[5] Zhang P, Zhou M, Fortino G, 2018, Security and Trust Issues in Fog Computing: A Survey. Future Gener. Comput. Syst, 88(2018): 16–27.

[6] Van Oorschot PC, Smith SW, 2019, The Internet of Things: Security Challenges, IEEE Secur. Privacy, 17(5): 7–9.

[7] Boroojeni KG, Amini MH, Iyengar S, 2017, Overview of the Security and Privacy Issues in Smart Grids, Smart Grids: Security and Privacy Issues, Springer, 1–16.

[8] Babar M, Tariq MU, Jan MA, 2020, Secure and Resilient Demand Side Management Engine Using Machine Learning for IoT-Enabled Smart Grid. Sustainable Cities Soc, 62: 102370.

[9] Wang J, Wang L, 2022, SDN-Defend: A Lightweight Online Attack Detection and Mitigation System for DDoS Attacks in SDN. Sensors, 22: 8287.

[10] Manso P, Moura J, Serrao C, 2019, SDN-Based Intrusion Detection System for Early Detection and Mitigation of DDoS Attacks. IEEE Access, 10: 106.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Construction of a Virtual Twin Testing Framework for Safety of the Intended Functionality in Intelligent Connected Vehicles

Quanyou Fu, Daxu Sun*

Foshan Polytechnic, Foshan 528137, China

**Author to whom correspondence should be addressed.*

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: This study aims to construct a virtual twin testing framework for the safety of the intended functionality of intelligent connected vehicles to address the safety requirements of intelligent driving and transportation systems. The research methods include the construction of a theoretical model of safety for intelligent connected vehicles based on the concept of virtual twins, the correlation study between key concepts and functional safety, and the application research of virtual twin technology in the safety testing of intelligent connected vehicles. The results reveal that the virtual twin testing framework can effectively enhance the functional safety of intelligent connected vehicles, reduce development costs, and shorten the product launch cycle. The conclusion suggests that this framework provides strong support for the healthy development of the intelligent connected vehicle industry and has a positive impact on the safety and efficiency of intelligent transportation systems.

Keywords: Intelligent connected vehicles; Safety of the intended functionality; Virtual twin; Testing framework; Safety theory model

Online publication: October 15, 2025

1. Introduction

With the rapid development of intelligent connected vehicle technology, the demand for safety in intelligent driving and transportation systems has become increasingly prominent. At the policy level, in recent years, national, regional, and international societies have all placed a high priority on the safety of intelligent connected vehicles. The introduction of regulations such as China's "Management Regulations for Intelligent Connected Vehicle Road Testing (Trial)" aims to establish a testing management system for intelligent connected vehicles. Additionally, international policies such as Europe's "Guidelines for the Safety of Intelligent Connected Vehicles" and the United States' "Guidelines for the Testing and Certification of Autonomous Vehicles" provide direction for safety research in the field of intelligent connected vehicles.

In this policy context, the construction of a virtual twin testing framework for the safety of the intended functionality of intelligent connected vehicles holds significant importance. On one hand, this framework can enhance the safety of intelligent connected vehicles during development, testing, and operation, thereby ensuring traffic safety and improving the user experience. On the other hand, by simulating real-world environments using virtual twin technology, it can reduce R&D costs and shorten the product launch cycle. However, research on the safety of the intended functionality of intelligent connected vehicles is still in its infancy, facing challenges in terms of technology, theory, and practice.

This paper aims to construct a virtual twin testing framework for the safety of the intended functionality of intelligent connected vehicles to meet the safety requirements of intelligent driving and transportation systems. From an academic perspective, this research will help expand the application of virtual twin technology in the field of intelligent connected vehicles and provide new insights into the development of safety theory for intelligent connected vehicles. From a practical perspective, this research will provide effective testing methods for intelligent connected vehicle development companies and promote the healthy development of the intelligent connected vehicle industry.

2. Research questions and innovative contributions

The core issue addressed in this paper is how to construct a virtual twin testing framework for the safety of the intended functionality of intelligent connected vehicles. Specifically, this paper will explore the following aspects:

(1) Construction of a theoretical model for the safety of intelligent connected vehicles based on the concept of virtual twins;

(2) Research on the relationship between key concepts and functional safety;

(3) Application research on virtual twin technology in safety testing for intelligent connected vehicles.

The innovative contributions of this paper are primarily reflected in the following aspects:

(1) The first application of virtual twin technology in the field of intelligent connected vehicle safety testing, proposing a novel testing framework;

(2) The construction of a theoretical model for the safety of intelligent connected vehicles based on the virtual twin concept, providing theoretical support for subsequent research;

(3) Overcoming the limitations of traditional testing methods by simulating real-world environments using virtual twin technology to enhance testing efficiency and quality.

This paper is aimed at intelligent connected vehicle development companies and government regulatory agencies. Based on virtual twin theory and utilizing virtual twin technology methods, it seeks to establish a testing framework for the safety of the intended functionality of intelligent connected vehicles, providing theoretical guidance and practical references for the development of the intelligent connected vehicle industry.

3. Review of domestic and international research

The focus is on three research areas: intelligent driving and assistance systems, intelligent transportation infrastructure, and the intelligent transportation field. In-depth discussions are conducted on safety and performance optimization, as well as safety and efficiency issues. First, the research focuses on intelligent driving and assistance systems to address safety and performance optimization issues. Awasthi et al. proposed a framework based on Bayesian optimization (BO) to accelerate the discovery of critical scenarios, demonstrating that this framework can significantly reduce the number of simulations and effectively identify dangerous situations such as boundary-crossing events, thereby improving the verification efficiency of autonomous driving functions (ADFs) and positively impacting the safety and public acceptance of autonomous vehicles (AVs) [1]. Mo et al. proposed that by conducting virtual testing on the CARLA platform and adopting a human-shaped scene generation (HSG) scheme, the limitations of scarce and non-reproducible dangerous behavior samples in real-world scenarios can be overcome, thereby significantly enhancing the realism and diversity of collision rate simulations for AVs. This scheme effectively validated the important role of HSG in improving autonomous driving safety performance [2]. Second, focus on intelligent transportation infrastructure to address safety and efficiency issues. Kloeker et al. proposed that the application of intelligent roadside infrastructure sensors is critical to the future development of connected and autonomous vehicles, and that the selection of sensor configurations significantly impacts data quality and downstream functionality. Their proposed multimodal framework effectively evaluates the performance of different sensor types in terms of accuracy, latency, and reliability, providing a reliable quality assessment tool for future intelligent transportation system applications [3]. Zhang et al. proposed the construction of a hybrid traffic environment active safety analysis platform based on digital twins, which integrates multi-source data such as drone lidar, OpenStreetMap, and vehicle sensors to generate high-resolution 3D road geometry. By utilizing the CARLA simulator, SUMO traffic model, and NVIDIA PhysX vehicle dynamics engine, the platform simulates real-world driving scenarios, significantly enhancing the effectiveness of active safety measures and driving traffic safety research toward deeper and more physically informed directions [4]. Third, focusing on the intelligent transportation field to address safety and efficiency issues. Wang et al. proposed that by introducing edge local digital twin (LDT) technology, information exchange and driving experience extraction in vehicle-road collaboration within intelligent transportation can be enhanced, significantly improving the safety and efficiency of autonomous driving systems. This results in a 10% improvement in traffic intersection safety performance and a 15% reduction in travel time, laying a solid foundation for the future development of autonomous driving [5]. Xia et al. proposed that due to the limited ability of existing autonomous vehicles to handle high-intensity computational tasks, coupled with the prevalence of malicious cyberattacks and threats to vehicle information privacy, designing an efficient evaluation system to ensure autonomous driving safety without compromising data security has become particularly urgent [6].

A review of domestic and international research indicates that studies on intelligent driving and assistance systems, intelligent transportation infrastructure, and the intelligent transportation field have all focused on safety and performance optimization, as well as safety and efficiency issues. Through innovative methods such as Bayesian optimization, virtual testing, and multimodal frameworks, significant improvements have been achieved in autonomous driving safety and transportation system efficiency [7]. The following issues remain to be addressed: (1) Insufficient innovation in research frameworks. Although innovative methods such as Bayesian optimization have achieved significant results in autonomous driving safety verification, existing research lacks systematic innovation in framework design and has failed to fully integrate interdisciplinary knowledge to achieve more comprehensive and efficient solutions [8]. (2) Limitations of simulated scenarios. Current virtual testing and human-like scenario generation schemes have limitations in the construction of simulated scenarios, making it difficult to fully reproduce the complexity and dynamic changes of real traffic environments, leading to discrepancies between autonomous driving performance evaluation results and actual applications [9]. (3) Cross-domain collaboration challenges. The collaborative development of intelligent driving and assistance systems, intelligent transportation infrastructure, and intelligent transportation faces challenges due to insufficient cross-domain knowledge integration, making it difficult to share research results across disciplines and limiting the overall performance improvement of intelligent transportation systems [10]. To address the above three issues, this paper first conducts an in-depth analysis of safety theory models based on actual operational data and traffic scenarios of intelligent connected vehicles, aiming to reveal the safety performance of intelligent connected vehicles in complex traffic environments. Second, this paper designs a testing framework based on virtual twin technology, simulating real driving environments to comprehensively evaluate the safety performance of intelligent connected vehicles. The results demonstrate that this framework can effectively identify potential safety risks, improving testing efficiency and accuracy. Third, this paper verifies the advantages of virtual twin technology in intelligent connected vehicle safety testing through comparative analysis and proposes corresponding solutions. Experimental results indicate that the proposed virtual twin testing framework can significantly reduce R&D costs and shorten the product time-to-market. Finally, this study introduces the following innovative points in addressing the problem: first, it constructs a theoretical model for the safety of intelligent connected vehicles based on virtual twin concepts, providing theoretical support for subsequent research; second, it proposes a new virtual twin testing framework, breaking through the limitations of traditional testing methods; finally, by simulating real-world environments, it improves testing efficiency and quality, providing a strong foundation for the healthy development of the intelligent connected vehicle industry.

4. Theoretical framework

This section aims to elaborate on the theoretical basis and core model for constructing a virtual twin testing framework for the safety of the intended functionality of intelligent connected vehicles.

First, virtual twin technology, as an emerging technology that has gained traction in recent years, holds broad application prospects in the field of intelligent connected vehicles [11]. Virtual twin technology involves creating a holographic replica of a physical system or entity from the real world in a virtual space through digital means, enabling real-time monitoring, analysis, and control of the entity's state. In the field of intelligent connected vehicles, virtual twin technology can be applied throughout the entire process of vehicle design, manufacturing, operation, and maintenance. Simulating real traffic scenarios in a virtual environment enables the assessment and optimization of the safety performance of intelligent connected vehicles [12]. However, the application of existing virtual twin technology in the field of intelligent connected vehicles is still in its infancy, with technical bottlenecks and deficiencies in theoretical frameworks. This study takes virtual twin technology as its foundation and combines the needs of intelligent connected vehicle safety performance assessment to construct a theoretical framework for the application of virtual twin technology in the field of intelligent connected vehicles.

Second, an intelligent connected vehicle safety performance assessment is a critical step in ensuring the safety of intelligent connected vehicles. Current methods for assessing the safety performance of intelligent connected vehicles primarily include experimental testing, simulation modeling, and virtual testing. Experimental testing requires on-road testing, which is costly and time-consuming; simulation modeling can reduce costs but struggles to fully replicate real-world traffic scenarios; virtual testing effectively combines the advantages of experimental testing and simulation modeling but faces challenges such as the difficulty of constructing testing scenarios and discrepancies between testing results and real-world applications [13–15]. This study is based on virtual twin technology to construct a virtual twin testing framework for the safety of the intended functionality of intelligent connected vehicles, aiming to address the shortcomings of existing safety performance evaluation methods and improve testing efficiency and accuracy.

Based on the aforementioned theoretical foundation, this paper proposes the following theoretical assumptions:

(1) Virtual twin technology can be effectively applied to the safety performance assessment of intelligent connected vehicles, enabling a comprehensive evaluation of their safety performance through the simulation of real-world traffic scenarios.

(2) A testing framework based on virtual twin technology can reduce R&D costs, shorten product time-to-market, and enhance testing efficiency and accuracy.

(3) The application of virtual twin technology in the safety performance assessment of intelligent connected vehicles will promote the healthy development of the intelligent connected vehicle industry.

The applicability of this framework is primarily reflected in the following aspects:

(1) Virtual twin technology can achieve real-time monitoring, analysis, and control of the safety performance of intelligent connected vehicles, thereby improving testing efficiency and accuracy.

(2) Virtual twin technology can simulate real traffic scenarios, providing reliable data support for the safety performance assessment of intelligent connected vehicles.

(3) Virtual twin technology can reduce R&D costs, shorten product time-to-market, and help promote the healthy development of the intelligent connected vehicle industry.

However, this framework also has certain limitations. First, the application of virtual twin technology requires a large amount of computing resources and has high hardware requirements; Second, virtual twin technology may encounter issues such as data synchronization and model accuracy in practical applications; finally, the application of virtual twin technology in the safety performance assessment of intelligent connected vehicles is still in its exploratory phase and requires further research and refinement.

This paper is based on virtual twin theory and relies on a virtual twin testing framework, aiming to promote in-depth exploration and understanding of safety of the intended functionality issues in intelligent connected vehicles.

5. Conclusion

This study established a virtual twin testing framework for the safety of the intended functionality of intelligent connected vehicles and conducted an in-depth exploration of its theoretical framework, application prospects, and implementation path. By analyzing the integration of virtual twin technology with safety testing for intelligent connected vehicles, this study revealed the importance and necessity of establishing this framework and proposed corresponding countermeasures and recommendations. The study demonstrates that the virtual twin testing framework can effectively enhance the functional safety of intelligent connected vehicles, reduce R&D costs, shorten product time-to-market, and provide strong support for the healthy development of the intelligent connected vehicle industry. In the future, this study will further explore the application of the virtual twin testing framework in the field of intelligent connected vehicles, aiming to provide more effective solutions for the safety and efficiency of intelligent transportation systems. The completion of this study not only enriches the theoretical foundation of intelligent connected vehicle safety but also provides valuable references and insights for the development of China's intelligent connected vehicle industry. Under the backdrop of the new era, this study holds significant theoretical and practical value, contributing to the continuous innovation and development of the intelligent connected vehicle industry.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Awasthi SS, Imran MIIS, Arrigoni S, et al., 2025, Bayesian Optimization Applied for Accelerated Virtual Validation of the Autonomous Driving Function, ArXiv, http://arxiv.org/abs/2507.22769v2

[2] Mo L, Hua M, Sun H, et al., 2023, Study on the Impacts of Hazardous Behaviors on Autonomous Vehicle Collision Rates Based on Humanoid Scenario Generation in CARLA, ArXiv, http://arxiv.org/abs/2307.10229v1

[3] Kloeker L, Liu C, Wei C, et al., 2023, Framework for Quality Evaluation of Smart Roadside Infrastructure Sensors for Automated Driving Applications, ArXiv.

[4] Zhang H, Yue X, Tian K, et al., 2025, Virtual Roads, Smarter Safety: A Digital Twin Framework for Mixed Autonomous Traffic Safety Analysis, ArXiv, http://arxiv.org/abs/2504.17968v1

[5] Wang K, Nonomura K, Li Z, et al., 2024, Augmented Intelligence in Smart Intersections: Local Digital Twins-Assisted Hybrid Autonomous Driving, ArXiv, http://arxiv.org/abs/2410.12163v2

[6] Xia L, Sun Y, Swash R, et al., 2021, Smart and Secure CAV Networks Empowered by AI-Enabled Blockchain: The Next Frontier for Intelligent Safe Driving Assessment, ArXiv, http://arxiv.org/abs/2104.04572v5

[7] Paranjape A, Patwardhan Y, Deshpande V, et al., 2023, Voice-Based Smart Assistant System for Vehicles using RASA, ArXiv, http://arxiv.org/abs/2312.01642v1

[8] Emami Y, Zhou H, Gaitan MG, et al., 2025, From Prompts to Protection: Large Language Model-Enabled In-Context Learning for Smart Public Safety UAV, ArXiv, http://arxiv.org/abs/2506.02649v1

[9] Collin A, Bilka A, Pendleton S, et al., 2021, Safety of the Intended Driving Behavior Using Rulebooks, ArXiv, http://arxiv.org/abs/2105.04472v1

[10] Wu K, Li P, Cheng Y, et al., 2024, A Digital Twin Framework for Physical-Virtual Integration in V2X-Enabled Connected Vehicle Corridors, ArXiv, http://arxiv.org/abs/2410.00356v2

[11] Nikouei SY, Xu R, Chen Y, et al., 2019, Decentralized Smart Surveillance through Microservices Platform, ArXiv, http://arxiv.org/abs/1903.04563v1

[12] Martins V, Rufino J, Fernandes B, et al., 2018, Personal Virtual Traffic Light Systems, ArXiv, http://arxiv.org/abs/1809.07829v1

[13] Czarnecki K, Kuwajima H, 2023, STEAM & MoSAFE: SOTIF Error-and-Failure Model & Analysis for AI-Enabled Driving Automation, ArXiv, http://arxiv.org/abs/2312.09559v2

[14] Limbasiya T, Das D, Sahay SK, 2019, Secure Communication Protocol for Smart Transportation Based on Vehicular Cloud, ArXiv, http://arxiv.org/abs/1912.12884v2

[15] Beaver LE, Chalaki B, Mahbub AMI, et al., 2019, Demonstration of a Time-Efficient Mobility System Using a Scaled Smart City, ArXiv, http://arxiv.org/abs/1903.01632v2

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Exploration of the Evolution of LiDAR Technology

Haotian Chen*, Tao Xi, Lei Wang

Armed Police Non-Commissioned Officer School, Hangzhou 311400, China

***Corresponding author:** Haotian Chen, chenht1019@163.com

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: Since its inception in the 1960s, light detection and ranging (LiDAR) technology has demonstrated great potential in various fields such as autonomous driving, robot navigation, and environmental monitoring due to its high precision, high resolution, and strong anti-interference capability. This paper reviews the development history, technical principles, application fields, and future development trends of LiDAR technology. It introduces the technical applications of LiDAR technology in autonomous driving, robot navigation, and environmental monitoring, and explores the development direction of SLAM algorithms in multi-sensor fusion and real-time map construction, providing a reference basis for the development and research of LiDAR.

Keywords: LiDAR; Autonomous driving; Robot navigation; Environmental monitoring

Online publication: October 15, 2025

1. Introduction

Light detection and ranging (LiDAR), as an advanced active remote sensing technology, measures the distance, position, velocity, and other information of target objects by emitting laser beams and receiving their reflected signals. It has the advantages of high precision, high resolution, and strong anti-interference ability. With the rapid development of technologies related to autonomous driving, robot navigation, environmental monitoring, and other fields, LiDAR technology has received widespread attention and application [1]. This paper aims to provide a reference for further research and application of LiDAR technology by systematically reviewing its development history, technical principles, application status, and future trends.

Since the first ruby laser was developed by American scientist Maiman in the 1970s, LiDAR technology has experienced rapid development. Foreign countries started research on LiDAR technology earlier and have accumulated rich technical experience. From the initial laser rangefinders to two-dimensional scanning LiDAR, and then to three-dimensional scanning vehicle-mounted LiDAR, foreign LiDAR technology has been continuously iterated and upgraded. Currently, in the global LiDAR industry, companies such as Velodyne (USA), TopoSys (France), IBEO (Germany), and Sick (Germany) occupy leading positions due to their technological advantages [2].

Compared to the research level of other countries, China started its research in the field of LiDAR technology relatively late. However, after rapid development in recent years, and with the support of national key research and development programs, it has achieved international leading results in the fields of single-photon avalanche diodes and superconducting nanowire single-photon detectors. Solid-state LiDARs developed by companies such as Huawei and DJI have achieved mass production. Domestic universities and research institutions have accumulated relatively deep technical expertise in laser technology through years of research, but are relatively lagging behind in the commercialization of laser products. With the rise of the autonomous driving industry, domestic LiDAR manufacturers such as RoboSense, SureStar, and LeiShen Intelligent have emerged, and their products have gradually gained market recognition.

However, compared with leading foreign products, domestic LiDAR still lags behind in accuracy and stability, requiring continuous investment in research and development.

2. Laser radar technology principles

2.1. Basic principles of laser radar

Laser radar emits laser pulses and receives echo signals reflected by target objects, using the time difference between the transmitted and received signals to calculate information such as the distance and velocity of the target objects. This process includes four main stages: laser emission, signal propagation, echo reception, and signal processing. According to different ranging principles, laser radar can be divided into triangulation laser radar and time-of-flight (TOF) laser radar.

Triangulation laser rangefinders use the triangular relationship formed by the reflection point of a laser beam on the surface of a target object and the transceiver to measure distance. These laser rangefinders consist of a laser emitter and a receiver, where the receiver can be regarded as a pinhole camera model. The emitter emits a laser beam, which is reflected when it reaches the surface of an object, and the reflected beam is captured by the receiver.

As shown in **Figure 1**, there is a certain distance between the emitter and the receiver, which is called the baseline length s. By measuring the position change of the reflected beam on the receiver, the distance of the target object can be calculated according to the trigonometric formula, combined with the known baseline length and angle information. However, the measurement accuracy of triangulation LiDAR is greatly affected by the distance of the target object, and the accuracy decreases significantly during long-distance measurement.

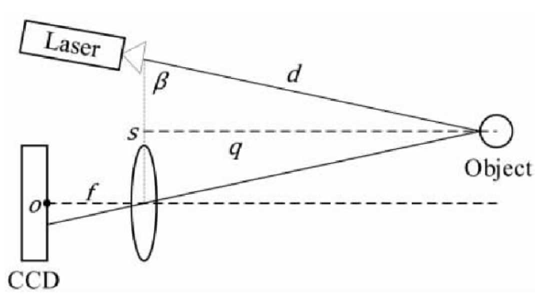


Figure 1. Principle diagram of triangulation laser rangefinder [2]

As shown in **Figure 1**, the linear distance between the emitter and the target object is d, the angle between the baseline and the emitter's optical axis is β, the distance between the target object and the baseline is q, the center of the linear CCD is point o, the focal length of the lens and the CCD chip is f, and the distance between the projection point of the reflected optical axis on the CCD and point o is x. Then:

TOF LiDAR calculates the distance between the LiDAR and the target object by measuring the time it takes for a laser pulse to travel from emission to reception (**Figure 2**) [3]. Its internal timer records the moment of laser emission and the moment of reception, and the measured distance is obtained by calculating the product of the time difference and the speed of light [4]. Although TOF LiDAR has the advantages of small size and strong anti-interference ability, its measurement accuracy is limited by the timing accuracy of the timer, and the power consumption is large.

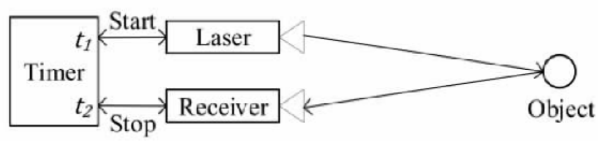


Figure 2. TOF LiDAR principle diagram [2]

The time difference t_2-t_1 calculated based on the laser emission time t_1 and reception time t_2 recorded by the internal high-precision timer, is called time of flight. Given the speed of light c, the observation distance d can be calculated as:

2.2. LiDAR key technologies

As the principle of LiDAR indicates, the precise measurement of LiDAR relies on the processing of information from its transmitted and received signals. Therefore, the accuracy of LiDAR's signal transmission and reception is one of the core key technologies to ensure its signal precision. High-power and high-stability laser transmitters can ensure the energy density and directionality of the laser beam, improving ranging accuracy and anti-interference ability [5]. At the same time, high-sensitivity receivers can accurately capture reflected beams, improving the accuracy and reliability of signal processing.

In practical application scenarios, the echo signals received by LiDAR are susceptible to environmental factors, often mixed with noise or interference signals generated by other devices, requiring signal processing algorithms to complete key processing steps such as filtering, denoising, and feature extraction [6].

To mitigate the negative impact of environmental noise and interference on signal quality, signal processing algorithms have become one of the core key technologies of LiDAR. Commonly used signal processing algorithms include basic algorithms such as threshold detection, peak detection, and correlation detection [7]. In addition, for multi-target detection and tracking requirements, higher-order signal processing algorithms such as cluster analysis and Kalman filtering are also required.

3. LiDAR technology applications

3.1. Autonomous driving field

As a benchmark field for the integration of artificial intelligence and the automotive industry, autonomous driving technology relies heavily on multi-sensor fusion solutions for its technological breakthroughs. Among them, LiDAR, with its centimeter-level ranging accuracy and 3D environment modeling capabilities, has become an indispensable core sensor in future L4/L5 autonomous driving systems. By integrating multi-line LiDAR, autonomous vehicles can build high-precision 3D environment models in real time, enabling functions such as obstacle detection, path planning, and decision control [8].

The perception of the surrounding environment can be achieved using LiDAR. The accuracy of environmental perception is affected by the number of lines, scanning frequency, wavelength, and algorithm processing capabilities of the LiDAR. High-line LiDAR can provide denser point cloud data, improving the detail and accuracy of environmental perception. Currently, L3 and autonomous driving systems are typically equipped with LiDAR. LiDAR-based environmental perception has the advantages of high resolution, high accuracy, and strong anti-interference ability, but it is costly, and its performance degrades in adverse weather conditions (such as heavy rain, heavy snow, and dense fog).

Although LiDAR has significant advantages in accuracy, its cost is high, accounting for nearly half of the entire perception system's cost. Breakthroughs in solid-state LiDAR technology will greatly reduce the cost of LiDAR, facilitating the popularization of LiDAR applications. How to reduce costs while ensuring accuracy, and improve the stability and reliability of LiDAR in harsh weather conditions, are the main challenges currently faced.

Autonomous driving technology is a hot research topic nowadays, and vehicle localization is a core problem that must be solved in autonomous driving technology. A fast and accurate localization system can not only effectively help vehicles achieve autonomous driving functions, but also significantly improve vehicle safety. High-definition maps are the technical guarantee for autonomous vehicles to achieve precise positioning and route navigation, and LiDAR is a crucial sensing device for building high-definition maps [9].

Compared to traditional vehicle navigation systems, modern vehicles with autonomous driving capabilities primarily rely on the matching accuracy of LiDAR and high-precision maps, as well as the performance of SLAM algorithms, for positioning and navigation accuracy. Using LiDAR for positioning and navigation has the advantages of high precision and real-time performance, but it relies on pre-built high-precision maps, and whether the maps can be updated accurately and in a timely manner in large-scale dynamic environments is a challenge.

3.2. Robotics navigation field

Lidar also has important applications in the field of robot navigation and is one of the key technologies for realizing autonomous robot movement and obstacle avoidance [10].

The robot's dynamic obstacle avoidance capability is affected by the scanning speed of the LiDAR, data processing capability, and obstacle avoidance algorithm. Fast scanning and efficient data processing can real-time perceive obstacles and make obstacle avoidance decisions [11]. Using LiDAR for dynamic obstacle avoidance has the advantages of real-time performance and accuracy, but false detections and missed detections may occur in complex environments [12,13]. How to improve the obstacle avoidance accuracy and robustness of LiDAR in complex environments, and how to combine other sensor data to achieve multi-sensor fusion obstacle avoidance, are current research hotspots.

The field of robot navigation also requires the support of SLAM technology, the performance of which is affected by the point cloud quality of LiDAR, algorithm complexity, and computing resources [14]. High-quality point cloud data and efficient algorithms can improve the accuracy and real-time performance of SLAM [15]. SLAM technology using LiDAR can realize the autonomous localization of robots and the map construction of the surrounding environment, but it requires a large amount of calculation and high hardware resources. How to optimize the SLAM algorithm to reduce the amount of calculation, and how to implement efficient SLAM on resource-constrained robot platforms are the main challenges currently facing.

3.3. Environmental monitoring field

In the field of environmental monitoring, LiDAR, with its significant advantages of strong beam directivity and high energy density, can achieve high-precision detection of atmospheric parameters, temperature and humidity, terrain and geomorphology, and other parameters [16,17].

However, in actual atmospheric environment monitoring, its measurement accuracy will be affected by multiple factors such as wavelength, emission power, and real-time atmospheric conditions.

When using LiDAR for atmospheric environment monitoring, its accuracy is affected by the wavelength and emitted power of the LiDAR, as well as atmospheric conditions. Appropriate wavelengths and emitted power can enhance the monitoring range and sensitivity. The application of LiDAR atmospheric environment monitoring technology has the advantages of high precision and high spatiotemporal resolution, but its equipment cost is high and requires professional operation and maintenance. These factors limit its large-scale application. Reducing equipment costs and improving the stability and reliability of LiDAR under different atmospheric conditions are current problems that need to be solved.

The accuracy of terrain and geomorphological monitoring using LiDAR is affected by the scanning method of the LiDAR, point cloud density, and data processing algorithms. The selection of airborne or terrestrial scanning methods, as well as efficient point cloud processing algorithms, can improve the accuracy and efficiency of monitoring. In atmospheric detection, the terrestrial scanning method utilizes the frequency of atmospheric backscattered light to measure temperature and wind speed in the direction of the light beam. Airborne LiDAR uses the differential absorption method to measure surface reflectance at different laser wavelengths to determine the mixing ratio of various gases in the atmosphere [2]. In terrain detection, LiDAR terrain and geomorphological monitoring have the advantages of high precision and rapid imaging, but are greatly affected by terrain complexity and vegetation cover.

How to improve the monitoring accuracy of LiDAR in complex terrain and vegetation-covered areas, and how to achieve automated processing and analysis of monitoring data, are current research directions.

4. Conclusion

As an advanced active remote sensing technology, LiDAR plays an important role in autonomous driving, robot navigation, and environmental monitoring. With the continuous advancement of LiDAR technology and the expansion of application scenarios, LiDAR technology is developing towards higher precision, lower cost, and smaller size. At present, significant achievements have been made in LiDAR technology research both domestically and internationally, but there is still room for improvement in the accuracy and stability of domestic LiDAR.

The future development trends of LiDAR technology will be as follows: First, high precision, which improves ranging accuracy and measurement resolution by optimizing laser emission and reception technology, signal processing, and algorithms; Second, low cost, which reduces the cost of LiDAR through large-scale production and technological innovation, promoting its market application; Third, miniaturization and integration, which achieves miniaturization and integrated design of LiDAR through micro-electro-mechanical systems and other technologies to meet the needs of more application scenarios. With the rapid development of autonomous driving, robot navigation, environmental monitoring, and other fields, the application prospects of LiDAR technology will be broader. In the field of autonomous driving, LiDAR will become one of the key sensors for achieving L4/L5 level autonomous driving; In the field of robot navigation, LiDAR will promote the intelligent upgrade of service robots, industrial robots, and other products; In the field of environmental monitoring, LiDAR will play a greater role in atmospheric environmental monitoring, topographic and geomorphic monitoring, and other aspects.

In conclusion, LiDAR technology is in a phase of rapid development, and its broad application prospects and continuous technological innovation will inject new vitality into the development of related fields. In the future, with the continuous maturity of technology and the expansion of application scenarios, LiDAR technology will play an important role in more fields, promoting the intelligentization and automation of related industries.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Shi Z, 2021, Research on Autonomous Navigation of Mobile Robot Based on Fusion of LiDAR and Binocular Vision, Master's thesis, Nanjing University of Science and Technology.

[2] Li X, Guo P, Zang C, et al., 2019, Research Status and Application of LiDAR Technology. Automotive Electric Appliance, (05): 4–6.

[3] Ma Z, 2022, Research and Design of TOF Readout Circuit for Single-Photon LiDAR Based on DLL-TDC, Master's thesis, Nanjing University of Posts and Telecommunications.

[4] Zhao J, Zhao Y, Tong J, et al., 2021, A Design of Multi-Channel Time-to-Digital Converter for TOF LiDAR. Semiconductor Optoelectronics, 42(03): 342–347.

[5] Anon., 2023, Chinese Scientists Have Overcome the Worldwide Challenges of Anti-Interference and High-Precision Parallel Detection for LiDAR. Electronic Quality, (03): 113.

[6] Li J, 2018, Study on the Modeling and Processing Algorithm of the Full Waveform of Pulsed Lidar, Master's thesis, Tianjin University.

[7] Chen Y, 2021, Research on FMCW Lidar Ranging and Velocity Signal Processing Algorithm, Master's thesis, Harbin Institute of Technology.

[8] Zhang S, Huang K, Gan L, et al., 2025, Research on Mapping Method Based on Autonomous Driving SLAM Technology. Special Purpose Vehicle, (02): 41–43.

[9] Yu R, 2023, Research on SLAM Algorithm Based on LiDAR, Master's thesis, Jilin University.

[10] Yin S, 2023, Research on Dynamic Obstacle Avoidance Algorithm Based on LiDAR, Master's thesis, Hebei University of Architecture and Engineering.

[11] Li W, 2019, Strategy of Intelligent Vehicle Path Following and Obstacle Avoidance Based on GPS and LiDAR, Master's thesis, Shandong University of Technology.

[12] Tang J, 2021, Research on Detection of 3D LiDAR by Using a Multi-Model Fusion Method under Challenging Conditions, Master's thesis, Hunan University.

[13] Guo Z, 2021, Development of SLAM and Dynamic Obstacle Avoidance Robot Based on LiDAR, Master's thesis, Heilongjiang University.

[14] Liu C, 2020, Research and Implementation of Robot Localization and Navigation Algorithm Based on SLAM, Master's thesis, Beijing University of Posts and Telecommunications.

[15] Guo X, 2022, Research on SLAM Map Construction Algorithm Based on LiDAR, Master's thesis, Shenyang

University of Technology.

[16] Liu Y, 2022, Application of LiDAR in Atmospheric Environment Monitoring and Meteorological Research. Cleaning World, 38(02): 67–68.

[17] Zhang Q, 2020, Application of LiDAR in Atmospheric Environment Monitoring and Meteorological Research. Science and Technology Information, 18(05): 19–20.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Research on Teaching Reform Strategies of Python Programming Course Based on Artificial Intelligence Technology

Hong Chen*

Changji Vocational and Technical College, Changji 831100, Xinjiang, China

**Author to whom correspondence should be addressed.*

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: As one of the core courses for computer-related majors, the Python programming course has become increasingly important in the era of artificial intelligence. It aims to help students develop good computer thinking and improve their abilities in programming and data analysis. The application of artificial intelligence technology in the teaching of Python programming courses is of great significance for optimizing the allocation of teaching resources, enriching students' learning experience, and significantly improving teaching quality. Based on this, this paper first briefly expounds on the importance of applying artificial intelligence technology in the teaching of Python programming courses. On this basis, it focuses on exploring effective strategies for the teaching reform of Python programming courses based on artificial intelligence technology, hoping to provide new ideas for the teaching of Python programming courses and contribute to cultivating more Python programming talents with artificial intelligence literacy.

Keywords: Artificial intelligence technology; Python programming course; Teaching reform; Effective strategies

Online publication: October 15, 2025

1. Introduction

In the era of artificial intelligence (AI), AI technology has become a key driver for promoting reform and innovation in education and teaching. For Python programming courses, traditional teaching models have the following drawbacks: First, teachers cannot accurately and comprehensively grasp each student's learning situation to provide personalized guidance. Second, teaching content fails to keep pace with the rapid development of computer technology, leading to a potential disconnect between what students learn and practical applications. The application of AI technology, however, not only helps teachers accurately identify students' individual needs but also facilitates dynamic adjustments to teaching content, ultimately significantly improving teaching effectiveness. Thus, the research in this paper holds significant practical relevance.

2. The importance of applying AI technology in Python programming course teaching

2.1. Facilitating optimal resource allocation and personalized learning experiences

In traditional teaching models, teachers often rely on textbooks and teaching aids to prepare lessons and design teaching processes, mostly adopting a "one-size-fits-all" approach that struggles to meet students' diverse and personalized learning needs. Students remain in a passive position of receiving knowledge, resulting in low learning quality [1]. The application of AI technology, however, helps optimize resource allocation while providing students with personalized learning experiences—achieving two goals at once. Specifically, first, AI-driven learning platforms enable teachers to offer personalized learning resources and tutoring plans to students at different levels, realizing intelligent resource allocation and recommendation to improve resource utilization efficiency [2,3]. Second, AI technology can intelligently analyze students' learning behaviors, homework completion rates, and exam scores to generate learner profiles. Based on these profiles, intelligent platforms can develop personalized learning paths, accurately identify weak areas, maximize learning effectiveness, and ultimately achieve personalized teaching.

2.2. Helping teachers identify teaching issues and adjust strategies in a timely manner

Python programming courses place great emphasis on monitoring students' real-time progress, but traditional teaching models may struggle to achieve this comprehensively. AI-based intelligent tools, such as smart learning platforms and classroom interaction systems, can promptly capture key indicators like students' response speed, accuracy, and review frequency, thereby truly reflecting their mastery of knowledge points, learning progress, and issues in the teaching process. This allows teachers to accurately grasp students' learning status and enhance the timeliness and relevance of feedback. For common problems faced by most students, teachers can provide centralized explanations; for individual issues, they can offer one-on-one targeted tutoring—greatly improving the relevance and effectiveness of teaching [4,5].

2.3. Stimulating students' interest in learning Python programming

For beginners, learning Python programming can be challenging. When faced with tedious grammar rules, abstract logical concepts, and monotonous code writing, some students may develop fear or resistance, affecting their final learning outcomes [6]. The integration of AI technology, however, can provide students with real-world cases and vivid projects for practice. It can also leverage AR, VR, and other technologies to create immersive programming environments, fully stimulating their desire to learn and explore—yielding twice the result with half the effort. For example, students can write and debug Python code in virtual scenarios and obtain visual results, transforming tedious code learning into an engaging exploration journey and infusing their learning path with sustained vitality [7,8].

3. Effective strategies for teaching reform in Python programming courses based on artificial intelligence technology

3.1. Enriching teaching content

Enriching teaching content requires teachers to promptly integrate AI cases into Python programming courses. Meanwhile, attention should be paid to the organic combination of compulsory and elective content. This ensures that teaching content remains aligned with the times, laying a solid foundation for students' future careers. First, teachers should timely incorporate AI cases into Python programming teaching to fully stimulate students' interest in learning and help them gain a deeper and more comprehensive understanding of AI. Python boasts rich AI libraries such as NumPy, Pandas, Matplotlib, Scikit-learn, TensorFlow, and PyTorch, which provide strong support for AI development, covering data processing, visual presentation, and algorithm implementation [9]. Therefore, in actual teaching, teachers can flexibly introduce AI cases based on these AI libraries. For example, using the Scikit-learn library to complete iris classification or implementing handwritten digit recognition through the TensorFlow library. These two cases help guide students to gradually explore the fields of machine learning and deep learning, enabling them to establish an overall understanding of core concepts, basic processes, and typical algorithms in AI. This, in turn, fully stimulates students' interest in learning and exploration, guiding them to transition from superficial learning to deep learning. Second, emphasis should be placed on organically combining compulsory and elective content to build a hierarchical and systematic learning system that meets students' personalized learning needs. In addition to compulsory content such as basic Python syntax, web crawlers, data analysis and processing, and data visualization, teachers should also push personalized learning resources to students through intelligent learning platforms, such as game development, network programming, natural language processing, and computer vision [10,11]. Students can freely choose learning content based on their personal interests, ensuring the cutting-edge nature of teaching content while providing solid support for their future career development. Teachers can provide students with diverse learning resources such as videos, tutorials, literature, software, and code examples through online learning platforms. They can also introduce real enterprise project tasks into the classroom to deepen personalized talent cultivation.

3.2. Optimizing teaching processes

To enrich students' learning experience and enhance the effectiveness of classroom teaching, teachers can actively adopt a diversified blended teaching model that integrates "in-class and after-class learning, online and offline engagement." This model divides the classroom into three stages—pre-class, in-class, and post-class—to optimize and restructure the teaching process. In the pre-class preview stage, teachers can release a list of preview tasks in advance. Students need to complete pre-class tasks according to the list, such as watching learning videos on the Classroom platform, finishing thinking questions and test questions, and forming a preliminary understanding of key and difficult knowledge through independent learning. During this process, teachers can use online teaching platforms to monitor students' preview progress in real time, accurately identify their weak points, and adjust subsequent teaching strategies accordingly. Taking the aforementioned "iris classification case" as an example, teachers can guide students to watch videos related to machine learning, require them to install Anaconda3 and learn about the Scikit-learn library, and set test questions and thinking questions, such as: What are the concepts of machine learning and supervised learning? What is the relationship between machine learning, artificial intelligence, and deep learning? At the same time, teachers can ask students to list specific scenarios of supervised learning in daily life and explain the reasons [12]. In the in-class learning stage, teachers should focus on explaining and practicing common problems among students and vividly demonstrating teaching cases. During this process, teachers should find appropriate opportunities to integrate ideological and political education content to cultivate more programming talents with both professional knowledge and excellent ideological qualities. Furthermore, teachers should flexibly use AI assistants to support students' learning and practice, giving full play to the enabling role of AI technology in Python programming teaching. For example, teachers can use tools like ChatGPT, Doubao, and DeepSeek to generate questions for students to answer, or encourage students to use tools like CodeGeek to compare the semantic similarity between their code and code generated by ChatGPT, preventing code plagiarism and educating students on professional ethics [13]. In the post-class consolidation stage, teachers can assign homework based on teaching content and students' overall preview and learning performance, and use AI technology for intelligent homework correction. Students can review knowledge points and practice programming skills through intelligent teaching platforms to achieve timely consolidation and internalization of knowledge and skills, significantly improving learning quality.

3.3. Reforming teaching methods

Given the differences in students' programming foundations and abilities, as well as the various problems they may encounter in programming practice, teachers can flexibly use generative AI to empower heuristic teaching, allowing students to become the masters of the classroom, meeting their personalized learning needs, and promoting the transformation of Python programming teaching toward digitalization and personalization. Specifically, teachers can use the AI assistant on the Classroom platform to guide students' learning and exploration [14,15]. This platform's AI assistant offers multiple functions, such as heuristic question answering, heuristic knowledge point learning, and code analysis, helping students solve problems in a timely manner and significantly improving their learning efficiency. The AI assistant can actively generate programming exercises for students to practice. If students encounter difficulties during practice, they can seek help from the AI assistant, which will then automatically enter the heuristic answering interface and generate prompts for reference, effectively inspiring students. Instead of directly providing answers, the AI assistant guides students step by step to establish problem-solving logic, which helps cultivate good thinking habits [16]. In addition to the above methods, teachers can adopt project-driven learning by introducing real enterprise programming cases into the classroom and encouraging students to complete projects in groups. This cultivates students' teamwork spirit and improves their comprehensive programming abilities. After completing tasks, teachers should guide students to participate in interactive discussions and share their experiences, including new insights during programming, problems encountered, and solutions. This deepens students' understanding of knowledge and mastery of skills, creates a positive learning atmosphere, and promotes the collaborative development of teaching and learning [17].

4. Conclusion

In conclusion, the application of artificial intelligence technology in the teaching of Python programming courses is gradually reshaping the classroom teaching ecology, injecting strong vitality and impetus into the reform and innovation of curriculum teaching, and helping to comprehensively improve the teaching and learning effects. In the future, teachers should continue to deepen the research on the application of artificial intelligence technology in the teaching of Python programming courses, properly handle the relationship between technical teaching and manual teaching, continuously optimize teaching methods, and contribute to the cultivation of more high-quality talents who firmly master the professional knowledge and skills of Python programming.

Disclosure statement

The author declares no conflict of interest.

References

[1] Su T, 2024, Analysis of Teaching Reform of Python Programming Course for Artificial Intelligence. Office Automation, 29(9): 47–49.

[2] Min Y, Xu C, 2022, Teaching Reform and Practice of "Python Programming" Course in the Era of Artificial Intelligence. Science and Technology Guide - Electronic Edition (Second Half), (12): 187–189.

[3] Liang S, Du R, Hang W, 2023, Exploration on the Practice of Intelligent Education of Python Programming from the Perspective of Artificial Intelligence. Software, 44(2): 42–45, 65.

[4] Li L, Li D, Yang H, et al., 2023, Curriculum Reform of "Python Programming" under the Background of Artificial Intelligence. Modern Information Technology, 7(17): 178–182, 188.

[5] Sheng G, Zhao M, Liu Z, et al., 2022, Exploration on the Curriculum Reform of "Python Language Programming" under the Background of Artificial Intelligence. Technology Wind, (30): 83–85.

[6] Li F, 2019, Preliminary Exploration on the Construction of Python Language Programming Course under the Background of Artificial Intelligence. Education Informatization Forum, (10): 27–28.

[7] Wang Y, Wang Y, Zheng Y, 2024, Development of a New Form of Python Programming Textbook Based on Artificial Intelligence and Multimedia. Knowledge Economy, 697(33): 214–216.

[8] Yang H, Wu G, Tian Z, et al., 2024, Analysis of Problems and Countermeasures in Python Programming Teaching. Industrial and Technological Forum, 23(17): 145–147.

[9] Song S, 2023, Teaching Practice of Python Programming Course. Application of Integrated Circuits, 40(9): 150–151.

[10] Wang Y, Fan D, Jiang J, 2022, Teaching Reform and Practice of Python Programming Course. Technology Wind, (34): 95–97.

[11] Wang Y, Liu W, Wang K, et al., 2025, Development and Application of a Teaching Tool for Python Language Programming Based on Knowledge Graph. China Information Technology Education, (1): 99–102.

[12] Chen J, 2018, Teaching Research on Python Programming Course. Southern Agricultural Machinery, 49(24): 9.

[13] Hao G, Yang J, Wu D, et al., 2019, Design and Implementation of an Intelligent Q&A System for the "Python Programming" Course. Computer Knowledge and Technology, 15(7): 86–88.

[14] Luo J, 2024, Reform of Higher Vocational Python Programming Course Based on Intelligent Robot Design with AI Assistance. Great Wuhan, (19): 129–130.

[15] Zeng X, Wang M, Li G, et al., 2024, Discussion on Problem-Solving-Oriented Python Programming Teaching Method. Computer Knowledge and Technology, 20(25): 117–119, 122.

[16] Lei L, Li P, 2025, Innovation of Teaching Method for Python Programming Language Based on Reinforcement Learning. Computer Education, (1): 110–115.

[17] Duan Y, Xu C, Wei L, et al., 2021, Exploration on Teaching Method of Python Language Programming Course. Technology Wind, (17): 40–42.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Correlation Analysis Between Investor Sentiment and Stock Price Fluctuations Based on Large Language Models

Guohua Ren[1], **Ziyu Luo**[1]*, **Naiwen Zhang**[2], **Yichen Yang**[1]

[1]Macau University of Science and Technology, Macau 999078, China
[2]Macao Polytechnic University, Macau 999078, China

***Author to whom correspondence should be addressed.*

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: The efficient market hypothesis in traditional financial theory struggles to explain the short-term irrational fluctuations in the A-share market, where investor sentiment fluctuations often serve as the core driver of abnormal stock price movements. Traditional sentiment measurement methods suffer from limitations such as lag, high misjudgment rates, and the inability to distinguish confounding factors. To more accurately explore the dynamic correlation between investor sentiment and stock price fluctuations, this paper proposes a sentiment analysis framework based on large language models (LLMs). By constructing continuous sentiment scoring factors and integrating them with a long short-term memory (LSTM) deep learning model, we analyze the correlation between investor sentiment and stock price fluctuations. Empirical results indicate that sentiment factors based on large language models can generate an annualized excess return of 9.3% in the CSI 500 index domain. The LSTM stock price prediction model incorporating sentiment features achieves a mean absolute percentage error (MAPE) as low as 2.72%, significantly outperforming traditional models. Through this analysis, we aim to provide quantitative references for optimizing investment decisions and preventing market risks.

Keywords: Large language model; Investor sentiment; Stock return prediction; Sentiment analysis; LSTM

Online publication: October 15, 2025

1. Introduction

In financial markets, investor sentiment has long been regarded as one of the key factors influencing stock price fluctuations [1]. Traditional methods of analyzing investor sentiment often rely on surveys, market indicators, etc., and suffer from various limitations [2]. For instance, survey methods are affected by sample size and respondents' subjective biases, and the low frequency of data updates makes it difficult to reflect real-time changes in market sentiment. The market indicator substitution method tends to confuse "sentiment-driven" factors with "fundamentals-driven" factors, while traditional text analysis methods exhibit poor adaptability to financial terminology. With the rapid development of artificial intelligence technology, large language models (LLMs) have provided a novel technical approach for analyzing investor sentiment. LLMs can deeply understand the semantic information within massive amounts of unstructured textual data, accurately capturing the emotional tendencies embedded therein, and thus offer a more efficient and accurate analytical tool for studying the correlation between investor sentiment and stock price movements. In recent years, notable progress has been made in applying LLMs, exemplified by ChatGPT, in the financial sector [2]. Relevant studies have shown that sentiment scoring factors constructed based on LLMs can effectively predict stock returns, achieving an annualized excess return of 9.3% in the CSI 500 Index domain. Meanwhile, combining sentiment analysis from LLMs with deep learning prediction models (such as LSTM) can reduce the mean absolute percentage error (MAPE) of stock price forecasts to 2.72%, significantly outperforming traditional statistical models. These research findings fully demonstrate the immense potential of LLMs in investor sentiment analysis and stock market forecasting.

The innovation of this paper lies in its systematic exposition of the technical principles underlying LLMs in sentiment analysis and its comprehensive use of factor backtesting and machine learning prediction models to provide multi-dimensional empirical evidence from both the A-share market and overseas markets regarding the correlation between investor sentiment and stock price movements. The following sections will first introduce the technical principles of LLMs and sentiment quantification methods, followed by an empirical analysis, and conclude with a summary of the research findings.

2. Technical principles of large language models in investor sentiment analysis

2.1. Core architecture and advantages of large language models

Large language models, represented by ChatGPT, employ the Transformer neural network architecture. Through training on massive text datasets, they possess powerful capabilities in natural language understanding and generation [3]. The core training process of these models consists of three key steps: supervised learning, reward model training, and proximal policy optimization-based reinforcement learning. This training process enables large language models to deeply comprehend semantic information in texts, accurately capture the emotional tendencies embedded within, and overcome the limitations of traditional sentiment analysis models.

Compared to traditional sentiment analysis models (such as BERT), large language models offer the following significant advantages in investor sentiment analysis:

(1) Continuity of emotional scores: Large language models can provide continuous emotional scores ranging from 1 to 10, rather than the discrete classifications (positive, negative, neutral) used by traditional models. This allows for a more nuanced reflection of the differences in emotional intensity in texts such as analyst research reports and news articles. For example, when analyzing two research reports both containing the phrase “performance growth,” large language models can assign different emotional scores based on contextual information such as the magnitude of growth and industry prospects, whereas traditional models might categorize both as the same emotional category.

(2) Contextual semantic understanding: Large language models can accurately comprehend the contextual semantics within texts, avoiding misjudgments caused by keyword matching in traditional models. For instance, in a text like “The company’s performance this quarter exceeded expectations, but increased industry competition will put pressure on future growth,” a traditional model might misjudge it as positive due to the keyword “exceeded expectations.” In contrast, a large language model can consider the

negative information in the latter part of the sentence and provide a more reasonable emotional score [4].

(3) Multilingual and multi-domain adaptability: Traditional models (such as BERT) are often constrained by English training corpora, leading to issues of information loss during translation when processing Chinese financial texts. Large language models, through extensive multilingual training data, can better cater to the analytical needs of Chinese financial texts, significantly improving the accuracy of sentiment recognition in areas such as analyst research reports and financial news [5].

2.2. Quantitative method for investor sentiment based on large language models

In practical applications, the quantification of investor sentiment based on large language models is primarily achieved through the following steps:

(1) Data collection and preprocessing: Unstructured textual data from the financial market is collected, including analyst research reports, financial news, social media comments, company announcements, etc. Taking the "A-Share Investor Sentiment Survey" conducted by the Applied Statistics Research Center at Shanghai University of Finance and Economics as an example, its data sample encompasses 250,000 analyst research report titles from 2016 to 2023. During the preprocessing stage, textual data needs to be cleaned, tokenized, and other operations performed to prepare for subsequent sentiment analysis.

(2) Sentiment scoring with large language models: Sentiment scores are assigned to the preprocessed textual data by calling large language models (such as GPT-3.5-turbo) through Python APIs. To ensure the stability and rationality of the scores, appropriate model parameters are typically set, such as setting the Temperature parameter to 0.5 to balance the randomness and consistency of the scores. Meanwhile, batch requests are employed to enhance processing efficiency and control API call costs.

(3) Sentiment factor construction: Based on the sentiment scoring results from large language models, sentiment factors suitable for stock market analysis are constructed. Common sentiment factors include equal-weighted average factors, exponentially weighted factors, and score volatility factors, etc. These factors can reflect the overall level, temporal trends, and volatility of investor sentiment from different dimensions, providing quantitative indicators for studying the correlation between sentiment and stock price movements [6].

(4) Validation of the effectiveness of sentiment factors: Through backtesting analysis, we verify the predictive power of sentiment factors on stock returns. For example, in the overall A-share market, the monthly information coefficient (IC) stability (ICIR) of the equal-weighted average factor is 1.6, with a monthly win rate exceeding 75% and an annualized excess return of approximately 5.2% for the long position. These indicators demonstrate that sentiment factors based on large language models possess significant stock-selection capabilities and can effectively predict the rise and fall of stock prices.

3. Empirical analysis on the correlation between investor sentiment and stock price fluctuations

3.1. Stock return prediction based on sentiment factors from large language models

To verify the correlation between investor sentiment and stock price fluctuations, scholars and research institutions both domestically and abroad have conducted extensive empirical studies based on sentiment factors constructed using large language models. Among them, the "Collision of ChatGPT and Research Report Text Sentiment: Quantitative Research Series Report Eleven" released by BigQuantAI Quantification provides rich empirical data, serving as an important reference for our analysis of the correlation between the two.

3.1.1. Performance of returns in the "GPT Exceeds Expectations" sample pool

This report uses a large language model to conduct sentiment scoring on the titles of analyst research reports from 2017 to 2023. Samples with scores ≥ 9 are defined as "GPT Exceeds Expectations" samples (representing strong optimism), and a corresponding stock sample pool is constructed (with monthly rebalancing, holding the top 20% of stocks by score). Meanwhile, samples identified as "Performance Exceeds Expectations" through traditional text analysis (based on keyword matching) are used as a control group (traditional "Text Exceeds Expectations" sample pool).

The backtesting results (see **Table 1**) reveal that from 2017 to 2023, the annualized return of the "GPT Exceeds Expectations" sample pool reached 8.0%, significantly surpassing the 7.8% achieved by the traditional "Text Exceeds Expectations" sample pool. A t-test was conducted to assess the significance of the return difference between the two groups, yielding a t-statistic of 2.31 and a P-value < 0.05. This indicates that the return advantage of the "GPT Exceeds Expectations" sample pool is statistically significant and not a random occurrence.

From the perspective of risk-adjusted metrics, the "GPT Exceeds Expectations" sample pool outperformed the traditional sample pool in terms of rebalancing win rate, information ratio, and maximum drawdown. This suggests that sentiment analysis based on large language models can more accurately identify stocks with the potential for excess returns, further corroborating the positive correlation between investor sentiment and stock price movements [7].

Table 1. Comparison of returns between the "GPT Exceeds Expectations" and traditional "Text Exceeds Expectations" sample pools (2017–2023)

Metric	"GPT Exceeds Expectations" sample pool	Traditional "Text Exceeds Expectations" sample pool	*t*-value	*P*-value
Annualized return (%)	8.0	7.8	2.31	< 0.05
Portfolio adjustment win rate (%)	68.3	62.1	-	-
Information ratio	1.8	1.2	-	-
Maximum drawdown (%)	-18.5	-25.3	-	-

3.1.2. Performance of sentiment factors across different index domains

To further evaluate the applicability of sentiment factors, their performance across different index domains was also examined. The results indicate that within the CSI 500 index domain, the coverage of sentiment factors based on large language models reached 70%, with the RankIC increasing to 3.6% and the annualized excess return of the long position reaching as high as 9.3%. Moreover, positive excess returns were achieved in each of the past three years (2021–2023), with values of 8.9%, 9.5%, and 9.1%, respectively. A t-statistic test was conducted to assess the significance of the returns generated by sentiment factors, resulting in a t-statistic of 3.12 and a P-value < 0.01, indicating a high level of statistical significance for these excess returns.

These findings suggest that in the mid-cap blue-chip stock market, investor sentiment exerts a more pronounced impact on stock price movements, and sentiment factors based on large language models demonstrate stronger stock-picking capabilities. In contrast, within the overall A-share market, the annualized excess return of sentiment factors was 5.2%, slightly lower than the performance observed in the CSI 500 index domain. This discrepancy may be attributed to the vast number of stocks and diverse industry distribution in the overall A-share market, leading to lower consistency in investor sentiment and a relatively dispersed impact of sentiment on stock price movements [8].

3.2. Stock price prediction model incorporating sentiment analysis

In addition to the construction and application of sentiment factors, researchers have also combined sentiment analysis from large language models with deep learning prediction models to further validate the correlation between investor sentiment and stock price fluctuations. A research paper published on CSDN Blog proposes a stock price prediction framework based on LSTM (long short-term memory network), which integrates historical trading data with sentiment scores generated by large language models, significantly enhancing the accuracy of stock price predictions.

3.2.1. Model construction and data sources

This study focuses on four technology companies listed on NASDAQ (Apple, Google, Microsoft, and Amazon), with data sourced from Yahoo Finance, covering daily trading data (opening price, highest price, lowest price, closing price, trading volume) from April 2024 to April 2025. Meanwhile, relevant financial news was collected through Bloomberg and Reuters, and sentiment scores were generated using a fusion approach of "VADER tool + fine-tuning of large language models": During the pre-training phase, a large language model (GPT-3.5-turbo) was fine-tuned using financial domain texts (100,000 financial news articles, 50,000 research reports) to optimize its understanding of financial jargon (e.g., "AI chip shipments," "cloud computing ARPU"). In the feature input phase, the sentiment scores (ranging from 1 to 10) output by the fine-tuned large language model were standardized to a range of -1 to +1 and then weighted and fused with the scores from the VADER tool (ranging from -1 to +1), with weights of 0.7 and 0.3, respectively, to generate a composite sentiment score as a quantitative indicator of investor sentiment.

In terms of model construction, this study employed a two-layer LSTM architecture: the first layer comprises 64 memory units, with return_sequences=True set to retain time series information, and incorporates a 20% dropout layer to prevent overfitting; the second layer contains 32 memory units, further refining data features; the output layer is a single linearly activated neuron, designed to predict the closing price on the 61st day. The model utilizes the Adam optimizer and Mean Squared Error (MSE) as the loss function, trained for 100 epochs with a batch size of 32.

3.2.2. Prediction results and analysis

Backtesting results demonstrate that the model achieved an average absolute percentage error (MAPE) of 2.72% on unseen test data, significantly outperforming the traditional ARIMA model (which had a MAPE of 20.66%). Among them, Google's stock exhibited the lowest MAPE, at just 2.65%. More importantly, sensitivity analysis reveals that when sentiment features are not utilized, the model's MAPE rises to 3.15%, indicating that sentiment analysis contributes approximately an 8–12% increment to the model's predictive accuracy. This result fully underscores the close correlation between investor sentiment and stock price fluctuations, demonstrating that incorporating sentiment factors into stock price prediction models can significantly enhance prediction accuracy.

3.3. Investor sentiment and stock market volatility amidst macro events

Macroeconomic events often exert a significant impact on investor sentiment, thereby triggering sharp fluctuations in the stock market [9]. The December 2024 Federal Reserve interest rate meeting serves as a quintessential example, where the policy signals released significantly altered investor sentiment, leading to substantial adjustments in global stock markets. This scenario provides a real-world context for analyzing the correlation between investor sentiment and stock price fluctuations. By selecting a three-day window before and after the event (December 16, 2024 to December 22, 2024) as the event window, we can enhance the quantitative support of our analysis by calculating the correlation coefficients between sentiment indicators and stock index returns and conducting Granger causality tests.

3.3.1. FOMC meeting and changes in investor sentiment

On December 19, 2024, the Federal Reserve released the minutes of its December interest rate-setting meeting, deciding to cut interest rates by 25 basis points (bp) to a range of 4.25–4.50%. However, the Federal Reserve simultaneously signaled a slower pace of rate cuts, with the dot plot indicating a reduction in the number of rate cuts for 2025–2026 from the previously anticipated five to three, while the long-term interest rate central tendency continued to rise by 50 bp. Additionally, Federal Reserve Chair Jerome Powell emphasized at the press conference that "future rate cuts will require the hard condition of sustained inflation falling to 2%," underscoring a more cautious approach to future rate reductions.

This policy signal starkly contrasted with prior market optimism, causing investor sentiment to abruptly shift toward caution. The U.S. stock market's "fear gauge," the VIX, surged 74% in a single day following the announcement, reaching a four-month high of 28.32—a larger daily spike than during the August "Black Monday." To verify the abruptness of the sentiment shift, a t-test compared the mean VIX values inside and outside the event window: the average VIX during the event window was 24.15, compared to a 15.82 average over the 30 trading days prior. The t-statistic was 5.73 (P-value < 0.001), indicating statistically significant emotional volatility triggered by the event.

3.3.2. Impact of changes in investor sentiment on the stock market

The sharp deterioration in investor sentiment directly led to a significant decline in global stock markets [10]. In the U.S. stock market, as of the close on December 18, 2024, the Dow Jones Industrial Average fell by 1,123.03 points, or 2.58%, to 42,326.87; the S&P 500 index dropped by 178.45 points, or 2.95%, to 5,872.16; and the Nasdaq Composite Index declined by 716.37 points, or 3.56%, to 19,392.69. Technology stocks, as a sensitive sector to market sentiment, were hit even harder, consistent with the earlier conclusion from the LSTM model that technology stock prices are significantly influenced by sentiment.

In the Asia-Pacific market, on December 19, 2024, the Korea Composite Stock Price Index (KOSPI) fell by 1.95%, while the Nikkei 225 Index dropped by 0.69%. Although the A-share market was affected to some extent, with the Shanghai Composite Index opening significantly lower by nearly 1%, it maintained a weak and volatile pattern throughout the day, with the closing loss narrowing to 0.36%, demonstrating strong resilience. This discrepancy may stem from significant differences in the investor structure and policy environment of the A-share market compared to overseas markets, leading to varying degrees of influence of investor sentiment on stock price fluctuations. For instance, a survey conducted by the Research Center for Applied Statistics at Shanghai University of Finance and Economics revealed that in the fourth quarter of 2024, the confidence indices for institutional and retail investors in the A-share market stood at 118 points and 101.42 points, respectively, both returning to the optimistic range. The relative stability of domestic investor sentiment, to some extent, offset the impact of overseas market sentiment volatility on A-shares. By calculating the Pearson correlation coefficient between the VIX Index and the returns of various stock indices within the event window, it was found that sentiment indicators were significantly negatively correlated with stock index returns, with a stronger correlation observed in overseas markets. Meanwhile, the Granger causality test results indicated that changes in investor sentiment triggered by macro events were a significant driver of stock price fluctuations, further validating the close relationship between the two.

4. Conclusion

Through an analysis of the correlation between investor sentiment based on large language models and stock price fluctuations, this paper draws the following conclusions: On the one hand, large language models, with their powerful natural language understanding capabilities, exhibit significant advantages in investor sentiment analysis. Compared to traditional sentiment analysis models, large language models can provide continuous sentiment scores, accurately understand contextual semantics, and better adapt to Chinese financial texts, offering a more efficient and accurate tool for quantifying investor sentiment. On the other hand, empirical research demonstrates a close correlation between investor sentiment and stock price fluctuations. The sentiment factors constructed based on large language models can effectively predict stock returns, achieving an annualized excess return of 9.3% within the CSI 500 Index domain. The LSTM model combined with sentiment analysis can reduce the MAPE of stock price predictions to 2.72%, significantly outperforming traditional models. Moreover, changes in investor sentiment triggered by macroeconomic events can lead to substantial short-term fluctuations in the stock market, further demonstrating the correlation between the two.

The theoretical value of this study lies in providing new evidence from cutting-edge large language model technology for the "investor sentiment" hypothesis in behavioral finance. Its practical value lies in offering actionable high-frequency sentiment factors for quantitative investment strategies and providing a new perspective for market regulators to prevent systemic risks. Future research could focus on higher-frequency sentiment data, more complex multimodal models (combining textual and audio/video sentiment), and robustness testing of models under different market conditions.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Baker M, Wurgler J, 2006, Investor Sentiment and the Cross-section of Stock Returns. The Journal of Finance, 61(4): 1645–1680.

[2] Zhu H, Lu X, Xue L, 2023, A BERT-Based Sentiment Analysis Model for Financial Texts. Journal of Shanghai University (Natural Science Edition), 29(01): 118–128.

[3] Wang D, Liang Y, 2025, The Technological Foundations, Application Scenarios, and Risk Prevention of Large Language Models in Artificial Intelligence: Taking the Banking Industry as an Example. Journal of Dongbei University of Finance and Economics, (04): 17–30.

[4] Lu M, 2024, Research on the Application Principles, Challenges, and Implementation Paths of Large Language

Models in the Financial Sector. Journal of Chongqing Technology and Business University (Social Science Edition), 41(04): 1–12.

[5] Liu M, Zhang L, Ping W, et al., 2025, Research on a Multi-Stage Network Public Opinion-Driven Group Consensus Decision-Making Method Based on Large Language Models. Chinese Journal of Management, 22(04): 750–759.

[6] Jiang F, Liu Y, Meng L, 2024, Large Language Models, Text Sentiment, and Financial Markets. Management World, 40(08): 42–64.

[7] Weng X, Lin X, Zhao S, 2022, A Long Short-Term Memory Network Stock Price Movement Prediction Model Based on Empirical Mode Decomposition and Investor Sentiment. Computer Applications, 42(S2): 296–301.

[8] Yang S, Guo W, 2018, Investor Sentiment, Excess Returns, and Market Volatility. Journal of Hubei Engineering University, 38(02): 85–90.

[9] Brown GW, Cliff MT, 2004, Investor Sentiment and the Near-Term Stock Market. Journal of Empirical Finance, 11(1): 1–27.

[10] Xu T, 2018, Research on the Impact of Investor Sentiment on the Stock Market in Online Social Media. Shanghai Management Science, 40(03): 67–74.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Effects of Manifold Structures on Velocity Distribution of V- and A-Type Microchannel Plates

Pingnan Huang*, Liqing Ye

School of Mechatronic Engineering and Automation, Foshan University, Foshan 528225, China

**Author to whom correspondence should be addressed.*

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: Flow velocity uniformity of the microchannel plate is a major factor affecting the performance of microchannel devices. In order to improve the velocity distribution uniformity of the microchannel plate, we designed two new microchannel structures: V-type and A-type. The effects of various structural parameters of the manifolds on the velocity distribution are reported. The V-type and A-type microchannel plates had a more uniform velocity distribution compared to the Z-type microchannel plate. The final result showed that it is beneficial for the V-type microchannel plate to obtain a more uniform velocity distribution when the manifold structure parameters are X_{in} = -1, X_{out} = 0, Y_{in} = 10, Y_{out} = 6, H_{in} = 4, H_{out} = 1, and R = 0.5.

Keywords: Microchannel; Velocity distribution; Manifolds; Structure optimization

Online publication: October 15, 2025

1. Introduction

Because microchannel devices have channel equivalent diameters of 500 and below [1], they have attracted extensive attention due to their small size, high specific surface area, and high heat and mass transfer efficiency. At present, the application of microchannel devices is focused on microchannel reactors [2,3] and microchannel heat sinks [4,5]. The uniformity of flow velocity distribution between microchannels has a large impact on the performance of microchannel devices. In microchannel reactors, the uneven distribution of flow velocity between the microchannels leads to uneven residence time, resulting in uneven contact area between the reactants and the catalyst inside the microchannel, thereby reducing the reaction rate and reaction efficiency of the microchannel reactor. In the microchannel heat sink, the uneven distribution of the flow velocity causes the overall pressure drop of the microchannel plate to rise, thereby increasing the pump power. However, the heat that the fluid carries away from the microchannel plate is uneven, resulting in overheating in some areas, affecting the heat dissipation effect. How to improve the flow velocity distribution uniformity of microchannel devices is a key area of research.

At present, it is possible to alter the microchannel plate flow velocity distribution by changing the microchannel structure, the inlet and outlet structure, the inlet flow velocity, and the number of stacking layers. The need to improve the uniformity of the microchannel plate velocity distribution has become the main restriction of the application of microchannel devices. The manifold structure of the microchannel plate is flexible and variable, and has different forms that also have a significant influence on the uniformity of the flow velocity distribution of the microchannel plate [6]. Thus, optimizing the manifold structures of the microchannel devices is one of the key ways to improve the velocity distribution uniformity.

In this article, based on the monolithic manifold structures, V-type and A-type microchannel plates are improved in order to simplify the manifold structures, decrease the flow resistance, and improve the flow uniformity. The effects of different manifold structure parameters on the velocity distribution are analyzed by simulation, and the velocity distribution of V-type, A-type, and Z-type microchannel plates is compared.

2. Computational fluid dynamics setting and analysis

The structures of V-type and A-type microchannel plates that we studied are respectively shown in **Figures 1(a)** and **1(b)**. They are formed by adding an inlet or outlet on the basis of the traditional Z-type microchannel plate. Both V-type and A-type microchannel plates are composed of a rectangular microchannel array, an inlet, an outlet, and corresponding manifold structures. The main difference between the V-type and A-type microchannel plates is that the V-type microchannel plate contains two symmetric inlets and one outlet, while the A-type microchannel plate contains two symmetric outlets and one inlet.

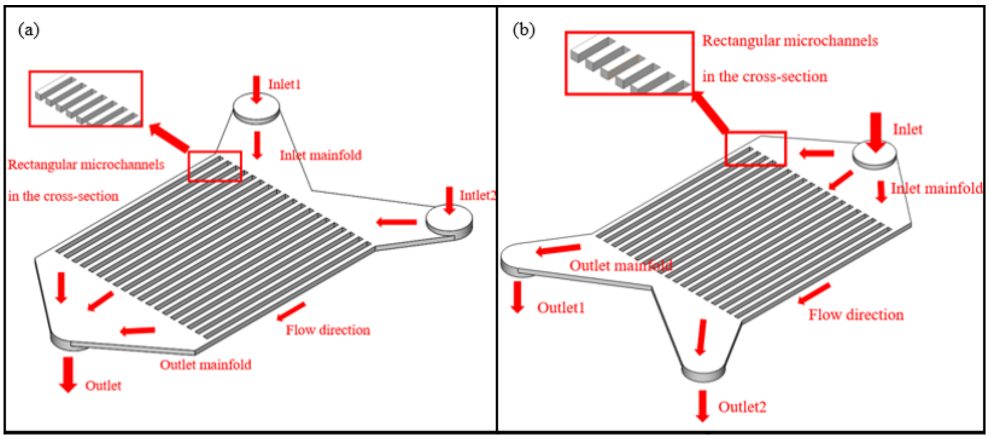


Figure 1. The structure of V and A types microchannel plates

Fluent17.0 was used to analyze the velocity distribution of the microchannel plate. The mesh division method divides the mesh by sub-areas. The microchannel array is divided by the sweep method, and the manifolds and the inlet and outlet are meshed by the automatic method. The independence test was carried out using a V-type microchannel plate reference model. The simulation results are shown in **Table 1**. The mesh unit size is controlled to be 0.05 mm in the reference model. It can be seen from the table that when the size of the control grid unit is 0.1 mm, the flow velocity distribution evaluation coefficient is only 0.064598, different from the reference model, so the overall grid is drawn by the control unit size of 0.1 mm. The result of the division is shown in **Figure 2**.

Table 1. Grid independence test

Element size (mm)	Element number	σ%	Differ
0.25	93726	0.698523	0.429523
0.1	1383475	0.520206	0.064598
0.05	10913538	0.488641	Baseline

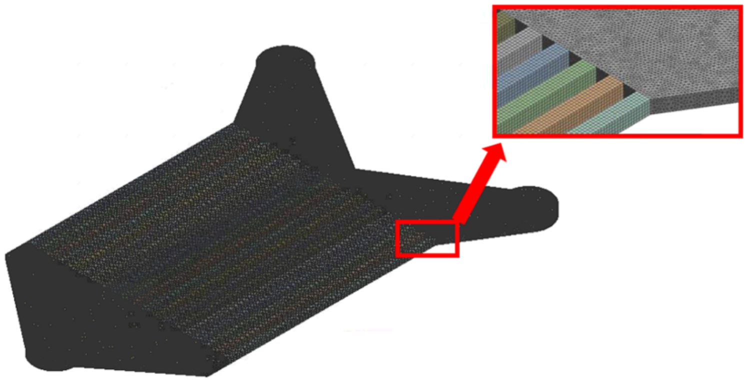


Figure 2. Mesh generation results

The inlet boundary condition is set as velocity-inlet. In order to make comparisons, the inlet quantity of flow is the same for V-type and A-type microchannel plates. Thus, the inlet velocity of the V-type microchannel plate is set as half of the A-type microchannel plate. The outlet boundary condition is set as "Pressure-Outlet." The static pressure of the outlet is 0 Pa. The other boundary condition is set as "wall." No slip condition is set. Flow liquid is liquid water. Density is 998.2 kg/m^3. The temperature is 300 K and the kinematic viscosity is 1.0032*10-3.

The Reynolds number is very small. Thus, the flow pattern in the microchannel plate is assumed to be laminar flow. The governing equation is:

$$\rho\frac{d\vec{v}}{dt} = -\nabla P + \mu\nabla^2\vec{v} + \rho\vec{F} \quad (1)$$

$$\nabla\vec{v} = 0 \quad (2)$$

where ρ, P, $\vec{v}$, μ, and $\vec{F}$ are the flow density, pressure, flow velocity, dynamic viscosity, and volume force per unit, respectively. ∇ is the differential.

In order to evaluate the flow velocity uniformity between microchannels, the flow velocity distribution evaluation coefficient σ% is defined, as shown in Equation 3. The smaller the value of σ% is, the more uniform the velocity distribution:

$$\sigma\% = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{U_c(i)}{U_m} - 1\right)^2} \times 100\% \quad (3)$$

Where $U_c(i)$ means the flow velocity of the ith microchannel, U_m is the average value of the flow velocity of all the microchannels. The definition is:

$$U_m = \frac{1}{N}\sum_{i=1}^{N} U_c(i)(=1, 2, 3\cdots N) \quad (4)$$

3. Results and discussion

3.1. Effects of X value

The effect of X value on the velocity distribution of V-type and A-type microchannel plates is shown in **Figure 3**. It can be seen that when the X value is -4, the flow velocity distribution curve of the V-type microchannel plate is inverted parabolic. The flow velocity of the middle microchannel is the smallest, and the flow velocity on both sides is the largest. As the value of X increases, that is, the inlet position is getting closer to the middle of the microchannel plate, the flow velocity begins to increase gradually, and the flow velocity on both sides decreases gradually. When the X value is 1, the flow velocity distribution curve is almost a parabolic shape, with minimum flow rates on both sides and large intermediate flow rates.

Figure 3(b) shows the velocity distribution on different X_{out} values of the V-type microchannel plate. When the microchannel plate is a positive V-type (X_{out} = 0), the velocity distribution curve is almost a horizontal line. The velocity distribution is the most uniform. As the X_{out} value increases, the outlet is far away from the middle channel, and the velocity distribution curve becomes an increasing slant, which means the velocity distribution becomes more and more uneven.

The velocity distribution of the A-type microchannel plate with the increase of the X value is shown in **Figures 3(c)** and **3(d)**. it can be seen that the bigger the X_{in} value is, the more uneven the velocity distribution is, which is similar to that of a V-type microchannel plate with different X_{out} values. It is seen that the flow velocity distribution of the V-type microchannel plate with different inlet parameters X_{in} values and the flow velocity of the A-type microchannel plate with different outlet parameters X_{out} values are similar.

3.2. Effects of Y value on the velocity distribution

The effect of Y value on the flow velocity distribution of V-type and A-type microchannel plates is shown in **Figure 4**. As for the V-type microchannel plate, when the inlet parameter Y_{in} = 3, the velocity distribution curve is an inverted parabolic shape, with minimum velocity in the middle channel and maximum velocity on both sides. With the increase of Y_{in} value, velocity in the middle microchannels gradually increases, and the velocity on both sides decreases gradually. When the Y_{in} value is 10, the velocity in the middle channels is almost the same, because as the Y_{in} value increases, the distance from the inlet to the intermediate channel 2 to channel 19 is almost equal, so the flow rates through channel 2 to channel 19 are also almost equal. The flow rate on both sides is slightly smaller. It may be that channel 1 and channel 20 are close to the inlet and outlet manifold turning point, and the local resistance is greater, so the flow rate is smaller.

The effect of Y_{out} value on the flow velocity distribution of the V-type microchannel plate is shown in **Figure 4(b)**. When Y_{out} value is 3, the flow velocity of the microchannel plate is distributed into a pagoda type. The flow velocity of the intermediate channel is the largest, and the flow velocity of both sides is the smallest. As the value of Y_{out} increases, the flow rate of the intermediate microchannel begins to decrease slowly, and the flow velocity at both ends begins to rise. When Y_{out} = 6, the velocity distribution curve is almost a horizontal line, and the velocity

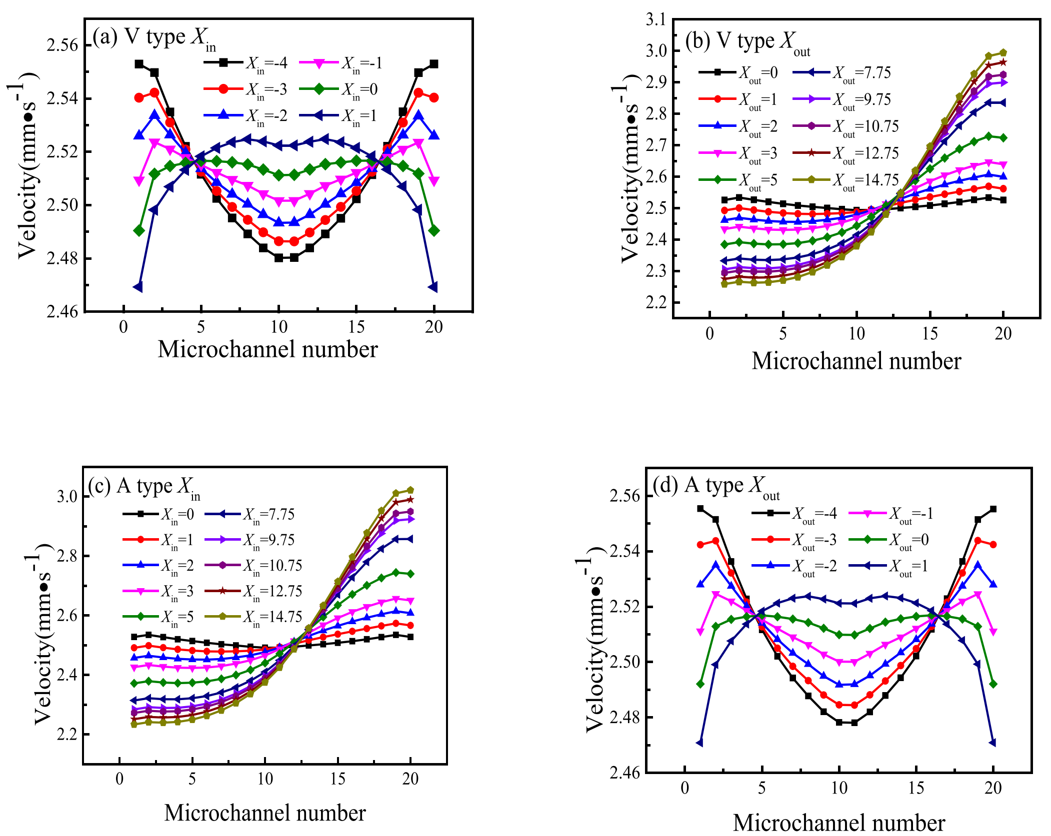


Figure 3. Velocity distribution of V-type and A-type microchannel plates at different X values. Effects of X_{in} values of (a) V-type and (c) A-type; Effects of X_{out} values of (b) V-type and (d) A-type

distribution is the most uniform. As the value of Y_{out} continues to increase, the length of the streamline flowing from the inlet to the outlet through the microchannels at both ends is longer than the length of the streamline of the intermediate microchannels at both ends is greater. Therefore, the flow rate of the intermediate microchannel continues to rise, and the flow velocity of the microchannels at both ends continues to decrease, eventually forming an inverted parabolic shape.

Comparison of **Figures 4(a)** and **4(d)**, **Figures 4(b)** and **4(c)**, it is seen that the influence of the inlet parameter Y_{in} of the V-type microchannel plate on the flow velocity distribution and the influence of the outlet parameter Y_{out} of the A-type microchannel plate on the flow velocity distribution are the same. Similarly, the influence of the V-type microchannel plate outlet parameter Y_{out} on the flow velocity distribution is the same as the influence of the A-type microchannel plate inlet parameter Y_{in} on the flow velocity distribution.

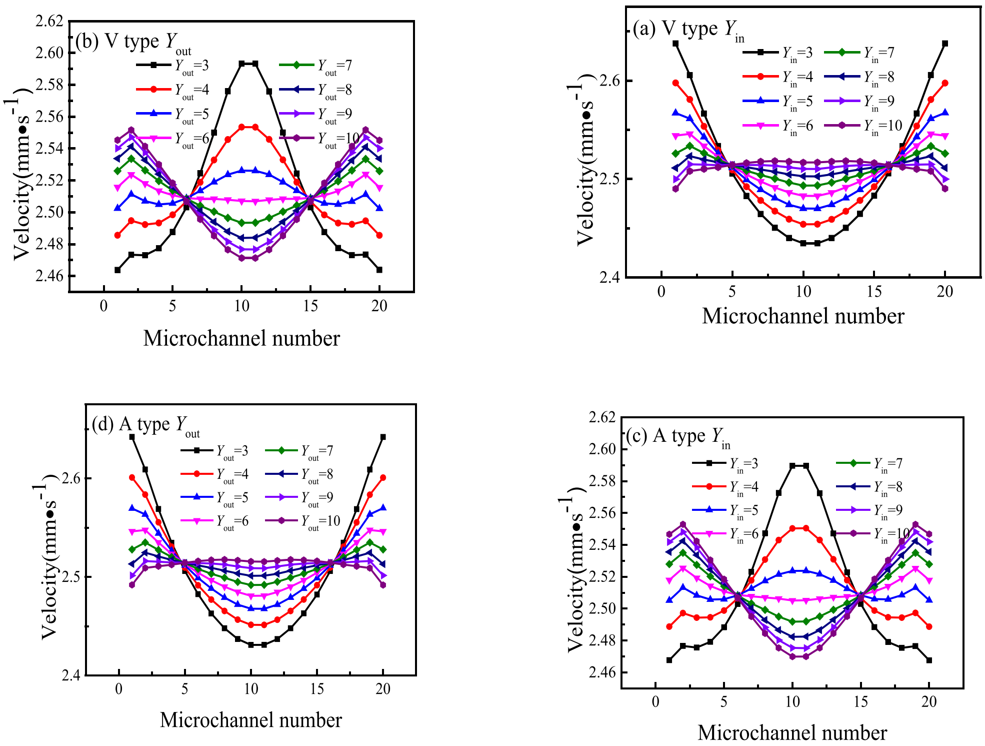


Figure 4. Velocity distribution of V-type and A-type microchannel plates at different Y values. Effects of Y_{in} values of (a) V-type and (c) A-type; Effects of Y_{out} values of (b) V-type and (d) A-type

3.3. Effects of H value

Figure 5 shows the change in velocity distribution of V-type and A-type microchannel plates with H value. It can be seen that when the inlet parameter H_{in} value is changed, the flow velocity distribution of V-type microchannel plate is similar to that of A-type microchannel plate when the outlet parameter H_{out} value is changed. When H_{in} (H_{out}) is 1, the flow velocity of V-type (A-type) microchannel plate presents a deep V-type. With an increase of H_{in} (H_{out}) value, the flow velocity of V-type (A-type) microchannel plate began to increase, and the flow velocity of both sides decreased. When the outlet parameter H_{out} of V-type microchannel plate and the inlet parameter H_{in} of A-type microchannel plate are changed, the change of velocity distribution is shown in **Figures 5(b)** and **5(c)**, respectively. When H_{out} (H_{in}) is 1, the flow velocity distribution of V-type (A-type) presents a flat M-type. As H_{in} value increases gradually, the flow velocity of the middle channels decreases, and the flow velocity of both sides increases gradually. When the value of H_{in} is 6, the velocity distribution is a deep V-type, with the smallest velocity in the middle channel and the largest velocity on both sides.

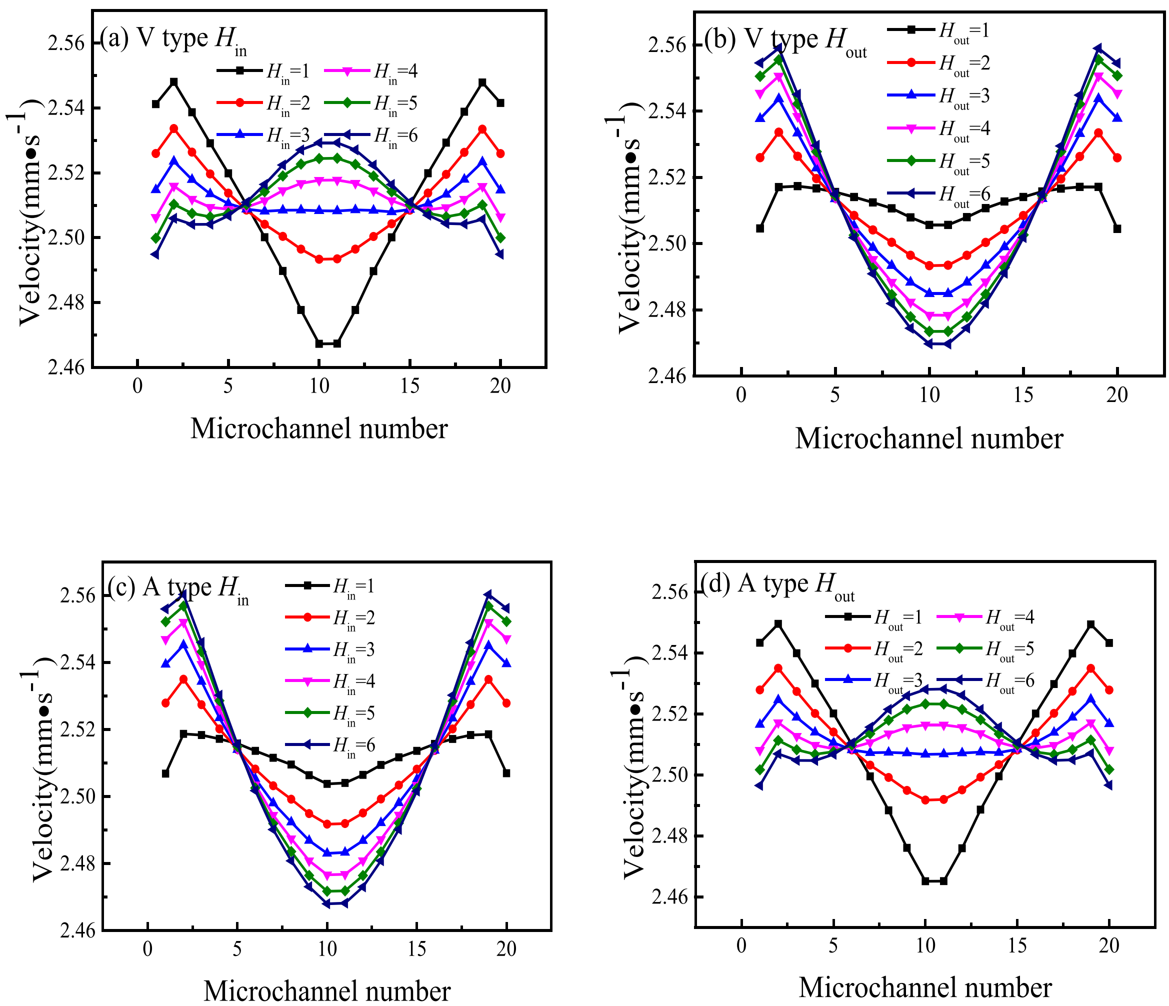


Figure 5. Velocity distribution of V-type and A-type microchannel plates at different H values. Effects of H_{in} values of (a) V-type and (c) A-type; Effects of H_{out} values of (b) V-type and (d) A-type

3.4. Effects of R value on the velocity distribution

Figure 6 shows the velocity distribution of V-type and A-type microchannel plates when the inlet and outlet radius are changed. As the radius increases, the flow velocity of the whole microchannel plate increases, because when the inlet velocity is constant, increasing the inlet radius, the inlet flow increases, so that the flow velocity of the whole microchannel plate increases. **Figure 6(c)** shows the relationship between the velocity distribution evaluation coefficient and the radius. With the increase of the radius R, the velocity distribution evaluation coefficient of V-type microchannel plates first increases and then decreases, while the velocity distribution evaluation coefficient of A-type microchannel plates monotonously increases. That is to say, as the increase of radius, the velocity distribution of V-type microchannel plate first becomes uniform and then uneven, but the velocity distribution of A-type microchannel plate becomes more and more uniform with the increase of the value of R.

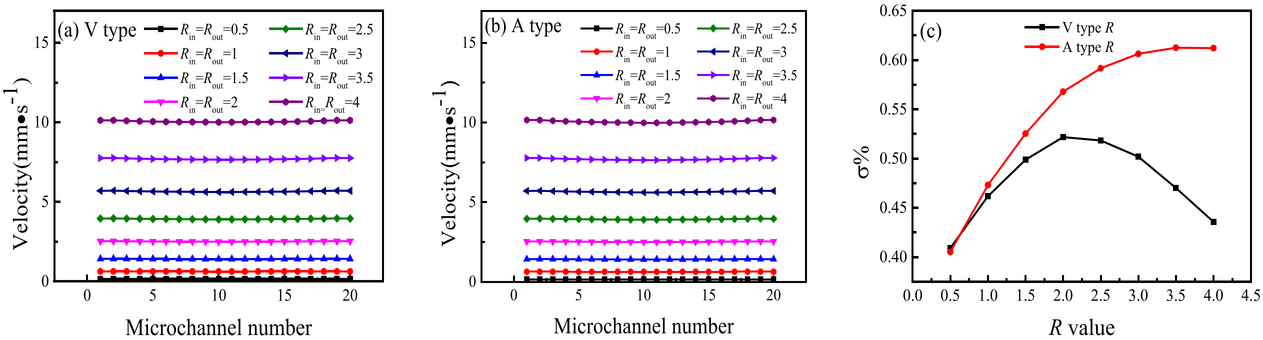


Figure 6. Velocity distribution of V-type and A-type microchannel plates at different R values. (a) V-type; (b) A-type; (c) Velocity distribution evaluation coefficient

3.5. Comparison of velocity distribution of V-type and A-type

From the above analysis, it can be found that the inlet parameters of V-type microchannel plate X_{in}, Y_{in}, and H_{in}, have the same influence on the velocity distribution as the outlet parameters of A-type microchannel plate, X_{out}, Y_{out}, and H_{out}. Similarly, changes in outlet parameters of V-type microchannel plate (X_{out}, Y_{out}, and H_{out}) have the same impact on velocity distribution as changes in inlet parameters of A-type microchannel plate (X_{in}, Y_{in}, and H_{in}). However, when the R value changes, the velocity distribution uniformity of V-type microchannel plate increases first and then decreases, while that of A-type microchannel plate increases monotonously.

3.6. Comparison of velocity distribution between Z-type and V-type and A-type

The flow velocity distribution of V-type and A-type microchannel plates is further compared to that of Z-type microchannel plates studied earlier [6] (**Figure 7**). Among them, both V-type and A-type microchannel plates adopt the reference model. It can be seen that the flow velocity distribution of V-type and A-type microchannel plates is basically the same, and the flow velocity distribution presents a shallow V-type. Compared with V-type and A-type microchannel plates, the flow velocity of the Z-type microchannel plate presents a deep V-type, with a larger flow velocity on both sides and a smaller flow velocity in the middle. Combined with **Figure 7(b)**, the evaluation coefficient of the velocity distribution of V-type and A-type microchannel plates is less than half that of Z-type microchannel plates. In other words, the velocity distribution of V-type and A-type microchannel plates is more uniform than that of Z-type microchannel plates.

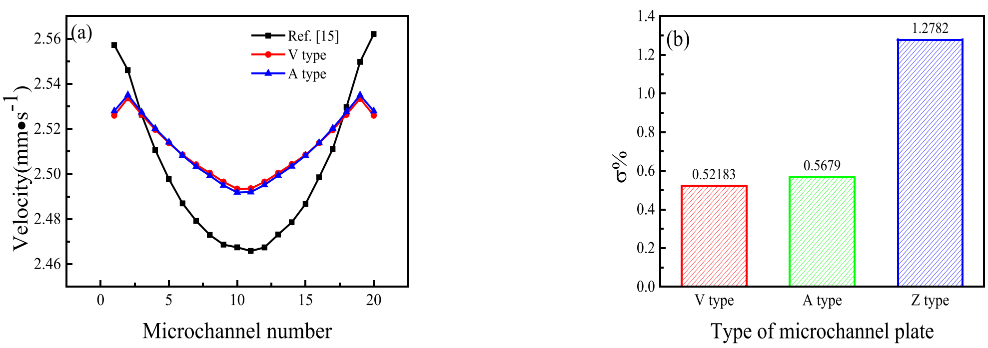


Figure 7. Velocity distribution of V-type, A-type, and Z-type microchannel plate

4. Conclusion

In this paper, two types of manifold structures, V-type and A-type, are used to study the flow velocity distribution under different manifold structure parameters. The results show that the inlet parameters of V-type microchannel plate, X_{in}, Y_{in}, and H_{in}, have the same effect on velocity distribution as the outlet parameters of A-type microchannel plate, X_{out}, Y_{out}, and H_{out}. When $X_{in} = -1$, $X_{out} = 0$, $Y_{in} = 10$, $Y_{out} = 6$, $H_{in} = 4$, $H_{out} = 1$, $R = 0.5$, the velocity distribution of V-type microchannel plate is the most uniform. Comparison of the flow velocity distribution of Z-type, V-type, and A-type microchannel plates reveals that the velocity distribution evaluation coefficient of V-type and A-type microchannel plates is less than half that of Z-type microchannel plates, and the flow velocity distribution is more uniform.

Funding

This research was supported by Scientific Research Project of Guangdong Provincial Department of Education (2024KQNCX152).

Disclosure statement

The authors declare no conflict of interest.

References

[1] Delparish A, Koc S, Caglayan BS, et al., 2018, Oxidative Steam Reforming of Glycerol to Synthesis Gas in a Microchannel Reactor. Catalysis Today, 299: 289–298.

[2] Engelbrecht N, Chiuta S, Bessarabov DG, 2018, A Highly Efficient Autothermal Microchannel Reactor for Ammonia Decomposition: Analysis of Hydrogen Production in Transient and Steady-State Regimes. Journal of Power Sources, 386: 47–55.

[3] Alfaryjat AA, Mohammed HA, Adam NM, et al., 2018, Numerical Investigation of Heat Transfer Enhancement Using Various Nanofluids in Hexagonal Microchannel Heat Sink. Thermal Science and Engineering Progress, 5: 252–262.

[4] Wang RJ, Wang JW, Li BQ, et al., 2018, Parameterization Investigation on the Microchannel Heat Sink with Slant Rectangular Ribs by Numerical Simulation. Applied Thermal Engineering, 133: 428–438.

[5] Cao J, Kraut M, Dittmeyer R, et al., 2018, Numerical Analysis on the Effect of Bifurcation Angle and Inlet Velocity on the Distribution Uniformity Performance of Consecutive Bifurcating Fluid Flow Distributors. International Communications in Heat and Mass Transfer, 93: 60–65.

[6] Pan M, Huang P, Vafai K, 2017, Investigation of the Blockage Conditions in a Laminated-Sheet Microchannel Reactor. Chemical Engineering & Technology, 40: 2283–2294.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Fluorescent Temperature Characteristics of CaMoO4:5%Tb^{3+} Based on Variable Temperature Excitations

Meilin Song*, Changwen Wang, Hongxia Tang, Changxing Yu, Yue Qiao

Suihua University, Suihua 152001, Heilongjiang, China

**Author to whom correspondence should be addressed.*

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: This study focuses on the fluorescent thermometric properties of CaMoO4:5%Tb3+ under different temperature excitations. At the detection wavelength of 544 nm, with the temperature varying from 293 K to 563 K, there is a broadband absorption peak in the range of 250 nm to 350 nm. The results indicate that this phenomenon is caused by the superposition of the 4f-5d transition of Tb3+ and the O2--Mo6+ charge transfer. It is considered that as the temperature rises, the luminescent intensity of the material shows an obvious continuous decreasing trend, which reflects a significant luminescent thermal quenching trend; thus, this quenching belongs to the "strong coupling" type. Based on the excitation spectrum results, two excitation wavelengths, 312 nm and 338 nm, were specifically selected to excite the samples, which correspond to the top of the charge transfer band, the redshift intersection of the charge transfer band, and the edge of the charge transfer band at 293 K, respectively.

Keywords: Strong coupling; Fluorescent temperature characteristics; Excitation

Online publication: October 15, 2025

1. Introduction

Rare earth elements have unfilled 4f5d electron configurations shielded from the external environment, thus possessing abundant electronic energy levels and long-lived excited states, which can generate a variety of radiation absorption and emission. Luminescence is a phenomenon in which an object directly converts internally absorbed energy (in some way) into non-equilibrium radiation without going through a thermal stage. Molybdate-based tricolor phosphors have attracted widespread attention due to their advantages, such as low sintering temperature, stable properties, high color rendering, good thermal stability, adjustable light color and color temperature, uniform light color that does not change with current, etc. Therefore, molybdate phosphors are regarded as a very promising fluorescent material. When Tb^{3+} is at the inversion symmetry center of the calcium molybdate lattice, atoms transition photons outward in the form of magnetic dipoles, and the magnetic dipole transition does not change with the intensity of the crystal field around Tb^{3+}. According to the Judd-Ofelt theory [1], when Tb^{3+} deviates from the inversion symmetry center of the calcium tungstate-molybdate lattice, atoms transition photons outward in the form of electric dipoles. Since the system is composed of two point charges with equal magnitude and opposite signs, the electric dipole rotates under the action of torque in the external electric field, causing the electric dipole moment to turn to the external electric field. Therefore, the luminescent intensity increases with the increase of the distortion degree of the crystal field around Tb^{3+}. The electric dipole transition is affected by the local electric field and is very sensitive to the symmetry around Tb^{3+}. This paper takes tungstate-molybdate as the main matrix; the other part is a small amount of component doped into the matrix, called the activator. The activator plays an important role in luminescent properties, and can even affect the luminescent color, brightness, and other characteristics, so it is also called the luminescent center. This study investigates the influence of doped Tb^{3+} ions on their luminescent properties and the energy transfer mechanism. Calcium molybdate belongs to the tetragonal scheelite structure, with advantages such as good thermodynamic and chemical stability, excellent luminescent properties, and low synthesis temperature. It is widely used in fluorescent lamps, display panels, solid-state lasers, etc., and is regarded as a perfect fluorescent matrix material [2].

In this study, rare earth molybdate materials with a scheelite structure were designed and selected, and the $^5D_4$–^7F energy level of Tb^{3+} was chosen as the research object. The CaMoO4:5%Tb^{3+} material was synthesized by the high-temperature solid-state method [3]. Through variable-temperature excitation spectra, excitation strategies with wavelengths of 312 nm and 338 nm were designed to excite molybdate, and then variable-temperature excitation spectra were obtained. It was found that the four energy levels of Tb^{3+} ions, i.e., $^5D_4$–$^7F_3$ emission, $^5D_4$–$^7F_4$ emission, $^5D_4$–$^7F_5$ emission, and $^5D_4$–$^7F_6$ emission, showed different changes with temperature, and the sources of these different changes were pointed out. By fitting the change trends of different behaviors, the fitting curves applicable to temperature measurement were obtained. The thermal population of the matrix ground state was explored to construct the opposite change trend of fluorescent intensity, realizing high-sensitivity temperature measurement, and a theoretical explanation was given.

2. Experiment

Samples of MoO_3, CaO, and Tb_4O_7 were weighed according to different stoichiometric ratios, respectively dissolved in an appropriate amount of distilled water, and stirred until completely dissolved. Then, rare earth ions were added to the sodium molybdate solution to allow the solution to fully react. After centrifugation, washing, drying, and grinding, $CaMoO_4$ powder doped with rare earth ions was obtained, which was calcined in a muffle furnace at 600°C for 3 hours. A series of $CaMoO_4$:5%Tb^{3+} samples with a Tb doping concentration of 5% was obtained.

3. Results and discussion

3.1. Temperature-dependent excitation spectra

When $CaMoO_4$:5%Tb^{3+} was excited using the energy levels corresponding to the matrix absorption and the $^5D_4$–^7F transition absorption of Tb^{3+} [4], the effect of different temperatures on the emission spectrum of the $CaMoO_4$:5%Tb^{3+} luminescent material at a Tb^{3+} concentration of 5% was investigated. As shown in **Figure 1**, at a detection wavelength of 544 nm and with the temperature varying from 293 K to 563 K, there is a broad absorption peak in the range of 250 nm to 350 nm, which is caused by the superposition of the 4f-5d transition of

Tb^{3+} and the O^{2-}-Mo^{6+} charge transfer. It is worth noting that as the temperature increases, the luminescent intensity of the material shows a very obvious continuous decreasing trend, indicating a significant luminescent thermal quenching trend. Therefore, this quenching [5] belongs to the "strong coupling" type.

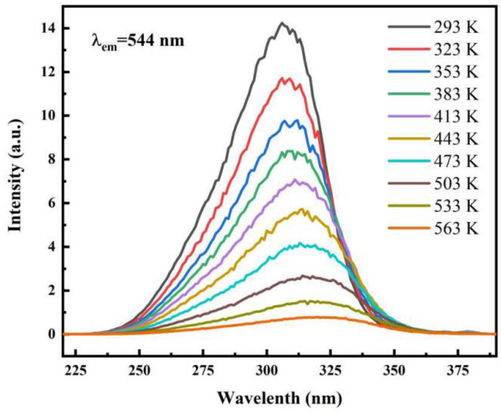


Figure 1. Temperature-dependent excitation spectra at a wavelength of 544 nm

As shown in the figure, the temperature-dependent excitation spectra at a wavelength of 544 nm were detected in the temperature range of 293 K to 563 K, corresponding to the characteristic emission of the $^5D_4 \rightarrow {}^7F$ transition. With the gradual increase in temperature, the excitation spectrum exhibits obvious characteristics: it shifts toward longer wavelengths (i.e., the wavelength increases and the frequency decreases), and the spectral lines weaken simultaneously. This indicates that in the temperature range of 293 K to 563 K, the Mo-O charge transfer weakens as the temperature rises. This is because the non-radiative process of the $^5D_4 \rightarrow {}^7F$ energy level is enhanced with increasing temperature, leading to a reduction in particles at the energy level and a consequent decrease in the characteristic emission of $^5D_4 \rightarrow {}^7F$. The laser spectrum of Tb^{3+} includes a strong absorption near 245 nm and weak absorptions in the range of 290 nm to 390 nm. Due to the overlap between the Mo-O charge absorption band and the charge absorption band of Tb^{3+} [6], no obvious charge absorption peak of Tb^{3+} was observed.

3.2. Temperature-dependent emission spectra excited at 312 nm and thermal coupling properties

Based on the excitation spectrum results, two excitation wavelengths, 312 nm and 338 nm, were selectively chosen to excite the samples. These wavelengths correspond to the top of the charge transfer band, the redshift intersection of the charge transfer band, and the edge of the charge transfer band at 293 K, respectively. We hypothesize that the redshift characteristic of the charge transfer band edge can be utilized to construct three opposite thermal behaviors in a single material, thereby exploring the influence of excitation light on thermal coupling [7].

To verify this hypothesis, the samples were excited at 312 nm and 338 nm, respectively. **Figure 2** shows the temperature-dependent emission spectra excited at 312 nm. The main emission peak near 545 nm corresponds to the $^5D_4 \rightarrow {}^7F_5$ transition emission. In addition, the emission at 488 nm corresponds to $^5D_4 \rightarrow {}^7F_6$, 585 nm to $^5D_4 \rightarrow {}^7F_4$, and 620 nm to $^5D_4 \rightarrow {}^7F_3$. With increasing temperature, all emissions corresponding to the $^5D_4 \rightarrow {}^7F_5$ energy level show a decreasing trend, as do those at 488 nm, 585 nm, and 620 nm. The emission intensity reaches the maximum when the calcination temperature is 293 K.

No emission from the $^5D_3 \rightarrow {}^7F_j$ transition was observed, which is attributed to the cross-relaxation effect between

Tb^{3+} ions. This process can be expressed as: Tb^{3+} ($^5D_3$) + Tb^{3+} ($^7F_6$) → Tb^{3+} ($^5D_4$) + Tb^{3+} ($^7F_0$). Additionally, it may be due to the large phonon energy of the molybdate material, where multi-phonon relaxation causes the quenching of the $^5D_3$→^7F_j transition. **Figure 2** more intuitively shows the electric dipole transitions at 293 K and 563 K.

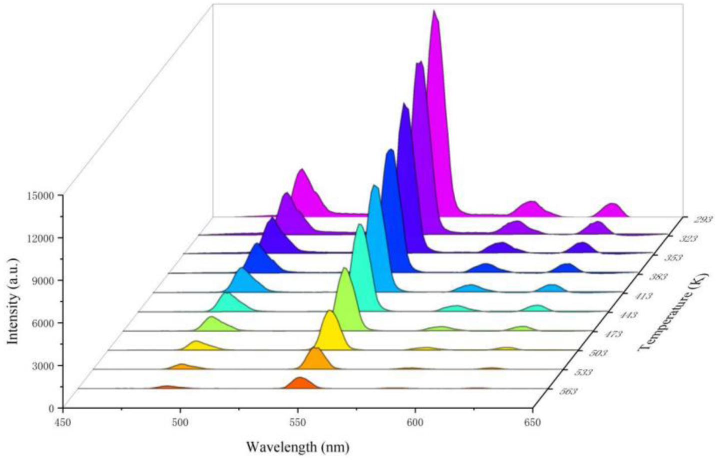


Figure 2. Temperature-dependent emission spectra with an excitation wavelength of 312 nm

To better observe the variation trend of fluorescence intensity and the thermal coupling properties of the two energy levels, further processing was performed on the emission spectral data. **Figure 3** shows the temperature-dependent integrated emission intensity under an excitation wavelength of 312 nm, where the marked points correspond to the emission from the $^5D_4$→$^7F_5$ transition. The two characteristic emission energy levels were fitted using Formula (1):

$$R=B\exp(-\frac{\Delta E}{kT}) \quad (1)$$

In formula (1), R is the fluorescence intensity ratio, B is a constant representing the energy level difference between two thermally coupled energy levels, k is approximately 0.695, which is the Boltzmann constant [8], and T is the thermodynamic temperature.

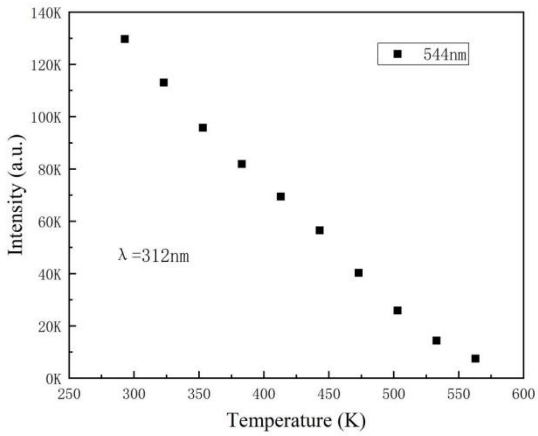


Figure 3. Temperature-dependent emission integrated intensity with an excitation wavelength of 312 nm

3.3. Temperature-dependent emission spectra under 338 nm excitation and thermal coupling properties

To construct the same thermal quenching behavior by utilizing the characteristic of red shift at the edge of the charge transfer band, temperature-dependent emission spectrum tests were also conducted with an excitation wavelength of 338 nm, as shown in **Figure 4**. The characteristic emission with the main emission peak around 545 nm corresponds to the $^5D_4 \rightarrow ^7F_5$ transition emission. With the increase of temperature, the emissions at 488 nm, 585 nm, and 620 nm first enhance and then weaken. **Figure 5** shows the emission spectra under an excitation wavelength of 338 nm at 293 K and 593 K. The emission at the left 1 with a central wavelength of 488 nm corresponds to the $^5D_4 \rightarrow ^7F_6$ transition emission; the emission at the left 2 with a central wavelength of 545 nm corresponds to the $^5D_4 \rightarrow ^7F_5$ transition emission; the emission at the left 3 with a central wavelength of 585 nm corresponds to the $^5D_4 \rightarrow ^7F_4$ transition emission; and the emission at the right with a central wavelength of 620 nm corresponds to the $^5D_4 \rightarrow ^7F_3$ transition emission.

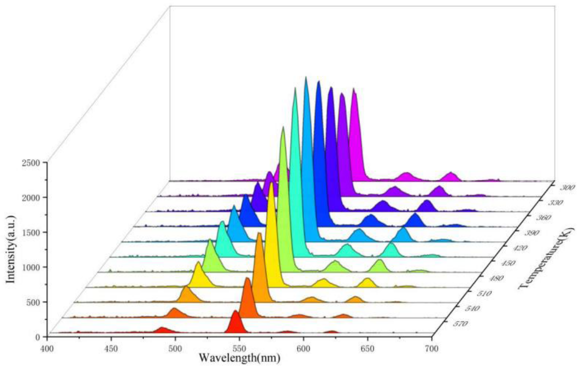


Figure 4. Temperature-dependent emission spectra with an excitation wavelength of 338 nm

The excitation at 345 nm shows an opposite thermal quenching behavior compared to that at 312 nm. To further verify whether different excitations affect the thermal coupling trend and alter the relative sensitivity of fluorescence thermometry [9], the fluorescence emission spectral data under 338 nm excitation were also processed and analyzed. As shown in **Figure 5**, the black markers in the temperature-dependent integrated emission intensity under 338 nm excitation correspond to the emission from the $^5D_4 \rightarrow ^7F_5$ transition. With the increase in temperature, the emission from the $^5D_4 \rightarrow ^7F_5$ transition first increases and then decreases, among which the emission at 544 nm rises rapidly, while the emissions from the other three emission peaks increase slowly.

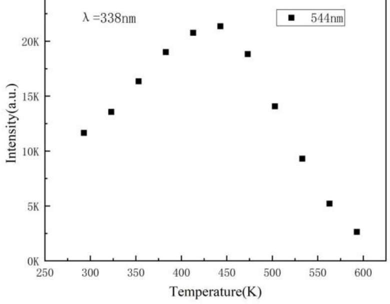


Figure 5. Temperature-dependent emission integrated intensity with an excitation wavelength of 338 nm

3.4. Temperature measurement sensitivity

The relative sensitivity S_r can describe the temperature sensing sensitivity more accurately than S. It can be expressed by formula (2):

$$S_r = \frac{1}{R}\frac{dR}{dT} \times 100\% \quad (2)$$

Relative sensitivity refers to the percentage change in the initial R value per unit temperature. When evaluating the performance of temperature-sensing materials [10], it is more meaningful to use relative sensitivity as a parameter. **Figure 6** below shows the relative sensitivity of the ratio of 338 nm to 312 nm. At 303 K, the relative sensitivity S_r reaches a maximum of 1.3%K^{-1}, and it is more sensitive at low temperatures.

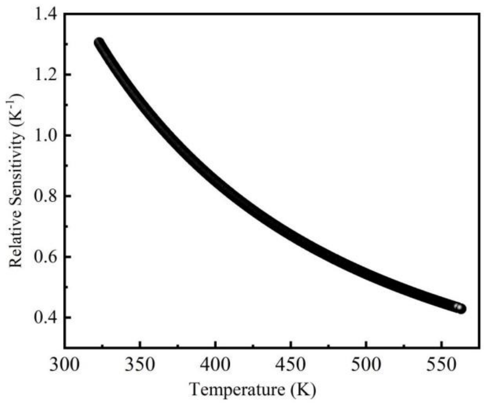


Figure 6. Sensitivity curves of excitation wavelengths

4. Conclusion

In this work, we designed and selected the $^5D_4$ energy level transition of Tb^{3+} ions as the research object. The CaMoO4:5%Tb^{3+} material was synthesized by the high-temperature solid-phase method. Through variable-temperature excitation spectroscopy, two excitation strategies were designed to excite the charge transfer band of MoO_4^{2+}. Furthermore, by measuring the variable-temperature emission spectra, it was found that the emissions of Tb^{3+} ions, i.e., $^5D_4$–$^7F_6$, $^5D_4$–$^7F_5$, $^5D_4$–$^7F_4$, and $^5D_4$–$^7F_3$, show different changes with the increase of temperature, and the sources of these different changes were pointed out. By fitting the trends of the two thermal behavior changes, the fitting curves applicable to temperature measurement were obtained. When the temperature is 303 K, the relative sensitivity of temperature measurement using thermally coupled energy levels reaches the maximum, and the maximum relative sensitivity of the ratio of the two excitations at 312 nm and 338 nm is 1.3%K^{-1}. The influence of matrix thermal population on the fluorescence intensity ratio was explored. The opposite trend of fluorescence intensity was constructed by using the thermal population of the matrix ground state, realizing high-sensitivity temperature measurement, and a theoretical explanation was given. This study may have certain reference value for the dual-excitation temperature measurement method based on thermally opposite behaviors.

Funding

(1) Basic Scientific Research Operating Expenses Project of Provincial Undergraduate Colleges and Universities

in Heilongjiang Province: Study on Luminescent Properties and Fluorescent Temperature Characteristics of Rare Earth Luminescent Materials Based on Tungstates (YWK10236210223)

(2) Provincial General Project: College Students' Innovation and Entrepreneurship Training Program Project: Design of a Multifunctional Intelligent Car (202310236033)

Disclosure statement

The authors declare no conflict of interest.

References

[1] Yang Y, Wang K, Zhao Z, et al., 2021, Research Progress of Rare Earth Europium and Terbium Complexes in the Field of Temperature Luminescent Sensing. Journal of the Chinese Society of Rare Earths, 39(1): 88.

[2] Zhou H, Wu F, Zhang Z, et al., 2022, Upconversion Luminescence and Temperature Sensing Properties of Lu_2O_3:Er^{3+}/Yb^{3+} Fluorescent Materials. Chinese Journal of Luminescence, 43(02): 192.

[3] Sun Y, Lu S, Meng Q, et al., 2017, Luminescence and Optical Temperature Sensing Properties of NaY(MoO4)2:Er^{3+} Nanocrystal Phosphors. Chinese Journal of Luminescence, 38(12): 1582.

[4] Hua J, Chen B, Sun J, et al., 2010, Upconversion Luminescence of Rare Earth Doped Materials. Chinese Journal of Optics and Applied Optics, 3(04): 301.

[5] Wang D, Xue B, Tu L, et al., 2021, Neodymium-Sensitized Multi-Shell Nanostructures for Enhanced Dye-Sensitized Upconversion Luminescence. Chinese Optics, 14(02): 418.

[6] Zhang Z, Ma J, Zhao W, et al., 2016, Hydrothermal Synthesis and Luminescent Properties of NaGd(MoO4)2:Eu^{3+} Phosphors. Bulletin of the Chinese Ceramic Society, 11(35): 3541–3545.

[7] Xu X, Gu Y, Kong M, et al., 2021, NaYF4: Nd^{3+}, Yb^{3+} Nanoparticles for Temperature Measurement in Biological Systems. Journal of the Chinese Society of Rare Earths, 39(1): 181.

[8] Yang X, Wu ZY, Yang ZN, et al., 2021, Flame-Made Y2O3: Yb^{3+}/Er^{3+} Up-Conversion Nanoparticles Mass Production Synthesis, Multicolor Tuning and Termal Sensing Studies. J. Alloys Compd., 854: 157078.

[9] Liang YJ, Chui PF, Sun XN, et al., 2013, Hydrothermal Synthesis and Upconversion Luminescent Properties of YVO4: Yb^{3+}, Er^{3+} Nanoparticles. J. Alloys Compd., 552: 289.

[10] Yan R, 2009, Temperature Characteristics of Erbium-Doped Yttrium Oxide Based on Fluorescence Intensity Ratio, dissertation, Master's thesis, Harbin Institute of Technology.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Exploring 3D Model Rendering Techniques for Cultural Relics Based on 3D Gaussian Splatting

Keran Yu*

Ara Institute of Canterbury International Engineering College, Shenyang Jianzhu University (Sino-New Zealand International Engineering Institute), Shenyang 110168, Liaoning, China

**Author to whom correspondence should be addressed.*

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: With the widespread application of 3D visualization in digital exhibition halls and virtual reality, achieving efficient rendering and high-fidelity presentation has become a key challenge. This study proposes a hybrid point cloud generation method that combines traditional sampling with 3D Gaussian splatting, aiming to address the issues of rendering delay and missing details in existing 3D displays. By improving the OBJ model parsing process and incorporating an adaptive area-weighted sampling algorithm, we achieve adaptive point cloud generation based on triangle density. Innovatively, we advance the ellipsoidal parameter estimation process of 3D Gaussian splatting to the point cloud generation stage. By establishing a mathematical relationship between the covariance matrix and local curvature, the generated point cloud naturally exhibits Gaussian distribution characteristics. Experimental results show that, compared to traditional methods, our approach reduces point cloud data by 38% while maintaining equivalent visual quality at a 4096×4096 texture resolution. By introducing mipmap texture optimization strategies and a GPU-accelerated rasterization pipeline, stable rendering at 60 frames per second is achieved in a WebGL environment. Additionally, we quantize and compress the spherical harmonic function parameters specific to 3D Gaussian splatting, reducing network transmission bandwidth to 52% of the original data. This study provides a new technical pathway for fields requiring high-precision display, such as the digitization of cultural heritage.

Keywords: 3D model; Dense point cloud; 3D Gaussian splatting

Online publication: October 21, 2025

1. Introduction

When conducting three-dimensional (3D) presentations in 3D data exhibition halls, 3D exhibition windows on websites, or virtual reality data displays, the issue of stuttering often arises. This stuttering significantly affects user experience. There are two primary reasons for stuttering: the first is hardware configuration limitations, and the second is the high computational cost of data processing, which increases data loading time. Additionally, in terms of visual effects, existing model data displays often become noticeably blurry when zoomed in to a certain extent, further affecting the presentation quality [1].

How can we effectively address these issues, reduce loading time, decrease power consumption, optimize display effects, and enhance user experience? The goal is to achieve a smooth and stutter-free display on web pages. Currently, 3D Gaussian splatting technology represents a rendering technique based on a 3D Gaussian distribution. It expresses point clouds or sparse data as dynamically adjusted Gaussian ellipsoids, enabling efficient and high-fidelity 3D reconstruction and rendering through projection and rasterization [2]. This paper will explore the conversion of 3D models into dense point cloud data, providing a foundation for the efficient rendering of existing massive 3D model datasets using 3D Gaussian splatting technology.

2. Related work

2.1. Bottlenecks of traditional rendering techniques

Traditional 3D model rendering primarily relies on polygonal meshes (such as the OBJ format) and their associated texture maps. However, when presenting high-precision cultural relic models in lightweight environments like the web, this approach faces significant bottlenecks [3]. Geometric complexity is the primary challenge: to capture the intricate details of cultural relics (such as engraved decorations), models often contain massive amounts of triangular faces (for example, the sample Buddha statue has 35,578 faces), resulting in an excessive burden on vertex processing and rasterization calculations, which can cause rendering stuttering. The pressure of texture resolution is also prominent: although high-resolution textures like 4096×4096 can ensure visual fidelity, they are limited by video memory bandwidth and sampling efficiency, which can easily cause loading delays and frame rate drops on the web. Additionally, the issue of blurred details upon zooming in is widespread—when a model is partially zoomed in, it relies on texture interpolation, and geometric details lose clarity due to inadequate sampling, making it difficult to meet the high-precision display requirements of cultural heritage. These bottlenecks restrict users' smooth interactive experiences in digital exhibition halls.

2.2. Research progress of 3D Gaussian splatting technology

3D Gaussian splatting, as an emerging rendering technique for radiance fields, has become a research hotspot in recent years, breaking through the limitations of traditional rendering methods. Its core innovation lies in representing scenes as dynamically adjusted collections of 3D Gaussian ellipsoids, where each ellipsoid encompasses spatial position (mean), shape and orientation (covariance matrix), and radiometric properties (such as spherical harmonic coefficients). Compared to methods based on implicit neural representations like NeRF [4], 3D Gaussian splatting offers significant advantages: (1) Rendering efficiency: Real-time frame rates are achieved through the projection of 2D Gaussian kernels and parallel rasterization utilizing GPUs. (2) Explicit controllability: The physical meaning of ellipsoid parameters is clear, supporting geometric editing and compression optimization. Existing research focuses on offline and desktop scene reconstruction, generating initial point clouds through Structure from Motion (SfM) and iteratively optimizing ellipsoid parameters. However, research on deploying high-precision, pre-built models (such as cultural relic OBJs) on the web remains unexplored. Specifically, the efficient conversion of massive mesh models into lightweight Gaussian representations suitable for 3DGS rendering, and addressing network transmission and browser-side computing constraints [5], has become an urgent direction to explore.

3. Point cloud generation method based on hybrid sampling

3.1. Steps for OBJ model parsing and point cloud generation

Extract vertex coordinates from the OBJ model file.

Determine the triangular faces of the model based on the vertex coordinates.

Calculate the total area of all triangular faces in the model using the vertex coordinates of the triangular faces.

Arbitrarily select a seed point within the area range corresponding to any triangular face.

Continue to sample points based on the total area of the triangular faces from the seed point, satisfying the pre-set number of sampling points (the number of sampling points meets the pre-set resolution), and record the coordinates of the sampled points.

Once sampling is complete, a dense point cloud is generated.

Based on the mapping relationship between the dense point cloud coordinates and the texture point coordinates, each pixel of the texture map is mapped to the corresponding point location.

This completes the point-to-point texture mapping process for the dense point cloud.

For example, consider the Zhou Ming Buddha statue from the Northern Song Dynasty's Chaozhou kiln: The OBJ model has 17,791 points and 35,578 faces; the Ply point cloud model has 17,791 points; and a total of 16,777,216 points (with a texture map resolution of 4096×4096) are sampled based on the number of pixels.

The traditional method used to be as follows: Retrieve the texture vertex coordinates corresponding to the OBJ model. Based on the triangular patches formed by these vertex coordinates and their corresponding triangular texture patches, determine the texture point coordinates within the triangular texture patches by taking point coordinates within the range of the triangular patches.

Complete the face-to-face texture mapping process based on these texture point coordinates.

3.2. Adaptive area-weighted sampling steps

Calculate the area of each triangular patch in the model.

Compute a random triangular patch based on the number of sampling points and the area of the triangular patches.

Randomly initialize any seed point r within the triangular patch, and randomly initialize r_1 and r_2. Calculate the three vertices of the new random triangular patch, denoted as $A(a_1, a_2, a_3)$, $B(b_1, b_2, b_3)$, and $C(c_1, c_2, c_3)$, using the following formulas:

$$a_1 = a_1 \times \sqrt{1-r_1}\ , a_2 = a_2 \times \sqrt{1-r_1}\ , a_3 = a_3 \times \sqrt{1-r_1} \quad (1);$$

$$b_1 = b_1 \times (1-r_2)\ , b_2 = b_2 \times (1-r_2)\ , b_3 = b_3 \times (1-r_2) \quad (2);$$

$$c_1 = \sqrt{r_1} \times (r_2 \times c_1 + b_1) + a_1\ , c_2 = \sqrt{r_1} \times (r_2 \times c_2 + b_2) + a_2\ , c_3 = \sqrt{r_1} \times (r_2 \times c_3 + b_3) + a_3 \quad (3)$$

Iterate the sampling process until the preset number of sampling points is reached.

4. 3D Gaussian splat data conversion

4.1. Estimation of ellipsoidal parameters

The parameters of a 3D Gaussian ellipsoid consist of a mean vector, a covariance matrix, and radiative characteristic parameters. The mean vector corresponds to the three-dimensional coordinates of the point cloud sampling points. The covariance matrix determines the spatial shape and orientation of the ellipsoid. The radiative characteristic parameters are associated with the color information of the texture mapping. In this study, the estimation of ellipsoidal parameters is brought forward to the point cloud generation stage, and adaptive parameter calculation is achieved through local geometric feature analysis.

The specific process is as follows: for each sampling point in the dense point cloud, the k-nearest neighbor algorithm is used to search for adjacent points in its local neighborhood (the value of k is dynamically adjusted based on the density of triangular patches, ranging from 16 to 64); spatial distribution features are calculated based on the neighborhood point set, and the normal vector and curvature information of the local surface are extracted through principal component analysis; the covariance matrix is decomposed into the product of a scaling factor and a rotation matrix, where the scaling factor is positively correlated with the local curvature (the scaling coefficient in high curvature areas is reduced to 0.3–0.5 times), while the rotation matrix remains consistent with the normal vector of the triangular patches.

The ellipsoid parameters generated by this method accurately reflect the geometric details of the cultural relic's surface, providing a precise geometric foundation for visibility judgment in the subsequent rendering stage.

4.2. Spherical harmonics compression

Spherical harmonics are used to encode the radiometric properties (such as color and lighting response) of a 3D Gaussian ellipsoid, and its higher-order parameters (typically using 4th-order spherical harmonics with 16 coefficients) are the main component of the data volume. Given the rich texture but strong spatial correlation of cultural relic models, this study proposes a three-level compression strategy to reduce network transmission load.

Firstly, layered quantization is applied to the spherical harmonic coefficients: the low-frequency components (0–1 order) are retained with 12-bit precision to maintain basic color fidelity, while the high-frequency components (2–3 order) are reduced to 8 bits, exploiting the human eye's low sensitivity to high-frequency lighting changes to reduce redundancy. Secondly, order reduction is employed by analyzing the surface lighting reflection characteristics of the artifact. In regions with gentle curvature, the 4th-order spherical harmonic function is reduced to 2nd order (preserving 5 coefficients), and full parameters are only retained in high-curvature detail areas. The ratio of order reduction is dynamically adjusted based on local curvature (with a threshold set at 0.02). Finally, entropy coding based on clustering is introduced. The quantized coefficients are subjected to K-means clustering ($K = 128$), and Huffman coding is used to compress the cluster indices, further reducing data redundancy.

5. 3D Gaussian splatting rendering based on WEBGL

To achieve 3D Gaussian splatting rendering in a WEBGL environment, a series of processing steps is required from the original 3D model to the final high-fidelity rendering effect, as follows: The original 3D model serves as the foundation for the entire rendering process, as shown in **Figure 1**. It contains complete geometric structure information about the artifact, including the shapes and positional relationships of various components, which are important bases for subsequent point cloud generation and rendering.

Texture images provide a guarantee for the authenticity of the rendering effect. The texture image presented in **Figure 2** contains detailed information such as colors and patterns on the surface of cultural relics. In the rendering process, this texture information needs to be accurately mapped to the corresponding point clouds to achieve realistic visual effects.

The conversion from the original model to the point cloud is one of the key steps. **Figure 3** shows a sparse

Figure 1. Original 3D model

Figure 2. Texture image

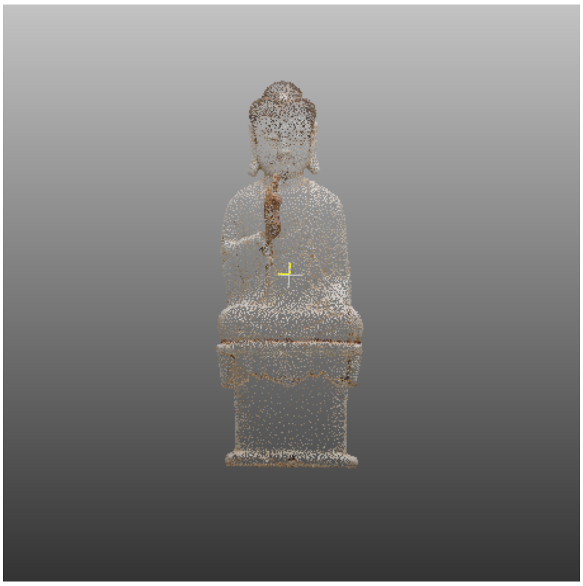


Figure 3. Sparse point cloud

Figure 4. Dense point cloud (without texture)

point cloud, which has a small number of points and can only roughly reflect the outline of the model, unable to well reflect the detailed features of the artifact. To achieve high-fidelity rendering, a dense point cloud needs to be generated.

Figure 4 shows a dense point cloud without texture. Through OBJ model parsing and point cloud generation methods, based on the area calculation and sampling of triangular patches, a large number of evenly distributed point clouds are generated. These point clouds can more accurately depict the geometric shape of cultural relics, laying a good geometric foundation for subsequent texture mapping and rendering.

Based on the dense point cloud, texture mapping is performed to obtain the textured dense point cloud shown in **Figure 5**. Through the point-to-point texture mapping process described in this paper, each pixel of the texture

Figure 5. Dense point cloud (with texture production)

Figure 6. Partial enlarged view of dense textured point cloud

image in **Figure 2** is accurately mapped to the corresponding position of the dense point cloud, enabling the point cloud to not only possess precise geometric information but also rich color and texture details, greatly enhancing the authenticity of the rendering.

Figure 6 shows a partial enlarged view of the dense textured point cloud, where fine textures and geometric details on the surface of the artifact can be clearly observed. This demonstrates that the hybrid sampling point cloud generation method and texture mapping process proposed in this paper can preserve high-precision details of the artifact model, meeting the demand for high-precision display in areas such as cultural heritage digitization.

In the WEBGL environment, efficient rendering is achieved through a mipmap texture optimization strategy (generating multi-resolution textures to adapt to different observation distances, balancing rendering effects and computational load) and a GPU-accelerated rasterization pipeline (utilizing parallel computing to quickly complete Gaussian ellipsoid projection and rasterization, achieving stable rendering of 60 frames per second to eliminate stuttering). Additionally, the spherical harmonic function parameters of 3D Gaussian splatting are quantized and compressed, reducing network transmission bandwidth to 52% of the original data to enhance the loading experience. In summary, this technology enables efficient and high-fidelity rendering of 3D artifact models based on 3D Gaussian splatting, providing strong support for the digital display of artifacts.

4. Conclusion

This study proposes a hybrid technical approach that combines traditional sampling with 3D Gaussian splatting. By improving the OBJ parsing process and implementing adaptive area-weighted sampling, it achieves the generation of a dense point cloud with adaptive triangle density. Innovatively, the study places ellipsoidal parameter estimation upfront, giving the point cloud inherent Gaussian distribution characteristics. Experimental verification shows that this method reduces the point cloud data volume by 38% at a 4096×4096 texture resolution. When combined with spherical harmonic function compression, the transmission bandwidth is reduced to 52%.

Additionally, relying on WebGL optimization, stable rendering at 60 frames per second is achieved. This research provides an efficient solution for high-precision digital display of cultural heritage, and further exploration can be done in dynamic scene adaptation and multi-scale model fusion technology.

Disclosure statement

The author declares no conflict of interest.

References

[1] Shi X, 2021, Focus: Smart Display Needs and Solutions for Museums. China Flights, (12): 61–63.

[2] Zhu D, Chen M, Mao Y, et al., 2025, Research Status and Challenges of 3D Gaussian Splatting Technology in Scene Reconstruction. Integrated Technology, 14(04): 1–20.

[3] Tong W, 2024, Optimization and Implementation of 3D Model Rendering Based on OpenGL, dissertation, Xidian University.

[4] Chen J, Huo Z, Yan Q, et al., 2024, Development Status of 3D Model Reconstruction System Based on NeRF Algorithm. Computer Knowledge and Technology, 20(30): 1–3.

[5] Chen H, Yan Z, Li W, 2025, Research on the Construction Method of Refined Single 3D Model Based on Oblique Photography Technology. Surveying and Mapping and Spatial Geographic Information, 48(S1): 175–177.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

High-Frequency Stable Wireless Amplitude Modulation System Based on a Pierce Circuit

Huiwen Xu*

School of Electronic Information, Soochow University, Suzhou 215001, Jiangsu, China

**Author to whom correspondence should be addressed.*

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: This paper designs a high-frequency stable wireless amplitude modulation (AM) system based on a Pierce circuit. The system utilizes an oscillator and comparator to generate a 20 kHz square wave with an adjustable duty cycle, combined with a 41 MHz carrier wave produced by a passive crystal oscillator Pierce circuit. A 100% modulation index amplitude modulation is achieved through the AD835 multiplier. The modulated signal is amplified by a power amplifier circuit and transmitted wirelessly via the transmitter antenna. Upon reception, the signal undergoes two-stage high-frequency amplification before passing through a Schottky diode envelope detector. The NE5532 shaping circuit then restores the square wave. Experimental results demonstrate reliable 11-meter transmission with carrier frequency deviation < 0.75% and demodulation error < 1%.

Keywords: Wireless transmission; Amplitude modulation; Pierce circuit; Low power consumption

Online publication: October 15, 2025

1. Introduction

With the continuous advancement of wireless communication technologies, amplitude modulation (AM) technology continues to demonstrate irreplaceable application value in medium and short-range signal transmission due to its advantages of simple circuit implementation and cost-effectiveness [1,2]. However, existing AM systems predominantly rely on digital auxiliary architectures. While these systems can improve transmission accuracy, their complex digital processing modules not only increase system power consumption but also raise hardware design complexity, making them inadequate for low-power and miniaturized applications [3]. In 2023, Chen *et al.* [4] developed an analog signal modulation system based on MATLAB and GUI that achieves digital AM modulation. However, research indicates that channel noise significantly impacts signal transmission, with noise reaching certain intensities potentially causing signal transmission failure.

To address the aforementioned challenges, this paper designs a high-frequency stable wireless AM system based on Pierce circuitry, with focused optimization of carrier generation, modulation, and demodulation circuits.

The system employs a passive crystal oscillator-based Pierce circuit to generate a stable 41 MHz carrier signal, utilizes an AD835 multiplier for 100% duty cycle amplitude modulation, and integrates Schottky diode envelope detection [5] with NE5532 shaping circuitry for signal reconstruction. This approach establishes a cost-effective, low-power, and reliable all-analog wireless AM solution, providing technical reference for medium-to-short range wireless communication scenarios.

2. System design

2.1. Overall design

The system's overall architecture is illustrated in **Figure 1**. The system first generates a 20 kHz square wave signal through a square wave generator circuit. This signal, combined with a 41 MHz carrier signal from the carrier generator circuit, is fed into an amplitude modulation circuit to produce an AM wave. After amplification by a power amplifier, the AM wave is transmitted wirelessly via the transmitting antenna. At the receiving end, the antenna captures the signal, which is then amplified by a high-frequency small-signal amplifier to extract the 41 MHz component. The extracted signal is demodulated using an envelope detection circuit to restore the original square wave. Finally, the signal undergoes amplification and reshaping through a shaping circuit to reconstruct the original 20 kHz square wave.

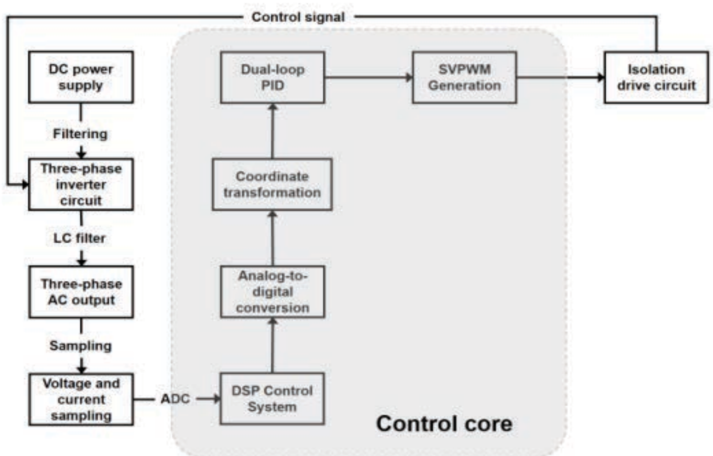


Figure 1. Overall system block diagram

2.2. Design of square wave oscillator circuit

The square wave oscillator circuit is shown in **Figure 2**. The square wave oscillator circuit first generates a rectangular pulse signal [6] from the inverter oscillator R_6 circuit (the circuit on the left side of **Figure 2**). The frequency of this signal can be adjusted through the potentiometer in **Figure 2**, and the calculation formula is:

$$f = \frac{1}{2\mathrm{RCln3}} = \frac{1}{2C_2\ (R_{f2} + R_6)} = 20\ \mathrm{kHz} \quad (1)$$

Take the C_2 = 10nF capacitance and substitute it into formula 1 to calculate:

$$R_{f2} + R_6 \approx 2.3\mathrm{k}\Omega \quad (2)$$

Set fixed resistor R_{f2} to 1 kΩ to protect the circuit, and it can be determined that the potentiometer R_6 needs to be set to approximately 2.3 kΩ. Although the calculation provides the theoretical value of the resistance, to allow precise adjustment of the final square wave frequency, a more stable fixed resistor is not used here; instead, potentiometer R_6 is chosen.

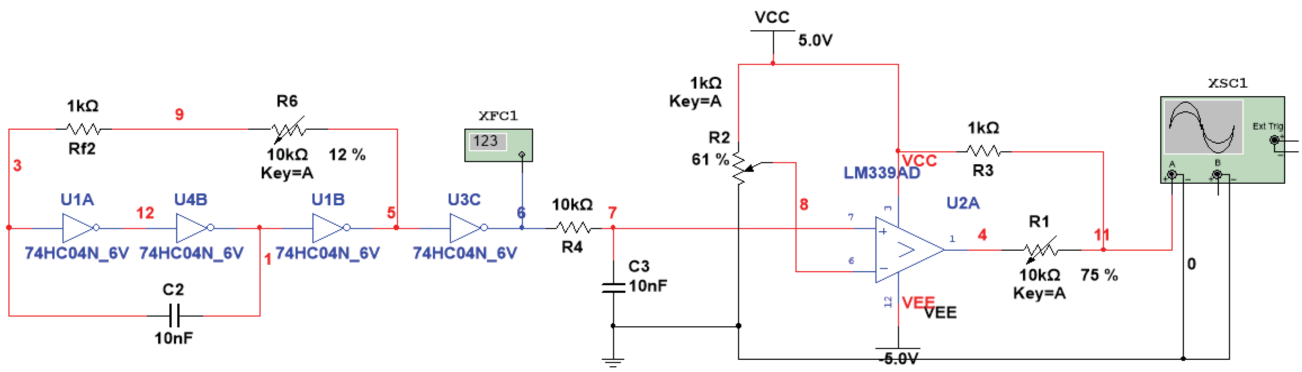


Figure 2. Square wave oscillator circuit

The inverter oscillator circuit utilizes the 74HC04N six-phase inverter chip. This component offers advantages such as low power consumption, high noise tolerance, fast response, and a wide voltage range. As the 74HC04N chip typically operates within a voltage range of 2 V to 6 V, the circuit supplies it with a 5 V power supply.

A comparator circuit is a device that compares the magnitude of currents or voltages between two input terminals. Typically composed of an operational amplifier in open-loop configuration, it features two input terminals and one output terminal. When the non-inverting terminal voltage exceeds the R_2 inverting terminal voltage, the chip outputs a high level; conversely, when the non-inverting voltage falls below the inverting voltage, it outputs a low level. As illustrated in **Figure 2**, the potentiometer regulates the duty cycle of the square wave signal by adjusting the inverting terminal voltage.

The comparator circuit utilizes the LM339D chip. This integrated device features four independent precision comparator channels, supporting both single and dual power supply configurations. With a typical input bias current of 25 nA and minimal quiescent current consumption, it is particularly suitable for applications requiring multiple comparison channels, low power consumption, and stable performance.

2.3. Carrier oscillation circuit design

A passive crystal oscillator (crystal resonator) is a kind of resonant device based on a quartz crystal. As a passive and high Q value element, it does not need an external power supply, and uses the inherent frequency characteristics of the quartz crystal to generate a stable frequency signal.

In a Pierce oscillator, the S9018 high-frequency transistor serves as the core active component, providing essential inversion functionality, gain, and amplification capabilities. Its superior high-frequency characteristics enable it to drive high-frequency crystals, which is why this design's carrier oscillation circuit specifically employs the S9018 transistor model. The detailed circuit configuration is illustrated in **Figure 3** [7].

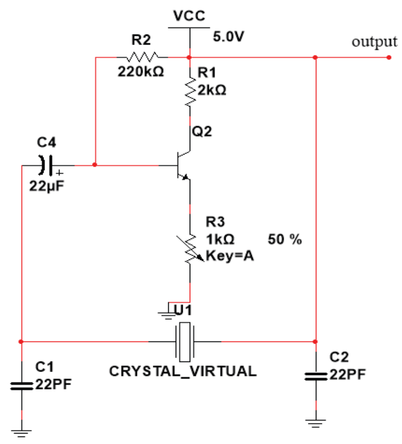


Figure 3. Carrier oscillator circuit

2.4. AM circuit design

The AD835 is a complete four-quadrant voltage-output analog multiplier designed to generate linear products of X and Y voltage inputs. With its wide bandwidth, it handles high-frequency signals effectively for high-speed signal processing applications. Its low-noise characteristic ensures minimal signal distortion in experimental scenarios requiring precise signal handling. Additionally, the device features high input impedance and low output impedance, effectively reducing interference from cascaded circuit stages. These advantages make the AD835 multiplier an ideal choice for amplitude modulation circuitry [8]. The amplitude modulation circuit is illustrated in **Figure 4**.

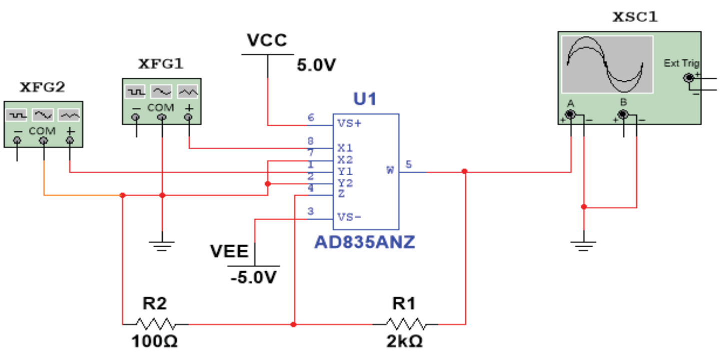


Figure 4. Amplitude modulation circuit

In an amplitude modulation circuit, the carrier and modulated signal are fed into the AD835's X and Y input terminals, with outputs at terminal W. Since both input signals have a range of ±1V, the carrier must undergo DC blocking before entering the multiplier. Through this component, the modulator achieves amplitude modulation of the carrier signal, with a modulation factor reaching up to 100%. The relationship between the multiplier's output signal and input signal is defined as:

$$W=XY+Z \tag{3}$$

2.5. Power amplifier circuit design

The power amplifier circuit is shown in **Figure 5**, which uses the S9018 transistor [9]. The characteristic frequency of S9018 is usually 100 MHz or higher. This means that in high-frequency circuits, S9018 can maintain high gain and stability, thus ensuring the quality of signal transmission.

Meanwhile, the S9018 features a low noise figure of approximately 2 dB, offering significant advantages in low-noise amplifier design. This capability effectively suppresses noise interference during signal processing, thereby enhancing signal clarity and quality. With its outstanding high-frequency performance, the S9018 is widely used in RF front-end circuits to amplify input high-frequency signals, thereby boosting signal strength and improving reception performance.

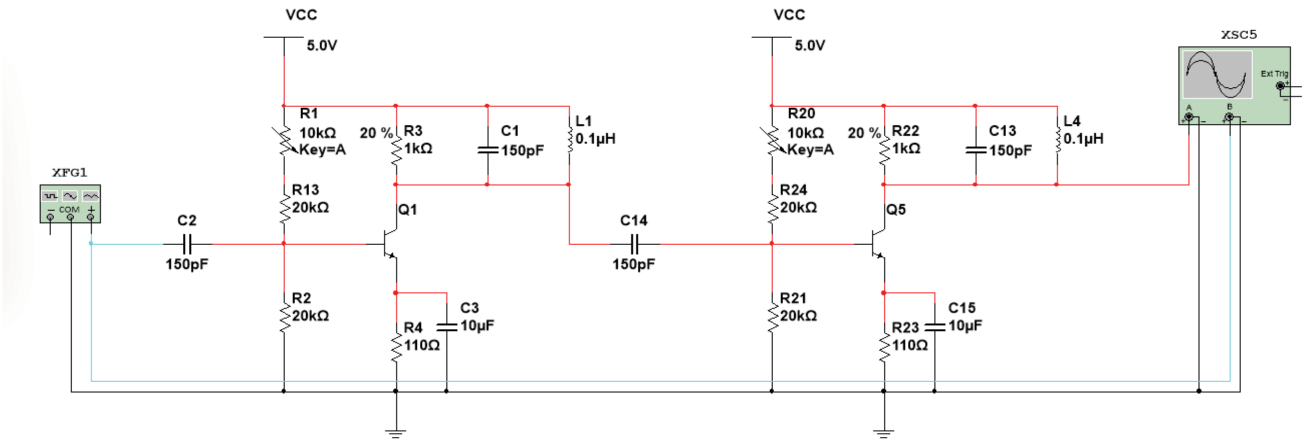


Figure 5. Power amplifier circuit

The circuit shown in **Figure 5** adopts a two-stage amplifier structure. Since the first and second stages share an identical circuit configuration, only the first stage will be analyzed below. In this circuit, C_1 and L_1 form a parallel resonant frequency-selecting network, which selectively amplifies the 41 MHz AM wave signal. Resistor R_1 serves as the base bias resistor for the transistor, with its primary functions being to provide an appropriate static operating point for the transistor to operate in the amplification region, while also offering a certain degree of input impedance matching.

2.6. Detection circuit design

The detection circuit, as shown in **Figure 6**, employs diodes for envelope detection based on their nonlinear characteristics and unidirectional conductivity. The AM wave consists of a high-frequency carrier and its low-frequency modulated signal. When the AM signal exceeds the diode's threshold voltage, the diode conducts, allowing current passage. At this point, the diode acts as a switch to extract the positive portion of the AM signal. Conversely, when the signal's amplitude falls below the diode's threshold voltage, the diode blocks the signal, thereby intercepting the negative portion. After diode detection, the resulting signal contains both high-frequency and low-frequency components. To extract the modulation signal that reflects the AM wave's envelope, this pulsating current must pass through a low-pass filter. The low-pass filter removes high-frequency components while retaining low-frequency ones, effectively producing the AM waveform envelope. This can be achieved using a simple RC filter circuit [10].

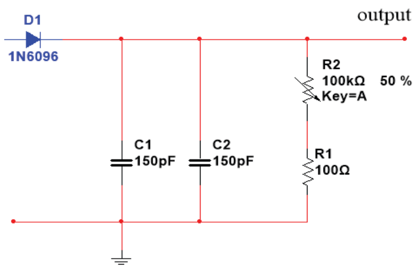


Figure 6. Detection circuit

The 1N6096 diode is a Schottky rectifier diode. A defining feature of Schottky diodes is their low forward voltage drop, typically around 0.45 V. This characteristic allows minimal voltage drop across the diode during forward conduction, effectively reducing power consumption and enhancing circuit efficiency. For these reasons, the 1N6096 diode is widely adopted in envelope detection circuits.

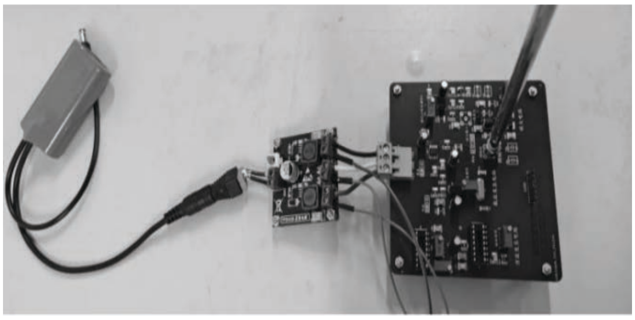
In **Figure 6**, components R_1, R_2, C_1, andC_2 form a low-pass filter. The cutoff frequency of this filter should be between 20 kHz and 41 MHz, and the calculation formula is:

$$f = \frac{1}{RC} = \frac{1}{(R_1 + R_2) * (C_1 + C_2)} \tag{4}$$

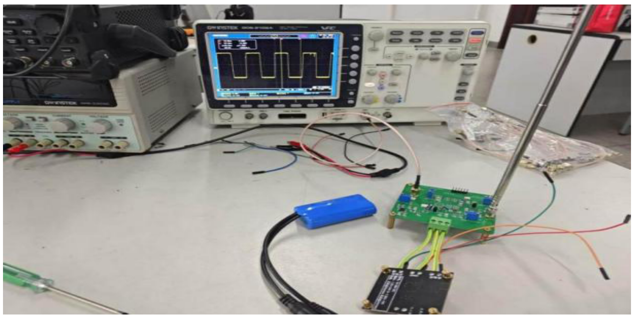
Set $C_1 = C_2 = 150pF$, $R_1 = 100\ \Omega$, and R_2 as an adjustable potentiometer of 100 kΩ. The detection circuit significantly affects the waveform of the receiver circuit. Therefore, an adjustable potentiometer is required to adapt the receiver to different transmission distances and achieve optimal detection performance.

3. Analysis of experimental results

The physical diagrams of the transmitter and receiver are shown in **Figure 7**.



(a) transmitter



(b) receiver

Figure 7. Physical diagram

The carrier test waveform is shown in **Figure 8**.

Figure 8. Carrier waveform

$$\frac{f_{\text{test}} - f_{\text{anticipated}}}{f_{\text{anticipated}}} = \frac{41.237 - 41}{41} * 100\% \approx 0.58\% \tag{5}$$

As shown in **Figure 8**, the actual carrier frequency generated by the circuit is 41.237 MHz, and the calculated carrier frequency deviation is:

The carrier output waveform exhibits a frequency deviation of merely 0.58%, significantly lower than the typical 1–2% range in conventional LC oscillator circuits and demonstrating superior performance. This achievement highlights the core advantages of Pierce oscillators integrated with quartz crystals in high-frequency stability.

The amplitude-modulated wave test waveform is shown in **Figure 9**.

Figure 9. Amplitude modulation waveform

The adjustment magnitude is calculated as:

$$m_a = \frac{V_{\max} - V_{\min}}{V_{\max} + V_{\min}} \tag{6}$$

As can be seen from **Figure 9**, the values of $V_{max} \approx 2.1$ V and $V_{min} \approx 0$ are obtained. Substituting these into Equation 7, the result is calculated as follows:

$$m_a = \frac{V_{max} - V_{min}}{V_{max} + V_{min}} = \frac{2.1 - 0}{2.1 + 0} = 1 \quad (7)$$

As shown in Equation 7, the modulation depth reaches 100% at this point, demonstrating the circuit's superior modulation capability. The waveform exhibits a clear envelope without flat-top distortion or overmodulation, indicating that the AD835 multiplier maintains excellent linearity and dynamic range even at 41 MHz operating frequency, with minimal crosstalk between X and Y channels. While achieving maximum transmission power efficiency through 100% modulation depth, this configuration also imposes higher requirements on the low-dropout characteristics of the receiver's detector circuitry.

The amplified waveform is shown in **Figure 10**.

Figure 10. Amplified waveform

As shown in **Figure 10**, the amplified square wave signal exhibits a peak-to-peak voltage of approximately 8.00 V, with a duty cycle of about 50% and a frequency of 20 kHz—nearly identical to the original signal. The low forward voltage drop characteristic of the Schottky diode 1N6096 enables efficient demodulation of small amplitude AM signals while minimizing detector loss. The waveform demonstrates sharp edges without oscillation and stable high-level performance without collapse, indicating that the envelope detection circuit successfully filtered out the 41 MHz carrier. Furthermore, subsequent amplification and shaping circuits effectively eliminated noise interference. The final output waveform perfectly reproduces the original 20 kHz square wave's amplitude, frequency, and duty cycle characteristics, achieving high-fidelity restoration of the modulated signal.

The transmitter and receiver are adjusted to a distance of 11 meters, and the final output waveform of the receiver is shown in **Figure 11**.

Figure 11. Receiver output waveform

As shown in **Figure 11**, the peak value of the square wave has decreased to approximately 6 V due to path loss in free-space electromagnetic propagation. The received signal strength attenuates by about 2.5 dB, consistent with theoretical expectations. Crucially, both the frequency and duty cycle of the output waveform remain stable despite the signal attenuation, showing no significant distortion or frequency drift.

The experimental results clearly show that the system can achieve reliable transmission within 11 meters, and all key indicators meet or exceed the design expectations, which verifies the feasibility and superiority of the full analog architecture in high-frequency radio amplitude modulation applications.

4. Conclusion

This paper designs a fully analog architecture-based wireless AM system, constructing circuits for square wave generation, carrier oscillation, amplitude modulation, power amplification, detection, and amplification to achieve signal modulation, transmission, reception, and demodulation. Experimental results demonstrate excellent system performance: A 20 kHz square wave exhibits a frequency error of 0.08% with a duty cycle approaching 50%; a 41 MHz carrier demonstrates 0.58% frequency deviation and good stability; the amplitude modulation circuit achieves 100% modulation depth; reliable transmission over 11 meters is attainable, maintaining stable square wave frequency and duty cycle at 11 meters with demodulation errors < 1%. This study validates the feasibility and effectiveness of the fully analog architecture for wireless AM transmission, providing valuable references for related field system designs. Future work could focus on optimizing circuit structures and exploring transmission performance under longer distances and more complex environments to expand its application scope.

Disclosure statement

The author declares no conflict of interest.

References

[1] Lai S, 2024, Analysis and Application Practice of Digital Technology for Amplitude Modulation Broadcast Transmitters. TV Technology, 48(10): 102–105.

[2] Jiang Z, Li F, Xu Y, 2025, Fast Amplitude Modulation Mode Decomposition for Adaptive and Robust Extraction of Rolling Bearing Compound Fault Characteristics, in IEEE Sensors Journal, 25(15): 28127–28136.

[3] Wen C, Li Y, 2020, Technical Features and Maintenance of Digital Amplitude Modulation Medium Wave Broadcast Transmitter. Electronic Manufacturing, (22): 71–72 + 18.

[4] Chen J, Lv J, 2023, Design of Analog Signal Modulation System Based on MATLAB/GUI. Modeling and Simulation, 12(4): 3450–3463.

[5] Li X, 2025, Design of Analog Signal Transmission Circuit in Rongjiayuan Driver, dissertation, University of Electronic Science and Technology of China.

[6] Yang Y, 2021, Decomposition and Synthesis of Periodic Square Wave Signals under Engineering Practice. Journal of Electrical and Electronic Teaching, 43(05): 138–140 + 182.

[7] Shao Y, Lu Z, 2021, Design of an 8 MHz Crystal Oscillator Circuit Based on Pierce Oscillator. Electronics & Packaging, 21(02): 96–101.

[8] Deng J, Xie L, Meng Z, et al., 2025, A “Digital-to-Analog” Modulation and Demodulation Experimental System Based on Multiplier Devices. Journal of Electrical and Electronic Teaching, 47(01): 218–221.

[9] Zhang H, Wang P, 2021, Simulation Design of RC-Coupled Common-Emitter Amplifier Circuit. Applied Science, 48(02): 19–22.

[10] Hou C, Zhang J, Yuan Y, et al., 2022, Reliability Demodulation Algorithm Design for Phase-Generated Carrier Signal. IEEE Transactions on Reliability, 71(1): 127–138.

Publisher’s note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Research on Quality Assurance and Testing Strategies of Quality Engineers in Software Product Development

Jialun Deng*

Shenzhen Institute of Measurement and Quality Inspection, Shenzhen 518000, Guangdong, China

**Author to whom correspondence should be addressed.*

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: Quality engineers play a key role in software product development, covering various stages such as requirements analysis, design, coding, testing, and delivery. Its responsibilities include formulating quality standards, writing test cases, conducting functional and performance tests, and optimizing the product based on feedback. In government procurement projects, quality evaluation focuses on process compliance, security, and functional compatibility. KPI evaluation trees are commonly used for quantitative assessment, and a dynamic adjustment mechanism for indicators needs to be established to cope with complex demands. In addition, risk-driven testing and agile development should be combined to set up quality access control to ensure that each iteration version meets expectations. The multi-dimensional quality assurance and verification scoring mechanism can effectively enhance product reliability and reduce project risks.

Keywords: Quality engineer; Government procurement; Software product quality

Online publication: October 15, 2025

1. Introduction

With the rapid development of information technology, software products are increasingly widely applied in various fields. In government procurement projects, software quality is even more of a concern. The "Measures for Promoting the Development of Small and Medium-sized Enterprises through Government Procurement" promulgated in 2020 emphasizes the fairness and impartiality of the procurement process as well as strict requirements for product quality. Quality engineers undertake multiple responsibilities in the software product development process. From requirement analysis to the final product evaluation, they need to ensure that the quality meets the standards. Meanwhile, in view of the special quality evaluation characteristics of government procurement projects, such as strict acceptance standards, service continuity, and compliance requirements, quality engineers need to establish a reasonable quality assurance system. On this basis, establishing a scientific KPI evaluation tree, a dynamic adjustment mechanism for indicators, and a risk-driven testing strategy, etc., is of vital importance for ensuring the quality of government procurement software products.

2. Functional positioning of quality engineers and construction of a quality assurance system

2.1. Quality assurance role in software product development

Quality engineers play an indispensable role in software product development, covering the entire process from requirement analysis, design, coding, testing, to delivery. During the requirements analysis stage, they closely communicate with stakeholders, sort out and verify the accuracy and completeness of the requirements, and formulate clear requirements documents to avoid subsequent development problems caused by ambiguous or missing requirements, ensuring that the development direction is consistent with expectations. During the design review stage, quality engineers need to conduct in-depth reviews of the software architecture design, module division, and technology selection to assess whether they meet quality standards and project requirements, and put forward optimization suggestions to enhance the system's maintainability and scalability. During the coding and process auditing stages, they regularly inspect the standardization of the development process, supervise the implementation of coding standards, promptly identify and correct potential issues, and prevent the accumulation of defects. In government procurement projects, the compliance review function of quality engineers is particularly prominent. They need to strictly compare with relevant regulations and standards, and focus on reviewing the performance of software in terms of security, reliability, and functional compatibility to ensure that the products meet the strict requirements of government procurement. They supervise the quality control throughout the project's entire lifecycle by formulating detailed quality inspection checklists and audit plans, ensuring the smooth progress of the project and achieving the expected quality goals. This multi-dimensional and full-process quality supervision and review mechanism not only enhances the reliability and stability of software products but also provides a solid guarantee for the compliance and successful implementation of government procurement projects [1].

2.2. Characteristics of quality evaluation for government procurement projects

The quality evaluation of government procurement projects has unique characteristics and is different from that of commercial projects. In terms of acceptance criteria, government procurement projects have formulated stricter and more detailed regulations, covering clear technical indicators and functional requirements, such as system performance, data security, user experience, etc. Each one needs to be verified to ensure that the product fully complies with the contract agreement [2]. Service continuity is an important evaluation dimension. Government projects usually require suppliers to provide stable and reliable operation and maintenance support over a relatively long period of time to ensure the long-term normal operation of the system and meet the continuity requirements of public services. Compliance requirements are particularly crucial. Projects must strictly abide by relevant laws, regulations, and policies, such as data protection, privacy, security, and the transparency of the procurement process, to ensure the legality and fairness of the implementation process. These characteristics place higher demands on quality engineers, who need to fully integrate these dimensions into the construction of the quality assurance system and design targeted quality control strategies. For instance, by formulating detailed acceptance criteria, establishing a service continuity assessment mechanism, and strengthening the compliance review process, quality engineers can effectively address the specific requirements of government procurement projects. This multi-dimensional quality evaluation system not only ensures the compliance and functional realization of the project, but also enhances the reliability of software products and user satisfaction, providing a solid foundation for the smooth implementation and long-term operation of government procurement projects.

3. Design of the scoring and rating system for contract performance quality

3.1. Construction of a multi-dimensional evaluation index system

The quality evaluation of contract performance for government procurement of software products needs to establish a multi-dimensional KPI index system, covering four dimensions: functionality, performance, security compliance, and service response, to ensure the comprehensiveness and scientific nature of the assessment. Functional indicators focus on whether the software meets the user's needs, involving the completeness of function realization, the accuracy of operation, and the friendliness of user experience [3]. Performance parameters focus on the operational efficiency of the system. By quantifying response time, throughput, and concurrent processing capabilities, they evaluate the software's performance in real-world scenarios. The security compliance dimension requires software to comply with relevant regulations and technical standards, safeguard data privacy and system security, and prevent potential risks. The service response dimension examines the problem-handling capabilities of suppliers, including the timeliness of fault response and the effectiveness of solutions. The KPI evaluation tree constructed by integrating these four dimensions can objectively reflect the quality level of software products and provide data support for the scoring and rating of contract performance. This system enhances the transparency and fairness of evaluation by refining indicators and conducting quantitative assessment, providing a scientific basis for the quality management of government procurement projects and ensuring that the procurement outcomes meet the expected goals.

3.2. Dynamic weight distribution model

The quality evaluation of contract performance for government procurement of software products should adopt a dynamic weight distribution model to adapt to the characteristics of different project stages. The Analytic Hierarchy Process (AHP) is the core method for constructing this model. By constructing a hierarchical structure, the evaluation indicators are decomposed into multiple levels and factors, and pairwise comparisons are made layer by layer to generate a judgment matrix, calculate the weights of each indicator, and establish the initial weight distribution [4]. This method ensures that the weight settings are scientific and reasonable, reflecting the relative importance of each indicator. In practical applications, the weights need to be dynamically adjusted according to the project stages to highlight the quality priorities at each stage. For instance, in the early stage of development, the accuracy of requirements analysis and design is crucial for subsequent development, and its weight should be appropriately increased. During the testing phase, the coverage rate of test cases and the defect discovery rate directly affect product quality, and corresponding weights need to be added. Through differentiated weight configuration, the evaluation system can accurately reflect the actual situation of performance quality at each stage, enhancing the pertinence and flexibility of the assessment. This model not only enhances the objectivity of quality scoring and rating but also provides a scientific basis for the dynamic management of government procurement projects, ensuring that quality evaluation is closely aligned with project progress [4].

4. Optimization of quality assurance strategies and test plans

4.1. Formulation of test strategies based on rating results

4.1.1. Risk-driven testing priority classification

Identifying key risk domains based on quality rating results is an important basis for prioritizing risk-driven testing. By conducting quality ratings on each module or function of the software, the degree of risk is clearly defined. For high-risk domains, they often involve core functions, safety-critical parts, and areas where problems frequently occur [5]. Once these areas malfunction, it may have a serious impact on the entire software system. Therefore, for these key risk domains, test resources should be prioritized for allocation, and more comprehensive and in-depth test cases should be designed. Establishing a test case priority matrix is a method to further quantify and standardize the priority division. In the matrix, factors such as the degree of risk, functional importance, and the possibility of problems are comprehensively considered to assign corresponding priority values to each test case, thereby ensuring that the testing work can be efficiently and orderly focused on the key parts, and improving the effectiveness and efficiency of the testing.

4.1.2. Optimization of the automated testing framework

The formulation of test strategies based on rating results should closely revolve around the quality characteristics and risk factors of software products. Based on the dynamic monitoring rating indicators, precisely locate the key test areas and prioritize the allocation of test resources. For high-risk modules, increase the coverage and execution frequency of test cases to ensure that potential problems can be exposed in a timely manner [6].

In terms of optimizing the automated testing framework, emphasis should be placed on enhancing the flexibility and scalability of the framework. Adapt to different software architectures and functional requirements, and be capable of conveniently integrating new testing tools and technologies. Optimize the writing and maintenance mechanism of test scripts, increase the reuse rate of scripts, and reduce testing costs. At the same time, enhance the framework's management capacity for test data to ensure its accuracy and completeness, providing strong support for the reliability of test results.

4.2. Quality assurance under the agile development model

4.2.1. Iterate the quality access control Settings

Under the agile development model, the iterative quality access control setting is a key mechanism to ensure the quality of software products. By embedding a quality rating threshold control system in the Continuous Integration and Continuous Deployment (CI/CD) process, the results of each iteration are strictly evaluated. Quality engineers need to set clear quality standards and thresholds, covering dimensions such as functional integrity, performance, and safety compliance. When the development team submits code to the version control system, the build and testing process is automatically triggered, and quality ratings are conducted based on the test results and predefined metrics. If the iteration results do not meet the threshold requirements, the system will prevent them from entering the next development stage to ensure that only the code that meets the quality standards can be advanced. This mechanism effectively identifies and fixes potential defects, preventing problems from accumulating to subsequent stages and thereby enhancing the overall quality of the software. The quality rating threshold needs to be dynamically adjusted based on the project characteristics and business requirements to adapt to the rapidly changing development environment and enhance the flexibility and pertinence of the assessment. This approach, through automated and quantitative means, strengthens quality control in agile development, providing a solid guarantee for the efficient implementation and reliability of government procurement software projects [7].

4.2.2. Construction of defect prediction model

Under the agile development model, building a defect prediction model is an important strategy for improving software quality assurance. Quality engineers collect historical rating data through the system, covering key information such as the number, severity, and occurrence stage of defects, and conduct in-depth analysis of the patterns and designs contained in the data. Defect density prediction models are developed by adopting data mining techniques and machine learning algorithms, such as decision trees, random forests, or neural networks. This model accurately predicts the distribution and density of potential defects by using input parameters such as project characteristics, code complexity, and development stage. The prediction results provide a scientific basis for the allocation of test resources, guiding quality engineers to focus on high-risk modules, optimize the design and execution frequency of test cases, and thereby enhance test efficiency and coverage. By identifying risk areas in advance, the model effectively reduces the probability of missed defect detection, lowers the cost of later repair, and enhances the stability and reliability of software products. To adapt to the rapid iteration characteristics of agile development, the model needs to update the training data regularly to ensure that the prediction accuracy is consistent with the actual requirements of the project. This method provides data-driven decision support for the quality management of government procurement software projects, significantly enhancing the pertinence and effectiveness of quality assurance [8].

5. Practice of contract performance in government procurement projects

5.1. Analysis of pain points in contract performance quality evaluation

5.1.1. Lack of quantitative assessments of deliverables

Subjective evaluation issues often exist in the traditional acceptance process, which is particularly prominent in the practice of contract performance for government procurement projects. Due to the lack of objective quantitative assessment standards, acceptance personnel often evaluate the quality of deliverables based on their own experience and subjective judgment, which is very likely to cause disputes. Different acceptance personnel may give different evaluations of the same deliverable, leading to conflicts between the supplier and the purchaser. At the same time, subjective evaluation is difficult to accurately measure whether the deliverables truly meet the contractual requirements, which may allow some products that do not meet the quality standards to pass the acceptance, while some products that meet or even exceed the standards do not receive the recognition they deserve. This situation has seriously affected the fairness and accuracy of the evaluation of contract performance quality and hindered the smooth implementation of government procurement projects [9].

5.1.2. Weak process traceability mechanisms

In the practice of contract performance in government procurement projects, there are weak points in the process traceability mechanism for evaluating the quality of contract performance. On the one hand, the relevant data records are incomplete and cannot comprehensively reflect each link of the performance process, making it difficult to accurately trace the root cause of the problem [10]. For instance, the operation records of some key nodes are missing, which makes it difficult to have a sufficient basis when evaluating quality. On the other hand, the correlation and integration of data are poor. Data at different stages are scattered and lack effective correlation analysis methods, making it impossible to form a complete process traceability chain. This makes it difficult to analyze the causes from the perspective of the entire performance process after quality problems are discovered, and thus, it is impossible to take targeted improvement measures, which affects the accuracy and effectiveness of the evaluation of contract performance quality.

5.2. Implementation of the dynamic quality evaluation mechanism

5.2.1. Milestone node rating system

In the practice of contract performance in government procurement projects, the implementation of dynamic quality evaluation mechanisms is of vital importance, among which the milestone node rating system is a key link. Quantitatively assess the progress of the project by setting clear milestone nodes. For instance, in software product development, the completion of the design phase can be regarded as a milestone. At this node, the design results are strictly reviewed in accordance with the pre-set quality standards. At the same time, link the result of this quality review to the contract payment. If the quality meets the standards during the design stage, normal payment shall be made as stipulated in the contract. If there are quality issues, the payment amount will be adjusted according to the severity of the problem, or the payment will be made after rectification is required. This mechanism prompts suppliers to attach great importance to quality at each milestone, ensuring that the overall quality of the project meets the requirements and guaranteeing the smooth implementation of government procurement projects.

5.2.2. Quality guarantee deposit calculation model

Establishing a dynamic deduction algorithm for the guarantee deposit based on the quality rating coefficient is the key to the quality guarantee deposit calculation model. By quantitatively assessing various quality indicators during the contract performance process of government procurement projects, the quality rating coefficient is determined. This coefficient takes into account multiple factors, such as the functionality, reliability, and ease of use of the product or service, comprehensively. Set the corresponding deposit deduction ratio based on different quality rating results. For instance, when the quality rating is higher, the deduction ratio is lower. When the quality rating is low, the deduction ratio shall be increased accordingly. This dynamic deduction algorithm can motivate suppliers to actively improve the quality of contract performance, ensure the smooth implementation of government procurement projects, and also provide a scientific basis for the reasonable calculation and management of quality guarantee deposits.

5.3. Design of a quality performance feedback system

5.3.1. Construction of supplier quality profile

The construction of supplier quality profiles is an important part of the design of quality performance feedback systems in the contract performance practice of government procurement projects. It is necessary to integrate multi-dimensional data to comprehensively depict the quality of suppliers. On the one hand, collect various indicator data during the contract performance process, including product quality, delivery timeliness, after-sales service, etc. On the other hand, standardize the data to ensure comparability among different indicators. By using data analysis techniques to mine the potential information behind the data, such as the strengths and weaknesses of suppliers. Based on this information, a quality profile is constructed to visually present the quality status of the supplier. Portraits can serve as an important basis for government purchasers to evaluate suppliers, which is conducive to making reasonable procurement decisions. At the same time, it can also prompt suppliers to improve their own shortcomings in a targeted manner and enhance the overall quality level.

5.3.2. Quality improvement decision support

The application of rating data mining technology can conduct an in-depth analysis of the quality and performance of contract fulfillment. By collecting and organizing a large amount of rating data, including multi-dimensional information such as the product quality, service level, and delivery time of suppliers. By using data mining algorithms, such as association rule mining and cluster analysis, the key factors influencing quality performance and the potential relationships among them can be identified. For instance, it may be found that certain specific supplier characteristics are associated with high-quality performance. Based on these mining results, generate a quality improvement suggestion report. The report elaborates in detail on the identified issues, related factor analysis, and targeted improvement suggestions, providing strong data support and decision-making basis for quality improvement decisions. It helps the purchasing party better manage the contract performance process and enhance overall quality performance.

6. Conclusion

Quality engineers are of vital importance in software product development. In terms of quality assurance, it is necessary to start from multiple dimensions, including precise control of requirements, strict supervision of the development process, and comprehensive assessment of the final product. For testing strategies, it is necessary to comprehensively consider the applicable scenarios of different testing methods, such as the reasonable application of unit testing, integration testing, system testing, etc. Meanwhile, the verification scoring and rating mechanism has practical value in enhancing the transparency of the quality of government procurement projects and strengthening the binding force of contracts, providing references for related fields. Future research directions can focus on building a government-enterprise quality data sharing platform. Quality engineers should also master data analysis skills to better adapt to the new requirements of digital quality management, thereby more effectively ensuring quality and implementing testing strategies in software product development.

Disclosure statement

The author declares no conflict of interest.

References

[1] Zhang Y, 2010, Design and Implementation of Online Problem Sheet System, dissertation, Beijing University of Posts and Telecommunications.

[2] Wang D, 2016, Research on the Construction and Application of Quality Engineer Competency Model, dissertation, Jiangsu University of Science and Technology.

[3] Cai C, 2012, Research on Competency Model of Quality Engineers in Automobile Enterprises, dissertation, Wuhan University of Technology.

[4] Guo W, 2018, Research on Quality Control Strategies in the Development Process of ADCR Software Products, dissertation, Southeast University.

[5] Lou Y, 2014, Research on the Application of Internet of Things Technology in Intelligent Buildings, dissertation, University of Electronic Science and Technology of China.

[6] Wang R, 2017, Discussion on Software Product Quality Comparison Testing and Comparison Schemes. China High-Tech Enterprises, 2017(12): 2.

[7] Pang Y, 2013, Analysis and Design of Software Project Management Information System Based on CMM13 Level, dissertation, Yunnan University.

[8] Gao X, 2019, Analysis of Quality Management and Risk Response in Software Testing Process. Wireless Interconnection Technology, 16(13): 31–32.
[9] Peng Z, 2013, Analysis of Test-Driven Development and Software Quality Assurance. Journal of Quanzhou Normal University, 31(6): 90–94.
[10] Lin W, Jing H, 2013, Software Product Testing Methods and Strategies. Science & Technology Vision, 2013(10): 37.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Research on Engineering Technological Innovation and Risk Management Strategies in Electric Power Construction

Runxian Zhou*

Shenzhen 518000, Guangdong, China

**Author to whom correspondence should be addressed.*

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: This article focuses on electric power engineering and expounds the development characteristics and applications of new electric power engineering technologies, including technologies such as smart grids and digital design platforms. It explores the identification and classification of risk elements in electric power engineering and analyzes the deficiencies of traditional risk assessment methods. It introduces the applications of new technologies such as intelligent sensor networks in risk management, proposes a dual-driven model of technology and management, and looks forward to the application prospects of artificial intelligence and blockchain technologies.

Keywords: Electric power engineering; Technological innovation; Risk management

Online publication: October 15, 2025

1. Introduction

With the increasing demand for energy transformation, the white paper "China's Energy Development in the New Era" released in 2020 emphasized the importance of energy technology innovation and risk management. The intelligent grid technology, digital design platform, and new energy access technology in the field of power engineering are constantly developing, while progress has also been made in BIM full lifecycle management, IoT monitoring technology, and so on. These technological innovations not only improve the efficiency and quality of power engineering, but also bring new risks, such as design defects, construction safety, equipment failures, and environmental factors. Therefore, studying engineering technology innovation and risk management strategies in power construction is of great significance.

2. Current status of technological innovation and development in power engineering

2.1. Development characteristics of new power engineering technology

New power engineering technologies, such as smart grid technology, digital design platforms, and new energy access technology, have shown many development characteristics. Smart grid technology has achieved intelligent monitoring, control, and optimized scheduling of the power system, improving the reliability and stability of the power grid. The digital design platform utilizes advanced computer technology and software tools to accurately model and virtually design power engineering, effectively reducing design errors and improving design efficiency. The new energy access technology adapts to the needs of energy transformation, solves the intermittent and fluctuating problems of new energy generation, and ensures the stable access and efficient utilization of new energy electricity [1]. The innovative breakthroughs in these technologies not only promote the progress of power engineering technology but also significantly improve construction efficiency, laying a solid foundation for the sustainable development of the power industry.

2.2. Key technological breakthrough direction

BIM full lifecycle management has made certain progress in power engineering, covering various stages from design, construction, to operation, which can effectively improve engineering efficiency and quality. By creating a three-dimensional information model, collaborative work among all parties involved can be achieved to detect and solve potential problems in advance. The Internet of Things monitoring technology has also made progress, which can monitor the operating status of power equipment in real time, such as temperature, humidity, and other parameters, providing a basis for equipment maintenance and management. The application of new insulation materials is another key direction [2]. These materials have better insulation performance and mechanical strength, which can improve the reliability and safety of power equipment, extend its service life, and have broad application prospects in the fields of high voltage, ultra-high voltage, and other power engineering.

3. Existing problems in the risk management of power engineering

3.1. Identification and classification of risk factors

The identification and classification of risk factors in power engineering are key to risk management. From the perspective of design defects, unreasonable design may lead to unstable operation of the power system, such as an unreasonable line layout, increasing the risk of faults [3]. In terms of construction safety, violations by construction personnel and inadequate safety protection measures may lead to accidents. Equipment failure is also an important risk factor, as equipment aging, poor quality, and other factors can easily cause power outages. Environmental factors cannot be ignored. Adverse weather conditions may damage power facilities, and complex geographical conditions increase construction difficulty and risks. By constructing a risk identification framework from these four dimensions, it is possible to comprehensively and systematically identify the risk elements of power engineering, providing a foundation for subsequent risk assessment and management.

3.2. Defects in the current risk management system

Traditional risk assessment methods are plagued by significant lag and are ill-equipped to adapt swiftly to the ever-increasing complexity of the power engineering environment and the rapidly evolving technical demands. These methods often rely on outdated evaluation indicators and models that fail to accurately reflect the current risk landscape, leading to potential misjudgments and inadequate risk mitigation strategies. Furthermore, the integrity of existing emergency plans is compromised due to their lack of comprehensive consideration for a wide range of unexpected scenarios. This deficiency results in insufficient response measures when these plans are implemented in real-world situations. Additionally, the current risk management system is severely lacking in research on the risk transmission mechanism [4]. Without a clear understanding of how risks propagate through different links in power engineering and the specific impact modes they adopt, it becomes nearly impossible to effectively prevent and control risks at their source. Consequently, this not only increases the likelihood of risk occurrence but also exacerbates the potential harm that these risks can inflict on power construction projects.

4. The impact mechanism of technological innovation on risk management

4.1. The positive risk control effect of technological innovation

4.1.1. Risk warning capability of real-time monitoring technology

Intelligent sensor networks, as a real-time monitoring technology, have proven to be highly effective in providing robust risk warning capabilities in power construction projects [5]. By strategically placing numerous sensor nodes throughout the construction environment, these networks are able to collect real-time data from key construction areas and critical links. The data gathered encompasses a wide range of parameters essential to the construction process, including temperature, pressure, stress, and more. When abnormal fluctuations are detected in these data streams, they serve as an early indication of potential construction quality deviations or emerging risks. Leveraging advanced data analysis algorithms, these anomalies can be swiftly and accurately identified, triggering timely warning signals. Armed with this critical information, construction management personnel are empowered to take immediate action [5]. They can adjust construction processes, enhance quality control measures, or implement other corrective actions to effectively mitigate risks and prevent their escalation. This proactive approach not only enables real-time monitoring but also ensures active risk prevention and control, significantly enhancing the safety and reliability of power construction projects.

4.1.2. Controllability of risks brought by standardized processes

Prefabricated assembly technology, as a standardized process, significantly enhances risk controllability in power construction. By completing the production of most components in a stable factory environment, this technology effectively reduces the risk of quality instability that is often caused by complex and unpredictable construction site conditions [6]. The standardized nature of prefabricated assembly ensures that the specifications and quality of components are more uniform, which not only facilitates smoother installation and debugging during the construction process but also minimizes delays and safety risks associated with mismatched components. Moreover, the application of prefabricated assembly technology helps to reduce hazardous work processes such as wet and high-altitude operations on construction sites. This further mitigates the personal safety risks faced by construction personnel, thereby improving the overall safety and reliability of the entire power construction process. In summary, prefabricated assembly technology plays a crucial role in enhancing the controllability of risks and ensuring the smooth progress of power construction projects.

4.2. New risks arising from technological iteration

4.2.1. Operational risks arising from technological complexity

As technological innovation progresses, power construction projects are increasingly confronted with new risks stemming from the complexity of advanced technologies. One prominent example is the emergence of network security risks due to the development of smart grid technology [7]. The intricate nature of modern power systems introduces a range of operational challenges. Firstly, the construction process becomes more convoluted, with a significant increase in the number of operational steps required. This complexity heightens the likelihood of human errors. For instance, during the installation and debugging of sophisticated power equipment, even minor mistakes can have severe consequences, potentially leading to equipment damage, project delays, or even safety incidents. Secondly, the growing sophistication of smart devices raises the operational threshold, placing higher demands on the quality and capabilities of construction personnel. Workers are now required to possess a broader range of professional knowledge and skills to effectively manage and operate these advanced systems. If the personnel do not meet these new requirements, they may struggle to operate intelligent equipment correctly, thereby increasing the risk of errors and subsequent complications. Addressing these operational risks requires not only enhanced training for workers but also the development of more robust error-prevention mechanisms.

4.2.2. Compatibility risks of multi-system integration

With the innovation of power construction engineering technology, compatibility risks are faced in the process of multi-system integration. There are differences in interface standards between different technical modules, which may lead to system conflict issues [8]. For example, in the power system, when integrating new intelligent monitoring systems with traditional power transmission systems, data transmission may encounter errors or interruptions due to inconsistent interface standards, which can affect the normal operation of the system. This compatibility risk not only leads to unstable power supply, but may also cause a series of safety issues such as equipment damage, personnel injury, etc. Therefore, in the process of technological innovation, it is necessary to fully consider the compatibility issues of multi-system integration and develop corresponding risk management strategies to ensure the smooth and safe operation of power construction projects.

5. Risk management strategies in the context of smart grid

5.1. Construction of a dynamic risk management model

5.1.1. Risk assessment system based on digital twins

Building a dynamic risk management model is crucial in the context of smart grids. The risk assessment system based on digital twins can establish a three-dimensional dynamic risk assessment model by integrating BIM and GIS technology [9]. Through the 3D modeling capability of BIM and the spatial analysis capability of GIS, the physical structure and geographical location information of power grid facilities can be accurately presented. This model can obtain real-time power grid operation data and simulate risk conditions under different working conditions. By utilizing digital twin technology, a virtual model corresponding to the actual power grid system is created to achieve dynamic monitoring and evaluation of risks. By continuously updating the virtual model and synchronizing it with the actual system status, potential risks can be identified in a timely manner, providing a scientific basis for the formulation and adjustment of risk management strategies.

5.1.2. Dynamic adjustment mechanism for risk warning threshold

In the context of the smart grid, constructing a dynamic risk management model requires considering the dynamic adjustment mechanism of risk warning thresholds. An intelligent warning algorithm that considers real-time changes in environmental parameters can be designed. By monitoring environmental parameters in real-time, such as temperature, humidity, wind speed, etc., these data can be input into the algorithm. The algorithm evaluates risks in real-time based on preset rules and models. When the risk level exceeds the normal range due to changes in environmental parameters, the warning threshold can be automatically adjusted. This dynamic adjustment mechanism can more accurately reflect the actual risk situation and avoid misjudgments or omissions caused by fixed thresholds. At the same time, this mechanism should be continuously optimized by combining historical and real-time data of power grid operation to improve the effectiveness and accuracy of risk management [10].

5.2. Technical risk management safeguard measures

5.2.1. Comprehensive technical talent training system

In the context of the smart grid, a comprehensive technical talent training system is crucial. On the one hand, attention should be paid to the imparting of engineering technology knowledge, including power technology and information technology related to smart grids, so that talents have a solid technical foundation and can cope with complex technological environments. On the other hand, knowledge of risk management is indispensable, covering risk identification, assessment, and response strategies. During the training process, teaching should be combined with practical cases to enable talents to understand the relationship between technology and risk in practice. At the same time, encourage talents to participate in practical projects, exercise their technical and risk management abilities in the projects, so that they can quickly adapt to the work requirements in the smart grid environment, and provide strong talent support for technical risk management.

5.2.2. Construction of technical application standards and specifications

In the context of the smart grid, the construction of technical application standards and specifications is crucial. It is necessary to establish quality control standards for the application of new technologies in power construction, clarify the qualified range of various technical indicators, and ensure the stability and reliability of technical applications. At the same time, the operating standards should cover in detail the construction process, equipment operation methods, personnel safety protection, and other aspects to ensure the standardization and safety of the construction process. Standardized construction should be combined with the characteristics of smart grids, fully considering the compatibility between new technologies and existing systems, as well as the potential risk factors that may arise. By strictly adhering to standards and regulations, constraining the application process of technology, reducing the risks caused by non-standard use of technology, improving the quality and efficiency of power construction, and providing strong support for the stable operation of smart grids.

5.3. Typical engineering practice verification

5.3.1. Risk management case of ultra-high voltage transmission projects

The ultra-high voltage transmission project faces various risks, such as high technical difficulty and significant environmental impact. Taking a certain ultra-high voltage transmission project as an example, an intelligent monitoring system is used for risk management. The system monitors the status of the line in real time, including parameters such as temperature and tension. When an abnormality occurs, it can be quickly alerted to. For example, during an extreme weather event, the system promptly detects an abnormal tension in a certain section of the line, and the staff takes quick measures based on this to avoid damage to the line. Meanwhile, the intelligent monitoring system can also monitor the construction process to ensure compliance with regulations and reduce safety accidents. Through this project practice, the effectiveness of the intelligent monitoring system in risk management of ultra-high voltage transmission projects has been verified, providing references for other similar projects.

5.3.2. Risk management plan for distributed energy access

The integration of distributed energy into the smart grid faces many risks and requires the development of effective disposal plans. From a technical perspective, it is necessary to strengthen the research on monitoring and control technology for distributed energy access points to ensure their stable access. For example, advanced sensor technology is used to monitor parameters such as power quality and power fluctuations in real-time, in order to adjust them in a timely manner. In terms of management, establish a sound approval and supervision mechanism for distributed energy access, strictly review its capacity, location, etc., to avoid causing excessive impact on the power grid. At the same time, emergency plans should be developed to quickly isolate the source of faults or abnormal situations, ensuring the normal operation of other parts of the power grid. It is necessary to strengthen the training of relevant operators and improve their ability to cope with the risks of distributed energy access.

6. Conclusion

This study systematically explores engineering technology innovation and risk management strategies in power construction. By summarizing the interaction patterns between the two, a technology management dual wheel drive control mode is proposed. Research has found that there are currently shortcomings in intelligent algorithm optimization and cross-domain risk transmission mechanisms. In terms of technological innovation, although some progress has been made, there is still room for improvement. In terms of risk management strategy, it is necessary to further improve the response mechanism for complex risks. Looking ahead, artificial intelligence and blockchain technology have broad application prospects in the risk management of power engineering. Artificial intelligence can improve the accuracy of risk prediction, while blockchain technology can enhance information transparency and security. The combination of the two will bring new opportunities and development directions for engineering technology innovation and risk management in power construction.

Disclosure statement

The author declares no conflict of interest.

References

[1] Okolelova E, Shibaeva M, Shalnev O, et al., 2018, Risk Assessment Models of the Use of Innovative Technologies in Construction as a Factor in the Development of Energy Management, Energy Management of Municipal Transportation Facilities and Transport, Springer International Publishing, Cham, 22–35.

[2] Albert A, Hallowell MR, 2013, Safety Risk Management for Electrical Transmission and Distribution Line Construction. Safety Science, 51(1): 118–126.

[3] ur Rehman O, Ali Y, Sabir M, 2022, Risk Assessment and Mitigation for Electric Power Sectors: A Developing Country's Perspective. International Journal of Critical Infrastructure Protection, 36: 100507.

[4] Kyriakopoulos GL, Arabatzis G, 2016, Electrical Energy Storage Systems in Electricity Generation: Energy Policies, Innovative Technologies, and Regulatory Regimes. Renewable and Sustainable Energy Reviews, 56: 1044–1067.

[5] Agunov AV, Titova TS, Kruchek VA, 2016, On the Construction of Power Quality Control Systems. Russian

Electrical Engineering, 87(5): 251–255.

[6] Rao R, Zhang X, Shi Z, et al., 2014, A Systematical Framework of Schedule Risk Management for Power Grid Engineering Projects' Sustainable Development. Sustainability, 6(10): 6872–6901.

[7] Hrinchenko H, Koval V, Shmygol N, et al., 2023, Approaches to Sustainable Energy Management in Ensuring Safety of Power Equipment Operation. Energies, 16(18): 6488.

[8] Li S, 2024, Risk Assessment and Coping Strategies in Power Engineering Project Management. Financial Engineering and Risk Management, 7: 114–118.

[9] Zheng G, Tianqi L, Qiang M, 2016, Cost analysis and risk supervision of project in electric power construction, International Society for Informationization and Engineering, Proceedings of 2016 4th International Conference on Management Science, Education Technology, Arts, Social Science and Economics (MSETASSE 2016), State Grid Liaoning Electric Power Company Limited Economic Research Institute, 1574–1577.

[10] Engineering Power Systems, 2018, New Findings from Isfahan University of Technology in the Area of Power Systems Reported (Evaluating the Operational Flexibility of Generation Mixture With an Innovative Techno-Economic Measure). Journal of Engineering, 2018: 578.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Data Elements and Trustworthy Circulation: A Clearing and Settlement Architecture for Element Market Transactions Integrating Privacy Computing and Smart Contracts

Huanjing Huang

440306198511260035, Shenzhen 518000, Guangdong, China

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: This article explores the characteristics of data resources from the perspective of production factors, analyzes the demand for trustworthy circulation technology, designs a fusion architecture and related solutions, including multi-party data intersection calculation, distributed machine learning, etc. It also compares performance differences, conducts formal verification, points out the value and limitations of architecture innovation, and looks forward to future opportunities.

Keywords: Data elements; Privacy computing; Smart contracts

Online publication: October 15, 2025

1. Introduction

With the rapid development of the digital economy, the importance of data as a key factor of production is becoming increasingly prominent. The Opinion on Building a More Comprehensive System and Mechanism for Market-Oriented Allocation of Factors, released in 2020, emphasizes the need to accelerate the cultivation of the data factor market. In this context, the marketization of data elements faces many challenges and opportunities. From the perspective of production factors, there are problems in defining property rights, forming value, and market-oriented transactions; There are dual technical requirements for analyzing trustworthy circulation from the perspectives of information security and process trust. This article focuses on these issues and proposes an architecture that integrates privacy computing and smart contracts. It explores its application and related technical solutions in the clearing and settlement of data element market transactions, which have important academic and practical significance.

2. Theoretical basis of the data element market and trustworthy circulation

2.1. The core characteristics and operational mechanism of the data element market

From the perspective of production factors, data resources have unique attributes. The definition of its property rights is complex, and due to the replicability and multi-source nature of the data, it is difficult to clarify ownership. The formation of value is influenced by various factors, including data quality, application scenarios, etc. In terms of market-oriented transactions, data ownership is a key prerequisite, but current technology is difficult to accurately determine the rights holders of data. Pricing also faces challenges, as the value of data varies depending on application scenarios and users, and there is a lack of unified pricing standards. There are also problems with supply-demand matching, as the diversity and dispersion of data make it difficult to efficiently achieve supply-demand integration. At the same time, technological constraints and institutional bottlenecks coexist, and there is a lack of effective data protection and transaction security measures in technology. Relevant laws and regulations in the system are not yet sound, which affects the healthy operation of the data factor market [1].

2.2. Dual technical requirements for trustworthy circulation

Analyze the dual technical requirements of trustworthy circulation from the perspectives of information security and process trust. In terms of information security, data privacy protection is crucial. With the development of the data element market, it is necessary to ensure that data is available and invisible in circulation to protect the privacy and rights of data owners. Privacy computing technology plays a crucial role in this regard, as it enables data analysis and utilization without compromising data privacy. In the dimension of process trust, the reliability and automatic execution of the transaction process are key [2]. Smart contracts, with their advantage of automated execution of transaction rules, can ensure the fairness and transparency of transactions, reduce human intervention and errors, enhance the trust of the transaction process, and promote the trustworthy circulation of data elements.

3. The overall design framework of integrated architecture

3.1. Hierarchical system architecture design

The fusion architecture design includes a hierarchical system architecture with four layers. The data resource layer is responsible for storing and managing data, providing a data foundation for the upper layer. This layer ensures that data is properly organized, indexed, and accessible, which is crucial for the smooth operation of the entire system. The privacy computing layer utilizes technological means to calculate and analyze data while protecting data privacy, ensuring the security and compliance of data during circulation and use. This layer incorporates advanced encryption techniques and secure multi-party computation methods to safeguard sensitive information without compromising functionality. The smart contract layer standardizes the transaction rules and processes of data elements through the form of smart contracts, achieving automated transaction execution and supervision [3]. This layer ensures that all transactions are conducted in a transparent, fair, and tamper-proof manner, reducing the need for intermediaries and minimizing the risk of fraud. The clearing and settlement layer is mainly responsible for clearing and settling data element transactions, ensuring the accuracy and fairness of transactions. This layer provides robust mechanisms for financial transactions, ensuring that all parties receive accurate and timely settlements. Each layer has clear functional positioning and interface standards, and efficient collaboration and data transfer between layers are achieved through reasonable interface design. This architecture not only ensures data security and privacy but also enhances the efficiency and transparency of data transactions, making it a robust framework for promoting the trustworthy circulation of data elements in a complex digital environment.

Additionally, the architecture is designed to be scalable and adaptable, allowing for future enhancements and integration with emerging technologies such as blockchain and artificial intelligence, which further strengthen its ability to support the evolving needs of the data-driven economy.

3.2. On-chain and off-chain collaboration mechanism

In the element market transaction clearing and settlement architecture that integrates privacy computing and smart contracts, the collaborative operation of the blockchain-based contract execution layer and privacy computing nodes is crucial. This collaboration ensures that the strengths of both blockchain and privacy computing are fully leveraged to enhance the security, transparency, and efficiency of the entire transaction process. By designing a reasonable collaborative operation model, the responsibilities and interaction methods of all parties involved can be clearly defined [4]. For data flow and control flow, specialized interaction protocols need to be established to ensure accurate transmission and interaction of data on and off the chain. Meanwhile, the verification mechanism is indispensable as it can rigorously verify the authenticity, integrity, and compliance of the interaction process of data, prevent data leakage and erroneous operations, and ensure the security and efficiency of clearing and settlement of factor market transactions. This architecture not only addresses the challenges of data privacy and security but also provides a robust framework for trustworthy and efficient transactions in the element market.

4. Key technology implementation path

4.1. The integrated application of privacy computing technology

4.1.1. Application of secure multi-party computation in data comparison

Design a multi-party data intersection calculation scheme based on secret sharing, utilizing secure multi-party computing technology to protect sensitive fields of transaction subjects during data comparison. By using a secret sharing mechanism, data is segmented and distributed to multiple participants for computation without leaking the original data. Each party performs calculation operations based on its own data shares, and finally combines the results through specific algorithms to obtain the intersection information of multi-party data [5]. This approach not only ensures the privacy and security of data but also provides a highly efficient and scalable solution for data comparison. It enables effective verification and processing of relevant data during the clearing and settlement process of data element market transactions while protecting sensitive information. Additionally, the scheme is designed to be resilient against potential collusion attacks among participants, ensuring robustness and reliability in a variety of complex scenarios. This makes it an ideal choice for enhancing the trustworthiness and security of data transactions in the digital economy.

4.1.2. Federated learning modeling framework

Build a distributed machine learning model training mechanism and use homomorphic encryption technology to ensure the privacy protection of model parameters during the aggregation process [6]. Homomorphic encryption allows specific computational operations to be performed directly on ciphertext, enabling parameter aggregation in model training without leaking the original data. In the federated learning modeling framework, the data of each participant is kept locally, and collaborative training of the model is achieved through encrypted transmission and aggregation of model parameters. This approach not only avoids direct data sharing but also enhances the security of the entire training process by ensuring that sensitive information remains encrypted throughout. It protects the privacy and security of data from all parties while fully utilizing the value of data to improve the performance and accuracy of the model. Additionally, this mechanism is designed to be highly scalable and adaptable to various data environments, making it suitable for large-scale distributed machine learning applications where data privacy is paramount.

4.2. Enhanced design of smart contracts

4.2.1. Verifiable calculation contract template

Developing a universal contract framework that supports zero-knowledge proofs requires consideration of multiple factors. Firstly, it is necessary to clarify the basic structure and functional modules of the framework, including the design of input and output interfaces, which should be able to adapt to different privacy computing scenarios and data formats. In the selection of zero-knowledge proof algorithms, it is necessary to comprehensively consider computational efficiency, security, and scalability. For the collaborative verification mechanism between on-chain verification and off-chain computing, it is necessary to establish an efficient communication protocol to ensure that the off-chain computing results can be accurately verified on-chain [7]. At the same time, it is necessary to design reasonable verification rules and processes, conduct comprehensive inspections of off-chain computing processes and results, and prevent malicious tampering and uploading of erroneous results. Through these technological means, the effective enhancement of smart contracts in a private computing environment can be achieved, ensuring the credibility and efficiency of data element market transaction clearing and settlement.

4.2.2. Automated clearing and settlement logic

In the enhanced design of smart contracts, automated clearing and settlement logic is crucial. One of the key paths is to design a multi-stage transaction processing model that includes fund freezing, account splitting settlement, and dispute arbitration. The fund freezing stage ensures the security of transaction funds and prevents issues such as fund misappropriation during the settlement process. During the sub-account settlement stage, funds are accurately allocated to various stakeholders based on preset rules. The dispute arbitration stage provides a mechanism for resolving potential transaction disputes. At the same time, a specification for off-chain data access based on oracle machines will be developed to enable smart contracts to obtain reliable external data, better support the execution of clearing and settlement logic, and ensure the accuracy and fairness of transactions [8].

5. System implementation and verification analysis

5.1. Prototype system construction

5.1.1. Implementation of distributed storage architecture

In the implementation of a distributed storage architecture, a hybrid storage solution combining IPFS and blockchain is adopted. IPFS has advantages such as content addressing and distributed storage, and can efficiently store data shards. By designing a data shard encryption storage mechanism, the security and privacy of data are ensured, and data leakage and tampering are prevented [9]. Meanwhile, establish an integrity verification mechanism to ensure the integrity of data during storage and transmission. By utilizing the tamper-proof and traceable properties of blockchain, the storage and verification of data information is recorded, providing a guarantee for the trustworthy circulation of data. This hybrid storage solution combines the advantages of IPFS and blockchain, providing a reliable distributed storage architecture foundation for the trustworthy circulation of data elements.

5.1.2. High concurrency processing optimization

It is crucial to adopt sharded blockchain technology to achieve high concurrency processing optimization. This technology can effectively improve system throughput and meet the needs of a large number of data element transactions [10]. Sharding divides the blockchain network into smaller, more manageable pieces called shards, each capable of processing transactions independently. This not only enhances the overall efficiency of the system but also ensures that the network can handle a significant increase in transaction volume without compromising on speed or security. At the same time, building a horizontally scalable computing node cluster architecture that can flexibly expand according to business load is essential. When transaction traffic increases, it is convenient to add computing nodes to ensure the stability and efficiency of the system in high-concurrency situations. This architecture allows for the dynamic allocation of computing resources, preventing any single node from becoming overloaded and thus improving the reliability and availability of the entire system. By doing so, it provides strong support for efficient transaction clearing and settlement of data elements in a trustworthy circulation environment, ensuring that the system can adapt to varying levels of demand while maintaining high performance and security standards.

5.2. Typical scenario testing verification

5.2.1. Joint risk control case of financial credit

In the case of joint risk control in financial credit, it is crucial to simulate the joint modeling process of multiple data sources. By constructing practical scenarios and integrating credit data from different financial institutions, including customer basic information, credit records, loan history, etc. Utilizing privacy computing technology for data fusion and analysis without disclosing sensitive data. At the same time, establish smart contracts to standardize the use of data and transaction settlement rules. During this process, the focus is on verifying the balance between privacy protection and model accuracy. By continuously adjusting privacy protection parameters and algorithms, observe the accuracy of the model in risk assessment. The results indicate that reasonable privacy protection settings can maintain high model accuracy while ensuring data security, effectively improve the effectiveness of financial credit joint risk control, and provide a more reliable risk decision-making basis for financial institutions.

5.2.2. Medical data trading scenarios

It is crucial to conduct testing on the encrypted transaction process for sensitive information, such as genetic data, in medical data trading scenarios. By constructing a simulated trading environment, verify the security and accuracy of the ciphertext in various stages of the transaction. Observe the processes of data encryption, transmission, decryption, and transaction confirmation to ensure that the privacy of genetic data is not compromised. At the same time, evaluate the effectiveness of implementing compliance audit functions. Check the completeness and accuracy of audit records, whether they can effectively trace transaction behavior, and ensure that every transaction involving sensitive information complies with relevant regulations and ethical standards. This not only ensures the legality of medical data transactions but also provides a reliable trust foundation for market participants, promoting the healthy development of the medical data element market.

5.3. Performance comparison analysis

5.3.1. Comparison of calculation efficiency indicators

Compare the performance differences between traditional centralized solutions and the architecture proposed in this paper in terms of response latency and throughput dimensions. In terms of response latency, traditional centralized solutions often suffer from congestion due to centralized processing of large amounts of data and transaction requests, resulting in high response latency. The architecture presented in this article utilizes privacy computing and the distributed processing capabilities of smart contracts to enable parallel processing of multiple tasks, effectively reducing the waiting time for individual tasks and significantly reducing response latency. In terms of throughput, traditional centralized solutions are limited by server performance and network bandwidth, resulting in limited throughput growth in high concurrency situations. The architecture of this article relies on the collaborative work of distributed nodes, which can flexibly expand processing capabilities and maintain high throughput when facing large amounts of data and transactions, demonstrating better computational efficiency and performance advantages.

5.3.2. Safety strength assessment

Based on the Dolev Yao threat model, formal verification is conducted to analyze the security of the architecture that integrates privacy computing and smart contracts through rigorous mathematical logic and reasoning. This model assumes that the attacker has strong computing power and complete control over network communication, and evaluates the security of the system in this extreme case. By analyzing the performance of the system in the face of various potential attacks, the effectiveness of enhancing its resistance to attacks can be quantified. From the perspective of privacy computing, verify its level of protection for data privacy to ensure that data is not leaked during circulation and transactions. From the perspective of smart contracts, verifying the accuracy and immutability of contract execution ensures the fairness and reliability of transaction clearing and settlement, comprehensively evaluates the security strength of the architecture, and provides strong guarantees for the trustworthy circulation of data element markets.

6. Conclusion

This architecture integrates privacy computing and smart contracts, bringing innovative value to the clearing and settlement of data element market transactions. By using relevant technological means to break down data silos and achieve trustworthy circulation of data, the trust cost of all parties involved in transactions has been reduced, and market efficiency has been improved. However, the current architecture has research limitations in cross-chain interoperability and dynamic access control. The imperfect cross-chain interoperability limits the data exchange and collaboration between different blockchain networks, affecting the widespread application of the architecture. Insufficient dynamic access control may lead to increased data security risks. Looking ahead, the integration of 5G edge computing and hardware envelope technology will bring new opportunities for architecture development. This integration is expected to further enhance data processing capabilities and security, provide new ideas and methods for solving existing problems, and promote the continuous improvement of the data element market transaction clearing and settlement architecture.

Disclosure statement

The author declares no conflict of interest.

References

[1] Wang S, Ouyang L, Yuan Y, et al., 2019, Blockchain-Enabled Smart Contracts: Architecture, Applications, and Future Trends. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 49(11): 2266–2277.

[2] Rouhani S, 2021, Data Trust Framework Using Blockchain and Smart Contracts, dissertation, University of Saskatchewan.

[3] Serrano W, 2022, Verification and Validation for Data Marketplaces Via a Blockchain and Smart Contracts. Blockchain: Research and Applications, 3(4): 100100.

[4] Perera S, Hijazi AA, Weerasuriya GT, et al., 2021, Blockchain-Based Trusted Property Transactions in the Built Environment: Development of an Incubation-Ready Prototype. Buildings, 11(11): 560.

[5] Afraz N, Wilhelmi F, Ahmadi H, et al., 2023, Blockchain and Smart Contracts for Telecommunications: Requirements vs. Cost Analysis. IEEE Access, 11: 95653–95666.

[6] Di Nenno A, 2019, Incentive-Compatible and Privacy-Preserving Data Analytics System enabled by Blockchain and Multiparty Computation, dissertation, Politecnico di Torino.

[7] Mvula F. A Conceptual Secure Blockchain-Based Settlement and Clearing House for Mobile Financial Services in Zambia, dissertation, The University of Zambia, 2020.

[8] Honari K, Rouhani S, Falak NE, et al., 2023, Smart Contract Design in Distributed Energy Systems: A Systematic Review. Energies, 16(12): 4797.

[9] Peters GW, Panayi E, 2016, Understanding Modern Banking Ledgers Through Blockchain Technologies: Future of Transaction Processing and Smart Contracts on the Internet of Money, Springer International Publishing.

[10] Asante M, Epiphaniou G, Maple C, et al., 2021, Distributed Ledger Technologies in Supply Chain Security Management: A Comprehensive Survey. IEEE Transactions on Engineering Management, 70(2): 713–739.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Structural Optimization and Innovative Practice in the Mechanical Design of Amusement Equipment

Bin Liu

Shenzhen Yingfeile Cultural Tourism Technology Co., Ltd., Shenzhen 518000, Guangdong, China

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: Materials mechanics and structural dynamics provide theoretical support for the structural optimization of amusement facilities. The design code system guides the design process, covering aspects such as strength and fatigue life. This paper introduces optimization methods like standardized module interfaces and variable density methods, as well as topics related to finite element simulation, reliability enhancement, innovative practices, and their significance.

Keywords: Amusement equipment; Structural optimization; Mechanical design

Online publication: October 15, 2025

1. Introduction

The safety and performance of amusement rides are of paramount importance. The newly promulgated "Safety Technical Code for Large-Scale Amusement Rides" (TSG 07-2023) in 2023 provides the latest regulatory basis for their design. The new code strengthens the application requirements of theories such as material mechanics and structural dynamics in the structural optimization of amusement rides, with particular emphasis on the validation standards for the rationality of material selection and the reliability of structural safety. The amusement ride design specification system established in accordance with the new code further refines the key technical requirements for structural strength, fatigue life, and failure mode analysis. The modular design standard, based on functional units, clarifies the requirements for interface standardization, effectively promoting the rapid reconstruction and maintenance of equipment. The new code encourages the use of advanced topological optimization techniques, such as the variable density method, and requires multi-condition verification through finite element simulation analysis. In addition, key design elements such as parameter sensitivity studies, redundant structural configuration, and fatigue damage assessment are listed as mandatory review items in the new code. These regulations significantly promote the development of mechanical design for amusement rides in terms of structural optimization and technological innovation.

2. The theoretical basis of amusement equipment structure optimization

2.1. Fundamentals of material mechanics and structural dynamics

Material mechanics provides an essential theoretical basis for the structural optimization of amusement rides. It investigates the deformation, stress, strength, stiffness, and stability of materials under various external forces. By gaining a deep understanding of the mechanical properties of materials, it is possible to select appropriate materials to ensure the safety and reliability of amusement rides when they are subjected to expected loads [1]. Meanwhile, structural dynamics focuses on the response of structures under dynamic loads. During operation, amusement rides are subjected to a variety of dynamic forces, such as vibrations and impacts. Structural dynamics can help establish accurate mathematical models for load distribution and dynamic response, analyze the vibration characteristics and dynamic response patterns of structures, and provide a basis for optimizing structural design to avoid adverse phenomena like resonance, thereby ensuring the normal operation of the equipment and providing a comfortable experience for passengers. In practical design, the application of material mechanics and structural dynamics requires not only theoretical support but also comprehensive analysis in combination with actual working conditions. For example, in the design of the drive unit for amusement rides, it is necessary to consider the fatigue life of materials and the complexity of dynamic loads. By integrating material mechanics and structural dynamics, the long-term performance of the equipment can be effectively predicted and optimized, thereby enhancing the overall safety and service life of the equipment. This multidisciplinary integrated design approach is particularly important in the development of modern amusement rides, as it can ensure that the equipment meets functional requirements while also providing a safe, reliable, and comfortable experience for passengers.

2.2. Play facility design code system

The design specification system for amusement rides is an important theoretical basis for the structural optimization of amusement equipment. Domestic and international safety technical specifications for special equipment provide clear guidance for the design of amusement rides. In terms of structural strength, the specifications stipulate the loads that amusement rides should be able to withstand under various operating conditions to ensure the safety and reliability of the equipment [2]. For example, based on the dynamic loads during operation, passenger weight distribution, and environmental factors (such as wind loads and seismic loads), strength calculations and verifications are carried out for key structural components. At the same time, fatigue life design requirements are also crucial. During long-term and frequent operation, structural components of amusement rides are subjected to alternating loads, which can easily lead to fatigue cracks. The specifications require the prediction and assessment of structural fatigue life through rational material selection, structural design, and stress analysis to ensure that the equipment can operate safely within the specified service life. These specifications cover not only the requirements for structural strength and fatigue life but also involve the overall design and optimization of the equipment, providing comprehensive theoretical support and practical guidance for the development of amusement rides.

3. Research on innovative design methods

3.1. Modular combination design principle

The standardization of modular interfaces based on functional units is key to the modular combination design of amusement rides. By conducting a detailed analysis of the functions of amusement rides and dividing them into several relatively independent functional units, each with specific functions and interfaces, a unified standard for these interfaces can be established. This ensures accurate and rapid connection and combination between different modules. This approach not only facilitates the rapid reconfiguration of the equipment to meet the changing demands of different sites and user needs but also greatly optimizes the maintenance process. When a module fails, it can be quickly replaced based on the standardized interface, reducing maintenance time and cost and enhancing the overall operational efficiency and reliability of the equipment. This provides strong support for the innovative design and sustainable development of amusement rides [3].

3.2. Application of topology optimization technology

The variable density method is an important technique in topology optimization. By applying the variable density method in three-dimensional space, the structure of amusement rides can be optimized. This method describes the topological form of the structure based on the distribution density of materials and uses optimization algorithms to achieve optimal performance targets under certain constraints. In the mechanical design of amusement rides, the variable density method is used to construct a new type of truss support structure system. Truss structures have good mechanical properties and stability. Through topology optimization, the distribution and connection of members can be determined reasonably, reducing the amount of material used while increasing the load-bearing capacity and reliability of the structure, thereby ensuring the safe operation of amusement rides. Moreover, the variable density method can be combined with finite element analysis to verify the optimized structure under multiple working conditions, ensuring its ability to withstand various complex loads in actual operation. This method not only improves design efficiency but also reduces design costs, providing strong support for the innovative design and sustainable development of amusement rides [4].

4. Key technology research

4.1. Finite element simulation analysis

4.1.1. Statics and dynamics simulation

Finite element simulation analysis is crucial in the mechanical design of amusement rides. For static and dynamic simulations, it is necessary to establish multi-condition finite element models that include wind loads and impact loads. Static simulation can analyze the stress and strain distribution of the structure under static loads to determine whether the structure meets the requirements of strength and stiffness. By conducting static analysis on key parts of amusement rides, potential structural weak points can be identified in advance [5]. Dynamic simulation, on the other hand, takes into account the dynamic characteristics of the equipment during operation, such as vibration and impact. Simulating the response of amusement rides under wind and impact loads can assess the structural stability and reliability. This helps to optimize the structural design and improve the safety and performance of amusement rides under complex working conditions. Moreover, through finite element simulation analysis, the fatigue life of the optimized structure can also be evaluated to further ensure the safety and reliability of the equipment during long-term operation. This comprehensive simulation analysis method not only improves design efficiency but also reduces testing costs, providing strong support for the innovative design and sustainable development of amusement rides.

4.1.2. Parametric sensitivity analysis

In the mechanical design of amusement rides, it is crucial to conduct parameter sensitivity studies based on finite element simulation analysis. The orthogonal array testing method can effectively identify key design variables and determine the direction for optimization [6]. By properly designing experimental schemes and conducting combined tests on multiple parameters, the impact of changes in each parameter on structural performance can be analyzed. This helps to identify the parameters that are most sensitive to structural response, such as the design variables corresponding to key indicators like stress and deformation. Subsequently, this provides a clear direction for subsequent structural optimization, focusing efforts on adjusting and improving key variables to enhance the mechanical performance and safety of amusement rides, ensuring that they can operate stably and reliably in actual service, meeting design requirements and usage standards.

4.2. Reliability design method

4.2.1. Redundant structure configuration strategy

The reliability of amusement rides is of utmost importance, and redundant structural configuration is one of the key strategies to enhance reliability. In the mechanical design of amusement rides, multiple protections for critical load-bearing components must be considered. For parts that are prone to failure, redundant structures should be provided, such as backup support members or reinforced structures. When the main structure is subjected to abnormal loads or suffers local damage, the redundant structures can promptly share the load to prevent catastrophic consequences. This requires adherence to strict design guidelines, taking into account the operating conditions of the equipment, types of loads, and potential failure modes. Through rational redundant structural configuration, the safety and reliability of amusement rides can be effectively improved, providing a safer and more enjoyable experience for visitors [7].

4.2.2. Fatigue life prediction model

During the long-term operation of amusement rides, fatigue damage is a key factor affecting their reliability. Establishing a cumulative damage assessment system based on the rainflow counting method is crucial for predicting fatigue life. The rainflow counting method can effectively simulate the stress cycle conditions of the equipment under complex working conditions and accurately identify the changes in stress amplitude and mean value [8]. By collecting and analyzing actual operating data, the rainflow counting method can be used to obtain the stress-time history curve. Based on this, combined with the fatigue characteristic curve of the material, the cumulative damage value of the equipment at different operating stages can be calculated using damage accumulation theories, such as Miner's rule. Subsequently, according to the relationship between cumulative damage and fatigue life, the fatigue life of key components of amusement rides can be predicted, providing a strong basis for reliability design.

5. Engineering practice and verification

5.1. Optimizing roller coaster tracks

5.1.1. Track support structure renovation

In the renovation of the roller coaster track support structure, an innovative approach was adopted by using honeycomb composite materials to replace the traditional steel structure. Honeycomb composite materials possess unique structural characteristics. Their internal honeycomb structure effectively reduces weight while maintaining strength. Through precise mechanical calculations and simulation analyses, the feasibility of applying this material in the track support structure was determined. The actual application results showed that this substitution achieved a weight reduction of 23%, significantly decreasing the self-weight of the track support structure. This not only reduced the pressure on the foundation structure but also enhanced the safety and stability of the roller coaster operation to a certain extent. Meanwhile, the application of this material has provided new ideas and directions for structural optimization in the mechanical design of amusement rides [9].

5.1.2. Dynamic characteristics test

A six-degree-of-freedom test platform was used to conduct dynamic characteristic tests on the optimized roller coaster track. Sensors were reasonably placed at key parts of the track to collect vibration data under different working conditions. The collected data were processed and analyzed using data analysis software, and the results were compared before and after optimization. The results showed that the track vibration was reduced by 40% after optimization. This indicates that the optimization measures have played a significant role in improving the dynamic characteristics of the track and effectively enhanced the safety and comfort of the roller coaster operation [10].

5.2. Sky wheel innovative design practice

5.2.1. Vane-type tension structure design

In the innovative design practice of the Ferris wheel, specifically in the design of the wheel spoke tension structure, the application of prestressed cable net structures is crucial. By reasonably applying prestress, the overall stiffness of the structure is effectively enhanced. The principle lies in the fact that the introduction of prestress alters the internal stress distribution within the structure, enabling it to better resist deformation when subjected to loads. Compared with traditional structures, this innovative design significantly reduces the amount of steel used by 18%. This not only lowers costs but also reduces the structure's self-weight to a certain extent, which is of great significance for improving the overall performance of the Ferris wheel. Meanwhile, the application of prestressed cable net structures also offers more possibilities for the aesthetic design of the Ferris wheel. On the basis of meeting mechanical performance requirements, it can better integrate aesthetics and functionality, providing visitors with a unique visual experience and riding sensation.

5.2.2. Intelligent monitoring system integration

In the integration of intelligent monitoring systems in the innovative design practice of Ferris wheels, the deployment of fiber-optic sensor networks is crucial. By reasonably arranging fiber-optic sensors at key structural locations of the Ferris wheel, it is possible to perceive real-time state information of the structure, such as stress and strain. These sensors transmit the collected data to the core processing unit of the monitoring system. The core processing unit employs advanced algorithms to analyze and process the data, determining whether the structure is in a healthy condition. Meanwhile, the system integration also needs to consider the stability and real-time nature of data transmission, using reliable communication protocols and transmission lines to ensure that the monitoring data is accurately and error-free delivered to the analysis end. In addition, to achieve efficient operation of the entire intelligent monitoring system, it is necessary to calibrate and optimize the system, improving the accuracy and reliability of monitoring, thereby providing strong support for the safe operation of the Ferris wheel.

5.3. Optimization of rotating equipment

5.3.1. Lightweight transmission system

Lightweighting the transmission system of amusement rides is crucial for enhancing equipment performance. Take the development of a magnesium alloy gearbox, for instance. Magnesium alloy boasts low density and high strength, which can significantly reduce the rotational inertia of the equipment when applied to the manufacture of the gearbox. Through precise design and process optimization, the rotational inertia was successfully reduced by 15%. During the design process, the material properties of magnesium alloy were fully considered, and the structure of the gearbox was reasonably arranged and reinforced to ensure sufficient strength and reliability when subjected to various loads during operation. Meanwhile, strict control was exercised over the manufacturing process to ensure that the quality and performance of the magnesium alloy gearbox meet the high standards required by amusement rides.

Moreover, this lightweight design also brings additional benefits, such as reducing energy consumption during equipment start-up and stop, improving transmission efficiency, and lowering operating costs. More importantly, the dynamic response performance of the equipment has been significantly enhanced through lightweighting, enabling it to better adapt to various complex working conditions that may arise during the operation of amusement rides, thereby further enhancing the stability and safety of the equipment. This design approach also offers a new direction for the sustainable development of future amusement rides and promotes technological progress in the industry. This innovative practice not only provides an effective solution for the lightweighting of amusement ride transmission systems but also offers valuable references for structural optimization and innovation in the mechanical design of amusement rides.

5.3.2. Innovation of safety protection devices

A series of engineering practices and validations have been carried out in the innovation of safety protection devices for rotating amusement rides. By conducting a detailed analysis of potential hazards that may arise during operation, a multi-stage hydraulic buffering system has been designed. This system plays a crucial role during emergency braking, effectively absorbing the kinetic energy of the equipment and preventing the significant impact forces caused by sudden stops from harming the device and passengers. In practice, after numerous simulation experiments and real-life operation tests, the results show that the multi-stage hydraulic buffering system significantly enhances the reliability of emergency braking. Whether it is sudden braking during normal operation or emergency shutdown under abnormal conditions, it can ensure that the equipment stops smoothly, providing strong support for the safe operation of amusement rides.

6. Conclusion

The structural optimization and innovative practices in the mechanical design of amusement rides hold significant importance in multiple aspects. The application of structural optimization methods in the design of large-scale amusement rides has effectively enhanced the performance and safety of the equipment. By making rational adjustments and improvements to the structure, the probability of failures has been reduced, and the service life of the equipment has been extended. The proposal of the intelligent design development direction based on digital twins has brought new ideas and technical means to the design of amusement rides, enabling more precise simulation and optimization. Innovative design not only raises the safety standards of the industry and ensures the safety of visitors' lives but also significantly improves operational efficiency, reduces maintenance costs and downtime, enhances the competitiveness of amusement rides in the market, and propels the amusement ride industry towards a higher level of development.

Disclosure statement

The author declares no conflict of interest.

References

[1] Molnar D, Blatnicky M, Ishchuk V, et al., 2024, Design of a Drive Unit for Technical Equipment Working in the Amusement Industry, International Conference on Machine and Industrial Design in Mechanical Engineering, Springer Nature Switzerland, Cham, 17–28.

[2] Ishchuk V, Blatnicky M, Molnar D, et al., 2025, Dimensional Calculation of the Main Beam of a Technical Equipment Intended for the Amusement Industry. Machine and Industrial Design in Mechanical Engineering: Proceedings of KOD 2024, 174: 68.

[3] Zhao Z, Song W, Wang H, et al., 2024, Development and Application of IoT Monitoring Systems for Typical Large Amusement Facilities. Sensors, 24(14): 4433.

[4] Wang L, Sun M, Fu Y, et al., 2022, Acceleration Test of Amusement Facilities Based on Sensing Equipment, 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT). IEEE, 2022: 702–705.

[5] Mojtabavi SM, Hashemi Evari MS, Hassankhouei E, 2023, Identifying the Components of Satisfaction in Tourists and Visitors of Amusement Parks with the Approach of Increasing Health. Geography and Human Relationships, 6(1): 160–176.

[6] Ye C, Yu Y, Zhang JJ, et al., 2021, Research on the Detecting Method of Motor Running State of Amusement Device Based on Infrared Thermography Technology, Advances in Condition Monitoring and Structural Health Monitoring: WCCM 2019, Springer Singapore, Singapore, 769–780.

[7] Woodcock K, 2022, Safety Evaluation of Amusement Rides Using Accumulated Accident Data: Accident Data Framework. Journal of Themed Experience and Attractions Studies, 2(1): 40–53.

[8] James AT, 2022, Service Quality Benchmarking of Amusement Parks Using a Hybrid Approach. International Journal of Quality & Reliability Management, 39(4): 1000–1019.

[9] Song W, Cheng P, Liu R, 2021, Research on Evaluation Method of Safety Integrity for Safety-Related Electric Control System of Amusement Ride, Advances in Condition Monitoring and Structural Health Monitoring: WCCM 2019, Springer Singapore, Singapore, 375–385.

[10] Hartono M, Ronyastra I, Fajrin TA, 2022, Human-Side Emotional Service Design for Experience-Centric Amusement Park, Proceedings of the Second Asia Pacific International Conference on Industrial Engineering and Operations Management Surakarta, Indonesia, September 14–16, 2898–2903.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Remote Diagnosis and Analysis of Rail Vehicle Status Based on Train Control Management System Data

Qiang Zhang, Feng Jiao, Fan Liu, Mengqi Yan, Xiaoyu Bai

The First Operation Branch Company Affiliated With Beijing Mass Transit Railway Operation Corp., Ltd., Beijing 102200, China

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: This article focuses on the remote diagnosis and analysis of rail vehicle status based on the data of the Train Control Management System (TCMS). It first expounds on the importance of train diagnostic analysis and designs a unified TCMS data frame transmission format. Subsequently, a remote data transmission link using 4G signals and data processing methods is introduced. The advantages of remote diagnosis are analyzed, and common methods such as correlation analysis, fault diagnosis, and fault prediction are explained in detail. Then, challenges such as data security and the balance between diagnostic accuracy and real-time performance are discussed, along with development prospects in technological innovation, algorithm optimization, and application promotion. This research provides ideas for remote analysis and diagnosis based on TCMS data, contributing to the safe and efficient operation of rail vehicles.

Keywords: Rail vehicle; TCMS data; Remote diagnosis; Data processing; Fault prediction

Online publication: October 17, 2025

1. Train diagnostic analysis overview

Train diagnostic analysis is a crucial step in ensuring the safe and efficient operation of trains, covering multiple aspects and methods. The main content of train fault diagnosis includes monitoring, diagnosis, and judgment of faults. When there are signs of a train malfunction, the diagnostic system needs to quickly detect and make judgments. This usually involves multiple steps, including collecting fault information, analyzing possible causes of faults, determining the most likely fault point, and so on. The modern train fault diagnosis system utilizes TCMS data and analyzes it to achieve rapid and accurate diagnosis and early warning of vehicle faults. These systems can monitor the operating parameters of trains in real time, such as speed, current, temperature, pressure, etc., and based on the real-time changes of these parameters, diagnose the cause of faults through preset fault models and issue warning information [1]. The computing performance of equipment in each train will have bottlenecks, and remote train diagnostic analysis can fully leverage the advantages of ground server clusters, providing a large amount of resources and computing power to meet the needs of multiple scenarios and big data [2,3].

2. TCMS data content

In order to unify the TCMS data transmission format, this article designs a standard data frame transmission format for reference in other projects in the future [3]. The data frame content is based on hexadecimal, and the overall structure is shown in **Table 1**.

Table 1. Data frame structure

MessageHeader	Type	MessageID	Content	ReservedFields	MessageTail
3 bytes	1 byte	4 bytes	N bytes	2 bytes	3 bytes

(1) Message header: Fixed padding bytes 0xAA, 0xAB, and 0xAC represent;

(2) Type: Fixed use 0x01;

(3) Message ID: Indicates the number of the message, starting from 0. For each new message sent, the message number is incremented by 1 and continues to accumulate until it reaches its maximum value before cycling;

(4) Data area: Loaded with status and fault data of vehicle TCMS;

(5) Reserved fields: fixed use of 0x00, 0x00;

(6) Message tail: Fixed padding bytes 0xBA, 0xBB, 0xBC. The specific message format is shown in **Table 2**.

Table 2. Wireless transmission message format

Byte No.	Content								Remarks
	0	1	2	3	4	5	6	7	
0	Message header 0xAA								
1	Message header 0xAB								
2	Message header 0xAC								
3	Message type 0x01								
4–7	Message ID								
8	0x02								
9	0x10								
10–13	Data Length								
14	0x01								
15	Line ID								
16–17	Train ID								
18	0x01								Equipment ID
19	Date (year)								
20	Date (month)								
21	Date (day)								
22	Date (hour)								
23	Date (minute)								
24	Date (second)								

Table 2 (Continued)

Byte No.	Content								Remarks
	0	1	2	3	4	5	6	7	
25–26	Time (ms)								
27–34	Reserved								
35–35+N-1	Vehicle operation data area, N bytes								Data area
35+N	Reserved								
35+N+1	Reserved								
35+N+2	Message tail 0xBA								
35+N+3	Message tail 0xBB								
35+N+1	Message tail 0xBC								

The data filling principle is based on variable filling (classified by navigation on all HMI status pages), and strives to continuously fill in data from unified subsystems, generally including traction system, auxiliary system, braking system, communication status, door system, air conditioning system, emergency communication unit, smoke and fire system, train operation parameters, axle temperature system, bypass status, and other important operational analysis variable data. The uploaded status data content should at least include the requirement to meet all status display variables of the display, for the purpose of reproducing the display interface on the ground. The uploaded fault data content should at least include all integrated subsystems and fault variables of each level in the onboard TCMS display unit [4].

The specific data filling situation of the data area needs to be compiled into a data list and submitted to the ground system, which will analyze the data based on the data list. The data list in the status area should include: variable name, variable code, byte offset, bit offset, data type, conversion expression, system, parsing rules, units, and other information. The fault area data list should include: fault name, fault code, fault code, byte offset, bit offset, data type, conversion expression, system, fault level, carriage number, fault cause, fault description, solution measures, etc.

3. Remote diagnosis and analysis with TCMS data

3.1. Data link

In order to use TCMS data as a data source to empower vehicle operation management, it is necessary to transmit the TCMS data of vehicles to the ground, and utilize the computing resources and capabilities of ground reconstruction to achieve data statistics, analysis, calculation, storage, mining, etc. A data transmission link based on 4G signal transmission was designed for this purpose [5,6]. The overall link is shown in **Figure 1**.

Figure 1. Vehicle analysis system data link

The data on the TCMS bus is transmitted to the server's public IP through the onboard wireless module and 4G channel. After passing through the server's firewall, the data is received, parsed, stored, counted, analyzed, and displayed by the server [7–9]. To prevent signal interruption during data transmission, a breakpoint continuation function is added on the onboard equipment side. When the train briefly reaches a section without 4G signals, the onboard equipment will briefly record the data that has not been sent out and wait for 4G to recover before retransmitting the data. The network layer uses TCP connections instead of unstable UDP, and the scheduled sending frequency of data packets is set to 1 second.

3.2. TCMS data processing

3.2.1. Missing value handling

(1) Delete: Delete samples (rows) or features (columns) with missing information attribute values to obtain a complete data table.

Advantages: Simple and easy to implement, very effective in situations where an object has multiple missing values for attributes, and deleted objects with missing values have a relatively small amount of data compared to the initial dataset.

Disadvantages: When the proportion of missing data is large, especially when the missing data is not randomly distributed, this method may lead to data deviation and lead to incorrect conclusions.

(2) Interpolation: Average value filling: If the null value is numerical, fill in the missing attribute value based on the average value of the attribute's values in all other objects; If the null value is non-numeric, use the mode of the attribute to fill in the missing attribute value.

3.2.2. Data normalization/standardization

Normalization and standardization of data are methods of feature scaling and key steps in data preprocessing. Different evaluation indicators often have different dimensions and units, which can affect the results of data analysis. In order to eliminate the dimensional influence between indicators, data normalization/standardization processing is needed to ensure comparability between data indicators.

(1) Normalization: Normalization is generally the process of mapping data to a specified range, used to remove the dimensions and units of data from different dimensions. The common mapping ranges are [0,1] and [-1,1], and the most common normalization method is Min Max normalization.

(2) Standardization: Standardization is processing data according to the columns of the feature matrix. There are various methods for data standardization, such as linear methods (such as extreme value method, standard deviation method), folded line methods (such as three line method), and curved methods (such as semi-normal distribution). Different standardization methods will have different impacts on the evaluation results of the system. Among them, the most commonly used is Z-score standardization. Z-score normalization provides the mean and standard deviation of the original data for data normalization. The processed data conforms to the standard normal distribution, with a mean of 0 and a standard deviation of 1.

3.2.3. Data denoising

(1) Wavelet transform: The widely used method for time series denoising is nonlinear threshold processing, which is based on the principle that the energy of useful signals is concentrated on a few wavelet coefficients after wavelet transform, while white noise is still dispersed on a large number of wavelet

coefficients in the wavelet transform domain. Therefore, relatively speaking, the wavelet coefficients of useful signals are inevitably larger than those of noise with dispersed energy and smaller amplitudes. Therefore, from the amplitude of the spectrum, useful signals and noise can be separated.

(2) Variational mode decomposition: Variational mode decomposition is an adaptive, completely non-recursive method for modal variation and signal processing. This technology has the advantage of determining the number of modal decompositions, and its adaptability is manifested in determining the number of modal decompositions for a given sequence based on actual conditions. During the subsequent search and solution process, it can adaptively match the optimal center frequency and finite bandwidth of each modality, and can effectively separate the intrinsic mode components (IMF), divide the signal into frequency domains, and obtain the effective decomposition components of the given signal, ultimately obtaining the optimal solution of the variational problem. It has a solid mathematical theoretical foundation, which can reduce the high complexity and strong nonlinearity of time series non-stationarity, decompose to obtain subsequences containing multiple different frequency scales, and is relatively stable, suitable for non-stationary sequences.

3.3. Advantages of remote diagnostic analysis

(1) Resource sharing and collaboration: The remote fault diagnosis system can achieve the sharing of fault diagnosis data for multiple subway lines, and multiple diagnostic systems can work together to improve the stability and accuracy of fault diagnosis. This means that a wider range of professional knowledge and resources can be fully utilized to jointly solve complex fault problems.

(2) Reduce labor and material costs: Remote fault diagnosis does not require technical personnel to personally go to the fault site, and only requires a remote connection to the equipment for troubleshooting. This greatly reduces labor costs and avoids additional expenses such as travel expenses. At the same time, it reduces the need for on-site operations and reduces the risk of equipment damage and maintenance.

3.4. Remote diagnostic analysis method

3.4.1. Correlation analysis

Correlation analysis can analyze which characteristic variables have a linear correlation with faults, including the fault itself and state data.

(1) Cross correlation: The correlation coefficient ranges from [-1,1], where -1 represents absolute negative correlation and 1 represents absolute positive correlation.

(2) Autocorrelation: To measure the correlation between current and historical failures, we need to use autocorrelation and partial autocorrelation functions. Autocorrelation is the cross-correlation of itself, representing the degree of correlation between different moments of the same sequence. At a delay of k, the autocorrelation function (ACF) is the correlation between sequence values separated by k time intervals, while the partial autocorrelation function (PACF) also considers values between intervals.

3.4.2. Fault diagnosis

(1) Fault decision tree: The decision tree method has a wide range of applications in fields such as classification, prediction, and rule extraction [10]. A decision tree is a tree-like structure where each leaf node corresponds to a classification, while non-leaf nodes correspond to partitions on a certain attribute.

Based on the different values of the sample on that attribute, it is divided into several subsets [11–13].

Fault decision trees require a lot of experience accumulation, and each type of fault needs to be sorted out as clearly as possible, and then diagnosed based on real-time data corresponding to the decision tree. As shown in **Figure 2**, determine the cause of the fault based on different judgment conditions.

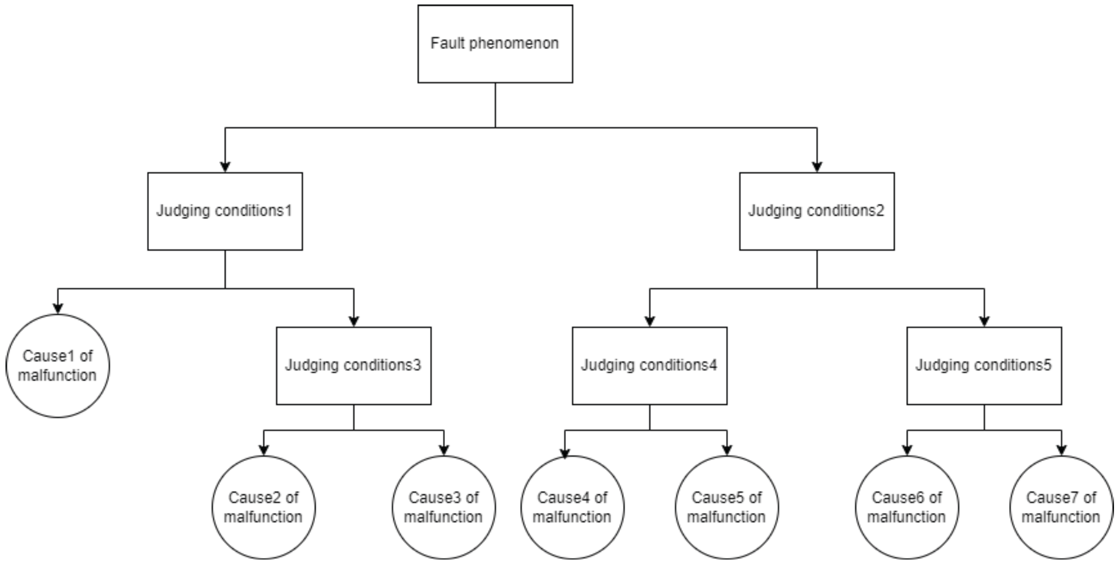


Figure 2. Fault decision tree

(2) Neural network classification: Use neural networks to classify each data. As shown in **Figure 3**, X represents each input data, which consists of fault codes and corresponding state feature data [14,15]. Z corresponds to each type of fault cause, and fault diagnosis is achieved by mapping the fault to the fault cause through a neural network.

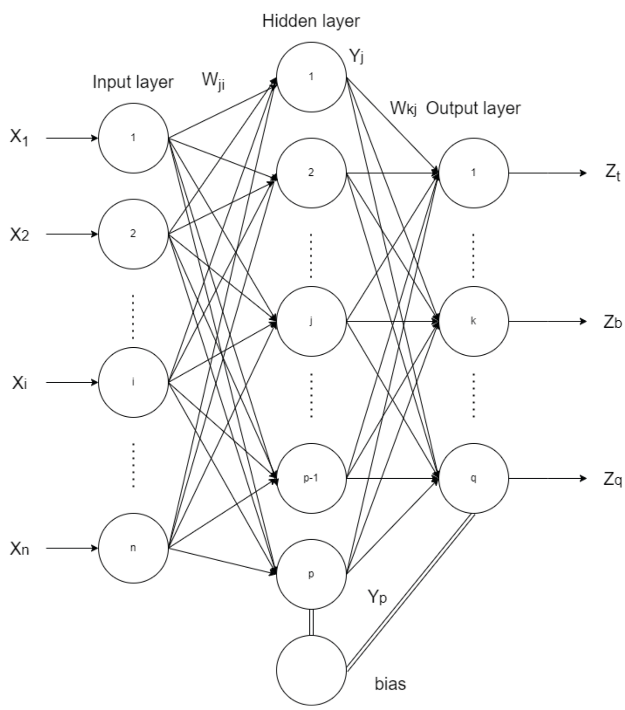


Figure 3. Principles of neural network classification

3.4.3. Fault prediction

Faults usually occur over time, and this type of data that describes one or more features that change over time is called time series data [16–18]. The use of historical data to predict the future is called a time series prediction algorithm. The general formula can be written as:

$$F\left(x_1, x_2, x_3, \ldots, x_n, Y'\right) = \hat{Y}$$

Among them, x_n represents state characteristic data such as vehicle speed and network voltage, Y' represents the historical fault dataset, and $\hat{Y}$ represents the predicted fault dataset. The fitting of prediction functions can be achieved through deep learning methods, thereby achieving fault prediction.

(1) Using LSTM prediction

Long short-term memory (LSTM) is a special type of recursive neural network. In order to solve the gradient problem during the training process of recurrent neural networks and the problem of information loss caused by excessive time, the LSTM network was developed [19,20]. The LSTM network is composed of multiple unit modules connected, each of which contains feedback-connected neurons and multiplication units. The unit modules include input gates, output gates, forget gates, and an internal memory unit. The structure of the LSTM network is shown in **Figure 4**.

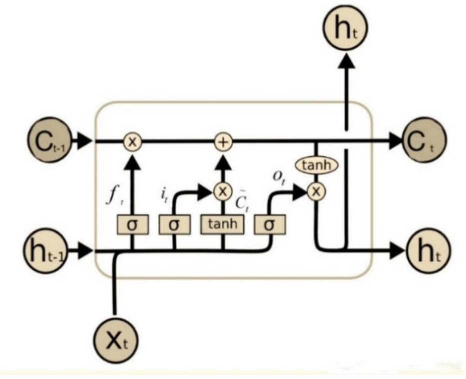


Figure 4. LSTM structure diagram

Use RNN models such as LSTM for autoregressive prediction, in order to predict the values of data in the future and achieve prediction of faults or other operational data. But single-point prediction can only represent the trend of data and cannot accurately predict the value of data, as shown in the following **Figure 5**.

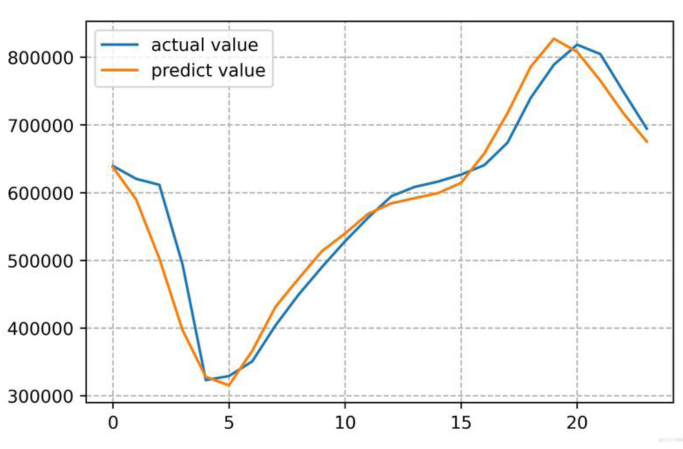


Figure 5. Schematic diagram of LSTM prediction

(2) Using DeepAR prediction

DeepAR is a prediction algorithm proposed by Amazon for unified modeling of a large number of related time series. DeepAR can generate probability predictions, with the goal of simulating conditional probability distributions $P(\mathbf{z}_{i,t_0:T} | \mathbf{z}_{i,1:t_0-1}, \mathbf{x}_{i,1:T})$, that is, modeling future sequences $\mathbf{z}_{i,t_0:T}$ using past time series $\mathbf{z}_{i,1:t_0-1}$ and covariates $\mathbf{x}_{i,1:T}$. Among them, t_0 is the time division point, $Z_{i,t}$ representing the value of the time series i in time t. Its overall structure is shown in **Figure 6**.

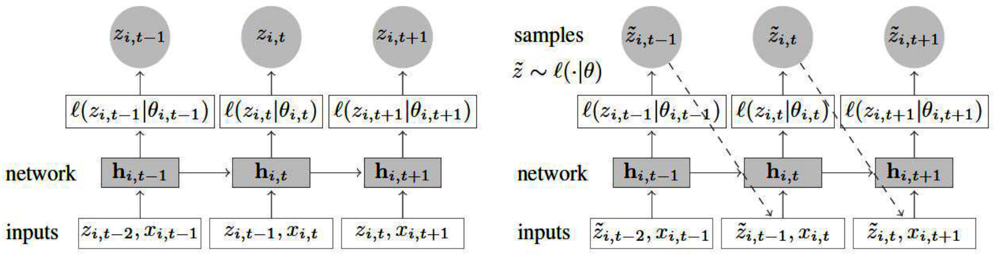


Figure 6. DeepAR structure diagram

If it is assumed that the data is approximately normally distributed, probability prediction can be used to obtain the probability distribution of the predicted results in different confidence intervals. From **Figure 7** below, it can be seen that predicting future data with different confidence intervals based on historical data is more effective in production environments than single-point prediction.

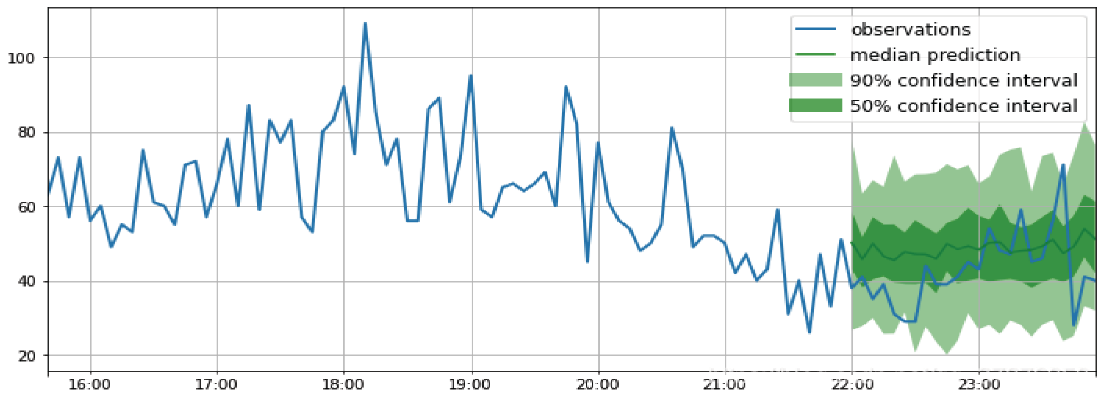


Figure 7. DeepAR prediction schematic diagram

4. Challenges and prospects of remote analysis and diagnosis of train status

4.1. Challenges faced

4.1.1. Data security assurance

(1) Data encryption: For data during transmission and storage, advanced encryption algorithms such as AES (Advanced Encryption Standard) should be used for encryption. This can effectively prevent hackers or other unauthorized third parties from intercepting or stealing data. Meanwhile, for the distribution and management of keys, public key encryption algorithms such as RSA can be used to ensure secure transmission of keys.

(2) Access control: Strict management and control of data access permissions, only authorized personnel

and institutions can access sensitive data. This can be achieved through the Role-Based Access Control (RBAC) model, which divides users into different roles and controls access based on their responsibilities and permissions. At the same time, two-factor authentication technology is adopted, such as the combination of a password and biometric information, to improve the security of the access system.

4.1.2. Balancing diagnostic accuracy and real-time performance

(1) High-speed network connection: Utilize high-speed and stable network connections, such as 5G or dedicated networks, to ensure fast and reliable data transmission between trains and remote diagnostic centers.

(2) Data compression and encoding: Adopting data compression technology to reduce data transmission volume, while adopting efficient data encoding methods to accelerate data processing speed.

(3) Algorithm optimization: Based on the characteristics of train fault diagnosis, optimize the algorithm model, reduce computational complexity, and improve real-time response speed.

4.2. Development prospects

4.2.1. Technological innovation

(1) Introducing new technologies: Introducing IoT technology to achieve a real-time connection between train equipment and remote diagnostic centers, improving data transmission efficiency and stability. Utilize cloud computing and big data technology to build a train fault data warehouse, providing powerful data support for fault analysis and prediction. Explore the use of artificial intelligence and machine learning technologies to achieve automatic fault identification and early warning, reducing errors and delays in human intervention.

(2) Device intelligence: Develop intelligent train equipment that can monitor and report its own status in real-time, reducing the frequency and intensity of manual inspections. Optimize communication protocols and interfaces between devices to achieve fast and accurate information transmission, and improve diagnostic efficiency.

4.2.2. Algorithm optimization

(1) Algorithm selection: Based on the characteristics and diagnostic needs of train faults, choose appropriate machine learning or deep learning algorithms, such as Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), etc. Develop customized algorithms to better meet the practical needs of train fault diagnosis.

(2) Model training and optimization: Utilize a large amount of train fault data for model training to improve diagnostic accuracy and generalization ability. Using regularization, dropout, and other techniques to prevent model overfitting, while utilizing optimization algorithms to improve training speed.

(3) Real-time diagnosis and prediction: Optimize the calculation speed and memory usage of algorithms to ensure fast response and accurate results during real-time diagnosis. Using methods such as time series analysis to predict train faults, develop maintenance plans in advance, and reduce the risk of faults occurring.

5. Conclusion

This article starts with an overview of remote diagnosis, followed by an introduction to the TCMS data structure and remote transmission link, explaining the advantages of remote diagnosis, and finally introducing the commonly used methods of correlation analysis, fault diagnosis, and fault prediction in remote diagnosis, providing a research approach for remote analysis and diagnosis based on TCMS data.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Zhu H, Yu H, 2024, Analysis and Troubleshooting of Urban Rail Vehicle Electrical Faults Based on TCMS Terminal Simulation Device. Internal Combustion Engine and Accessories, (4): 78–80.

[2] Lin B, Zhang J, Zheng D, et al., 2023, Fully Automatic Unmanned Train Air Conditioning System Control Scheme Based on TCMS. Electric Locomotive and the Urban Rail Vehicles, 46–48(6): 78–82.

[3] Zhang S, Zheng B, Pan Y, et al., 2019, Analysis on TCMS Data Organization and Transmission Mechanism of "Fuxing" EMU. Railway Rolling Stock, 39(06): 1–4.

[4] Sun J, Yang X, Liu B, et al., 2018, Remote Fault Diagnosis and Recovery of Fully Automatic Train Control System. Urban Rail Transit Research, 21(08): 149–151.

[5] Feng Q, 2017, Design and Implementation of Remote Monitoring and Diagnosis System for Urban Rail Vehicles, dissertation, Dalian University of Technology.

[6] Wang H, 2015, Research on Subway Vehicle Remote Diagnosis Ground System, dissertation, Beijing Jiaotong University.

[7] Song J, 2016, Metro Vehicle Monitoring and Control System (TCMS) research, dissertation, Nanjing University of Science and Technology.

[8] Xu J, Zhang T, 2015, Design of TCMS State Monitoring Simulation System for Tram. Railway Locomotive and EMU, (04): 5–7.

[9] Nie D, 2016, Application Analysis of WTB Process Data Messages Based on Time Division Multiplexing. Technology Outlook, 26(36): 120.

[10] Li R, 2021, Optimization Method for TCMS Signal Transmission Delay of Light Rail Trains Exported to Israel. Smart Rail Transit, 58(05): 19–24.

[11] Shen R, Gong L, 2017, Research on the Overall Plan of China Locomotive Remote Monitoring and Diagnostic System (CMD System). China Railway, (03): 9–15.

[12] Goyal D, Sharma A, Houran AM, 2024, Intelligent Machinery Fault Diagnostics and Prognostics: The Future of Smart Manufacturing, CRC Press.

[13] Rigatos G, Abbaszadeh M, Hamida M, et al., 2024, Fault Diagnosis for Electric Power Systems and Electric Vehicles, CRC Press.

[14] Yuan X, Liu H, Yang F, et al., 2024, Symmetric Circulant Matrix Decomposition-Based Multivariable Group Sparse Coding for Rolling Bearing Fault Diagnosis. Measurement Science and Technology, 35(6).

[15] Ding X, Deng A, Deng M, et al., 2024, A Dual-Weighted Adversarial Network for Partial Domain Fault Diagnosis of Machinery. Measurement Science and Technology, 35(6).

[16] Lin Y, Wang Y, Zhang M, et al., 2024, Information-Guided Signal Multi-Granularity Contrastive Feature Learning for Fault Diagnosis with Few Labeled Data. Advanced Engineering Informatics, 61102471.

[17] Wen T, Gao S, Wang J, et al., 2024, Signed-Data Reinforced Observer-Based Fault Diagnosis for Virtually-Coupled Electric Multiple Units Trains. Control Engineering Practice, 147105921.

[18] Zhang X, Wang Z, Du M, et al., 2024, Research on Data-Driven Model for Power Grid Fault Diagnosis Fusing Topological Quantification Information. Engineering Applications of Artificial Intelligence, 133(PC): 108286.

[19] Zhang Y, Cao Y, Zhao Y, 2024, Research on Fault Location Method of Train Operation Control System Based on Delta Debugging. Journal of Railway, 46 (11): 84–92.

[20] Zhang J, 2024, Simulation Model Analysis of Train Communication Network Fault Diagnosis. Heilongjiang Science, 15(18): 91–93 + 97.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Exploring the Development Model of UAVs Empowered by the Low-Altitude Economy

Fan Shi[1]*, Yuyao Zu[2]

[1]The High School Affiliated to Shaanxi Normal University, Xi'an 710062, Shaanxi, China
[2]Nankai University, Tianjin 300071, China

**Author to whom correspondence should be addressed.*

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: The "14th Five-Year Plan" and the Long-Range Objectives Through the Year 2035 propose to strengthen the construction of strategic emerging industrial clusters, promote the deep integration of the internet, big data, artificial intelligence, blockchain technology, etc. with the real economy, facilitate the development of advanced manufacturing, and consider unmanned aerial vehicles (UAVs) as an important breakthrough, providing significant opportunities for the development of the UAV industry. Therefore, this article takes the current status of the UAV industry development as a starting point, analyzes the exploration and practice of the UAV development model based on the low-altitude economy, and discusses strategic suggestions to promote the development of UAVs empowered by the low-altitude economy. Through analysis, this article aims to provide theoretical references and practical guidance for promoting the sustainable development of the UAV industry under the wave of the low-altitude economy.

Keywords: Low-altitude economy; UAV; Development model

Online publication: October 15, 2025

1. Introduction

An unmanned aerial vehicle (UAV) refers to an aircraft that can autonomously complete the entire flight mission, including takeoff, cruising, landing, and recovery, either by radio remote control equipment or devices with radio remote control capabilities, independent of ground personnel. UAV technology originated in the military field but has gradually been applied in the civilian sector due to its advantages, such as easy operation, low cost, high efficiency, and high safety factor. In recent years, with the implementation of low-altitude airspace management reform policies, China's UAV industry has entered a rapid development stage. However, the further development of the UAV industry still faces many bottlenecks. In this context, deeply exploring the development model of UAVs empowered by the low-altitude economy not only helps to break through the development dilemma of the UAV industry but also has important theoretical and practical significance for promoting the high-quality development of the low-altitude economy.

2. Current status of the development of the drone industry

Since 2016, the entire drone industry has ushered in a period of rapid development. By 2024, the total industrial scale had exceeded 200 billion yuan, representing an increase of over 50% compared to the 130 billion yuan in 2023. Civil drones accounted for 76% of the market share. As of August 2024, the number of real-name registered drones in China reached 1.987 million, and the number of pilot licenses issued was 220,000, an increase of 720,000 drones and 139,000 licenses compared to the end of 2023 [1]. Simultaneously, the usage scenarios of drones have become increasingly diversified, including agricultural plant protection (with over 150,000 drones in use in 2024), power inspection, emergency rescue, and emerging areas such as drone lifting (with a predicted global market size exceeding 300 billion yuan by 2025), express logistics (SF Express and Meituan's drone pilots), and medical emergency response. In terms of policies, "drone transportation" has been listed as a key industry supported by the state, with 487 logistics and distribution pilot areas currently planned.

However, despite drones being applied in multiple fields and demonstrating significant potential, they still face several issues and challenges. Firstly, the existing regulatory framework is not sufficiently comprehensive, and airspace management is relatively strict, limiting the widespread use and commercial development of drones [2]. Secondly, there is a homogenization of drone products and services, lacking innovative capabilities and differentiated advantages to meet market demands. Finally, there is inadequate collaboration between upstream and downstream industries in the supply chain, and a lack of unified standards and norms, resulting in an imperfect industrial chain and inefficiencies. Effective measures are needed to address these issues to promote better development of the drone industry.

3. Exploration and practice of UAV development models based on low-altitude economy

3.1. Technology innovation-driven model

In the booming development of China's low-altitude economy, technological innovation has become a significant driving force for advancing the UAV industry. For instance, 5G and higher-level communication technologies are revolutionizing UAV control and data transmission methods. Compared to the past, 5G communication technology enables UAVs to possess characteristics such as low latency, high bandwidth, and widespread connectivity. In the low-altitude logistics and distribution environment of urban areas, utilizing 5G technology allows UAVs to stably transmit massive amounts of information, including flight status and cargo details, to the dispatch center in real-time. Additionally, they can receive instructions promptly from the dispatch center, enabling precise obstacle avoidance and enhancing transportation efficiency and safety performance.

High-precision navigation and positioning technologies are also propelling UAV development. In complex terrain conditions such as urban high-rises, valleys, and mountainous areas, conventional GPS positioning is prone to signal deviations or interference. Nowadays, a composite navigation and positioning system that integrates satellite navigation, inertial navigation, visual navigation, and Real-Time Kinematic (RTK) positioning enables UAVs to achieve centimeter-level precise positioning [3]. In agricultural plant protection, this precise positioning allows UAVs to accurately spray pesticides or seeds evenly across fields according to preset routes, preventing repeated operations or missed applications. This not only enhances agricultural production efficiency but also reduces pesticide usage and environmental pollution [4].

With the deep application of artificial intelligence, UAVs have evolved from simple remote-controlled aircraft into intelligent entities with independent decision-making capabilities. Through machine learning algorithms, UAVs can rapidly analyze and recognize data such as images and videos [5]. In the process of power inspection, adopting artificial intelligence image recognition technology, combined with high-definition cameras and infrared thermal imagers, allows UAVs to automatically detect faults like loose wire screws, damaged insulators, and wire heating. They can promptly issue alarms, enabling monitoring of the safe and stable operation of the power system. Additionally, in security monitoring, artificial intelligence analyzes object tracking and behavior, enabling UAVs to realize real-time monitoring and alarming of suspicious individuals and abnormal activities, improving the intelligence of urban security [6].

3.2. Industrial integration development model

In the context of the low-altitude economy, the integration of multiple industrial domains has become a prominent feature of UAV development, providing a broad application prospect. The "UAV + Agriculture" model promotes the development of traditional agriculture and accelerates its transition towards smart agriculture. For large-scale field pest monitoring, equipping UAVs with multispectral cameras allows fast and comprehensive acquisition of crop growth status. By analyzing the data, accurate identification of pest species, distribution range, and severity can be achieved, guiding farmers in scientific pest control measures. Additionally, UAVs exhibit high efficiency in plant protection operations. A single plant protection UAV can spray pesticides on over 100 acres of farmland per day, which is more than ten times more efficient than manual methods. This achieves precise fertilization, saves chemical fertilizers, and reduces agricultural production costs [7].

The "UAV + Logistics" model aims to establish an efficient and low-cost air logistics system. In remote or mountainous areas with poor conditions, traditional ground logistics and distribution methods face issues like high transportation costs and low efficiency. UAV logistics can effectively address these problems, providing technical support for precise "last mile" and even "last 100 meters" delivery. For example, in certain situations where time-sensitive materials such as medical emergency supplies and fresh products are required, UAV transportation can significantly reduce delivery cycles, ensuring timely arrival. Furthermore, during major e-commerce promotions, UAVs can be utilized for urban cargo distribution, alleviating ground transportation congestion and improving delivery efficiency and consumer shopping experiences [8]. However, current UAV logistics still faces challenges like short range, limited payload capacity, and complex flight area management, requiring further technological breakthroughs and improvements in related policies and regulations.

"UAV + Culture and Tourism" presents a new opportunity and experience for the development of the tourism industry. At tourist attractions, UAVs can perform spectacular light shows. By programming and controlling the flight paths, light colors, and blinking rhythms of multiple UAVs, colorful patterns and shapes can be displayed in the sky, providing visitors with a visual feast and offering more options for local nighttime tourism [9]. Additionally, UAV aerial photography services offer tourists a brand-new perspective to appreciate scenery, capturing creative and aesthetic photos and videos that meet personalized travel experience demands. For some historical and cultural relics, UAV technology can utilize high-precision measurement instruments to achieve 3D modeling of the sites, providing data support for relic protection and monitoring, and promoting sustainable development in the tourism industry.

3.3. Policy support and guidance model

Policies play a crucial role in guiding and supporting the development of drones in the low-altitude economy. The gradual advancement of low-altitude airspace opening policies has opened up vast opportunities for the widespread application of drones [10]. In the past, strict airspace regulations limited the flight range and operating time of drones, significantly restricting their development. Nowadays, with China's pilot reform of low-altitude airspace classification management in some regions, low-altitude airspace is divided into controlled airspace, monitored airspace, and reporting airspace, clarifying access conditions and management rules for different airspaces. This allows drones to carry out various operational activities more freely and efficiently, provided they comply with the corresponding regulations. For example, in low-altitude tour projects in some tourist attractions, drones can safely provide air sightseeing services to tourists within the prescribed monitored airspace and along established routes. In the field of logistics and distribution, drones can also successfully complete cargo transportation tasks within the reporting airspace, thanks to the support of low-altitude airspace opening policies.

Industrial support policies have injected strong momentum into the growth of drone enterprises. The government encourages enterprises to increase investment in technological research and development and improve product performance and quality through various means such as financial subsidies, tax incentives, and funding for scientific research projects. For example, for enterprises focusing on key drone technology research and development, the government provides tax incentives such as additional deductions for research and development expenses, reducing enterprise research and development costs. It also establishes special industrial development funds to provide financial support for innovative and potential drone projects, helping enterprises solve funding shortages. Simultaneously, in terms of industrial park construction, the government offers preferential policies for hardware facilities, such as land and factories, to drone enterprises, attracting enterprises to gather and develop, forming industrial cluster effects, and promoting collaborative innovation and cooperation among upstream and downstream enterprises in the industry chain.

Safety regulatory policies are the cornerstone of ensuring the healthy development of the drone industry. With the rapid growth in the number of drones, flight safety issues have become increasingly prominent. To address this, the government has formulated a series of strict safety regulatory policies to regulate various aspects of drone production, sales, registration, and flight. Drone manufacturers are required to strictly adhere to relevant safety standards to ensure product quality and safety. After purchasing drones, consumers need to register them under their real names, facilitating effective management of drones by regulatory authorities. During flight, clear regulations stipulate no-fly zones, flight altitude limits, and flight qualification requirements for drones. Violations of these regulations are subject to severe penalties according to the law. The implementation of these safety regulatory policies has effectively reduced the incidence of drone flight safety accidents, maintained public safety and air order, and created a favorable environment for the sustainable development of the drone industry.

4. Strategic suggestions for promoting the development of UAVs enabled by the low-altitude economy

4.1. Strengthening top-level design

To promote the healthy development of the UAV industry, it is necessary to strengthen the planning and guidance of UAV development from the national level and formulate corresponding development policies and implementation plans. Simultaneously, a sound legal and regulatory framework should be established to clarify the scope and limitations of UAV flights, and standardize UAV operation behaviors and safety management. Furthermore, successful experiences from other industries, such as intelligent connected vehicles, can be leveraged to provide more convenient and efficient airspace management and services for UAVs.

4.2. Establishing an industrial alliance

To promote industrial collaborative innovation and resource sharing, a UAV industrial alliance can be established, integrating resources from upstream, midstream, and downstream enterprises in the industry chain, as well as research institutes and universities, to form a synergistic force. The alliance can organize technical research, product development, and application promotion to enhance the overall level and competitiveness of the industry chain. Additionally, the alliance can build a platform to facilitate communication and cooperation between upstream and downstream enterprises, promoting the transformation of innovation achievements into practical applications.

4.3. Fostering technological innovation and standardization

As an emerging industry, UAVs require continuous innovation in technology, products, and business models. By increasing research and development investment, breakthroughs and upgrades in UAV technology can be achieved, improving their intelligence and automation levels. Simultaneously, cooperation with upstream and downstream enterprises should be strengthened to jointly develop new application scenarios and expand market space. In terms of standardization, a unified set of standards and specifications for UAVs can be established, covering various aspects from design, manufacturing, to usage, to ensure the consistency of product quality and performance.

4.4. Optimizing regulatory environment

To ensure the safe operation of UAVs in public airspace, a scientific and reasonable regulatory mechanism needs to be established, achieving the goal of "effective pre-event management and post-event handling" [11]. Firstly, it is essential to strengthen the access management of UAVs, formulate strict approval processes and requirements to prevent illegal flight incidents. Secondly, routine supervision of UAVs should be enhanced, utilizing satellite positioning, data monitoring, and other means to timely detect and address illegal flight behaviors. Additionally, emergency response plans for UAV accidents should be established to ensure rapid response and proper handling in case of accidents. Finally, publicity and education should be intensified to popularize UAV knowledge and safety awareness, guiding the public to view and use UAVs correctly.

5. Conclusion

With the continuous advancement of technology and the expanding application scenarios, drones have become an indispensable part of modern society. Based on the concept of the low-altitude economy, this article proposes a drone development model that aims to promote the development and application of the drone industry by strengthening top-level design, establishing industrial alliances, fostering technological innovation, and enhancing standardization. Under this model, drones will integrate with other low-altitude sectors to form a complete ecosystem, providing more convenient and efficient travel options and services. Simultaneously, through intelligence, digitization, networking, and other means, drones can achieve a wider range of functions and applications, such as smart delivery, smart cities, emergency rescue, and more. In summary, the drone development model based on the low-altitude economy concept is expected to bring broader market prospects and development opportunities to the drone industry, making it an indispensable part of modern society.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Wang L, 2025, Discussion on the Innovation Model of Light and Small Drone Logistics and Distribution in the Greater Bay Area Driven by the Low-Altitude Economy. Logistics Engineering and Management, 47(02): 19–22.

[2] China Energy Network, 2024, The Scale and Level of the Drone Industry Continue to Improve (New Perspective), viewed November 27, 2024, https://www.zhihu.com/question/457765989

[3] Jia H, 2025, Challenges and Responses of Drone Logistics in the Development of Low-Altitude Economy. China Aviation Weekly, (04): 51–53.

[4] Chen H, 2024, Discussion on the Role and Significance of Drone Channel Surveying and Mapping in the Development of Low-Altitude Economic Strategy. Pearl River Water Transport, (22): 11–14.

[5] Zhang J, Chen L, Gao Z, et al., 2025, Research Status and Prospects of Low-Altitude Drone Technology. Engineering Sciences, 27(02): 73–85.

[6] Han Z, Lu W, Wang A, 2025, Application and Technological Innovation of Drones in the Agricultural Low-Altitude Economy. Shanghai Informatization, (06): 35–38.

[7] Zhang M, 2025, Yishui, Shandong: Exploring the Development Model of Plant Protection Drones Empowered by the Low-Altitude Economy. China Rural Science and Technology, (05): 46–47.

[8] Li J, Zhao Y, 2024, The Concept and Practice of the County-Level Drone Government Service Flight System in the Context of the Low-Altitude Economy. China Security, (06): 83–86.

[9] Hu Q, 2025, Media Application of Drones in the Context of Low-Altitude Economy. Shanghai Informatization, (01): 18–21.

[10] Qiao L, Huo D, Gu W, et al., 2025, Research on the Optimization of the Development Path of Low-Altitude Economic Modernization—Based on the Perspective of 3D Drone Flight. Scientific Management Research, 43(02): 29–38.

[11] Chen Y, Yang J, Zhang Y, et al., 2025, Policies, Standards, and Technologies for Drone Communication Spectrum Management for the Low-Altitude Economy. Data Acquisition and Processing, 40(01): 2–26.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Design of a Pressure Sensor Array System Based on Minecraft

Ximing Luo*

Shanghai No.4 High School, Shanghai, China

**Author to whom correspondence should be addressed.*

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: Multimodal information sensing becomes increasingly critical under the rapid development of automation and information technology. With the ability to provide high-density and high-sensitivity pressure detection, pressure sensor arrays have been applied to a variety of fields, including intelligent robotics, medical monitoring, and industrial automation. This study proposes a pressure sensor array system based on the Minecraft game platform. The simulation and testing of the pressure sensor arrays system have been conducted using redstone circuits and pressure plates in Minecraft to simulate real-world piezoelectric pressure sensor arrays. A series of experiments verified the feasibility and effectiveness of the system.

Keywords: Pressure sensor array; Minecraft; Redstone circuits

Online publication: October 15, 2025

1. Introduction

With the swift advancement of automation and information technology, accurate environmental sensing and control have become increasingly important. Pressure sensors are a crucial type of sensing device that enables the detection and measurement of pressure changes. Therefore, they are widely applied to various automation control systems. While most traditional pressure sensors can only enable single-point measurements, pressure sensor arrays enhance data acquisition by enabling multi-point measurements, thereby facilitating more precise environmental sensing.

In fields such as intelligent robotics, medical monitoring, and industrial automation, there is an increasing demand for pressure detection systems that exhibit high precision and sensitivity. Consequently, pressure sensor arrays have emerged as a prominent topic of research, owing to the capacity to deliver multipoint, high-density pressure measurements.

Capacitive, resistive, and piezoelectric sensors are prevalent types of pressure sensors, each characterized by distinct operational principles and specific application scenarios.

Minecraft is a highly creative sandbox game with a built-in redstone circuit system that provides a unique platform for simulating electronic circuits. The redstone circuit system is programmable, versatile, and creative, allowing players to create complex circuits and mechanical systems through the arrangement and configuration of diverse redstone components. In Minecraft, the implementation of redstone circuits in conjunction with pressure plates allows for the simulation and evaluation of real-world pressure sensor arrays, including capacitive, resistive, and piezoelectric types [1].

Therefore, this study aims to investigate the application of redstone circuits and pressure plates within the Minecraft environment to simulate the pressure sensor array system. The research focuses on the following elements.

2. Principle of redstone circuits

The working principle of redstone circuits is based on the signal transmission of redstone dust. Redstone dust can connect various redstone components, thereby facilitating the creation of intricate circuit systems. Therefore, the fundamental principles of redstone circuits help to understand the function of each individual component and comprehend the signaling mechanisms inherent in these circuits.

Additionally, investigations were conducted on associated digital logic circuits. The operational principles and applications of basic logic gates were researched, including AND Gates, OR Gates, and NOT Gates. Furthermore, the significance of flip-flop and latch in data storage and processing was studied.

The pressure sensor was built based on Minecraft redstone circuits. Utilizing a combination of redstone dust and a pressure plate, a redstone signal can be activated upon the presence of a player or other entity on the pressure plate. A simple circuit connecting the pressure plate to the redstone dust can ensure effective signal transmission.

Design and implementation of a test circuit for the pressure sensor array. An array equipped with multiple pressure sensors was designed. Each sensor is responsible for detecting pressure changes in a specific area. Using redstone repeaters and comparators, signals generated by each sensor were integrated to form a unified output signal. According to the design diagram, test circuits were constructed using components such as redstone dust, repeaters, and comparators.

3. Components of redstone circuits

3.1. Redstone circuits

Redstone circuit is a unique circuitry system in Minecraft that enables players to create complex logic control and signal transmission systems by placing and configuring a variety of redstone components [2]. Fundamentally, redstone circuits possess programmability, versatility, and creativity, enabling users to construct a diverse array of mechanical and automated systems. Redstone components include, but are not limited to, redstone dust, redstone torch, redstone repeater, redstone comparator, etc., which together form the basis of redstone circuits.

3.2. Redstone components

Redstone components are fundamental for constructing redstone circuits. The detailed description of some key components is listed below:

(1) Block of redstone: As a permanent power source, the block of redstone provides a constant redstone signal. Therefore, it often functions as a power source or signal source for circuits.

(2) Button: Button is a common input component. The player can click it to generate a brief signal pulse that activates the connected circuits.

(3) Daylight detector: This component generates redstone signals of varying intensity depending on the daylight intensity. It is often used to simulate daylight changes or be a functional part of a timer.

(4) Detector rail: The detector rail detects the passage of a mine car. It sends out a signal to activate the connected circuits when the car passes.

(5) Levers: Similar to buttons, levers can be operated by the player to activate or deactivate circuits. Moreover, their design facilitates a more intuitive method of operation.

(6) Pressure plate: When the actions that a player or creature steps on are detected, the pressure plate sends out signals for the activation of circuits.

(7) Redstone torch: The redstone torch not only functions as a light source, but also functions as a signal inverter and transmission component for the circuits.

3.3. Digital logic principles

The digital logic principles of redstone circuits involve basic logic gates that are critical to complex circuits. The introduction to some of the basic logic gates is listed below:

(1) NOT gate: NOT gate is one of the most basic logic gates. It is a gate used when an opposite output is wanted from the input given. When the switch, or input, is set to "on" (1), the output toggles to "off" (0), and vice versa.

(2) AND gate: AND gate requires that the output is toggled to "on" only when all inputs are "on." This logic gate is often used in cases where multiple conditions need to be met simultaneously.

(3) OR gate: An OR gate is a gate that, whenever any input is "on," the output is also "on." This logic gate is often used for "at least one" condition determination.

(4) XOR gate: An XOR gate is a gate that uses two inputs and the output is toggled to "on" when one switch is "on" and one switch is "off." This type of logic gate is very useful in figuring out whether two signals are equal [3].

Flip-flops and latches are the crucial components for temporal logic implementation in redstone circuits. They can retain states and modify their outputs in response to input signals. Some of the common flip-flops and latches are introduced as follows:

(1) T flip-flop: A T flip-flop is a basic flip-flop whose output toggles its state whenever the input signal T changes from OFF to ON. This type of flip-flop is commonly used to implement frequency dividers or pulse generators.

(2) RS latch: The RS latch functions as a circuit storing a two-bit binary number. Its output depends on the state of the input signals R and S. This latch is significant in data storage and transmission.

(3) D flip-flop: The D flip-flop is a data flip-flop whose output varies in accordance with the state of the input signal D. This type of flip-flop is highly beneficial for data synchronization and timing regulation.

4. Pressure sensor based on Minecraft redstone circuits

4.1. Construction of a single pressure sensor

4.1.1. Principle of construction

The simulation of the piezoelectric pressure sensor employs the signal generation capabilities inherent in redstone circuits. The redstone signal is generated when the pressure plate is pressurized, simulating the charge generated by the piezoelectric material when it is pressurized. The simulation can detect and measure pressure. In Minecraft, the redstone signal generated by a pressure plate can activate a diverse array of redstone components, including gates and pistons. Therefore, the simulation can facilitate pressure detection and subsequent responses.

4.1.2. Simulation of a single pressure sensor

Initially, we conducted a simulation involving a single pressure sensor. The operational response of the pressure sensor is simulated by positioning a pressure plate and monitoring the redstone signal it generates. We established a connection between the pressure plate and a redstone lamp using redstone dust, thereby enabling the redstone lamp to activate and deactivate in response to the signal generated by the pressure plate [4].

As shown in **Figure 1**, when there is no object placed on the pressure plate, the redstone lamp is not illuminated, and no energy is generated. As shown in **Figure 2**, when the game character steps on the pressure plate, the redstone lamp is illuminated and energy is generated. The experimental results indicate that the redstone signal generated by the pressure plate exhibits a strong correlation with variations in pressure.

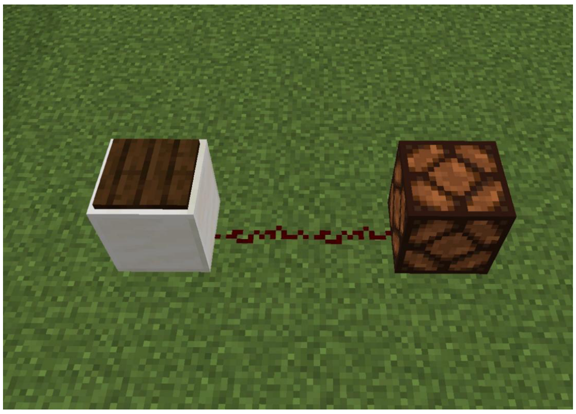


Figure 1. Single pressure sensor (no object on the pressure plate)

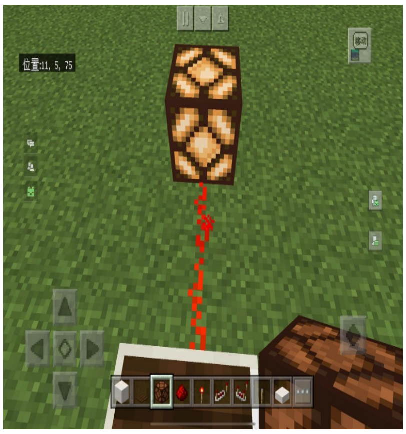


Figure 2. Single pressure sensor (game character standing on the pressure plate)

4.2. Construction of sensor array

Subsequently, a 3×3 pressure sensor array was developed. As illustrated in **Figure 3**, nine blocks of quartz were systematically organized in a 3×3 matrix configuration, with a pressure plate positioned atop each block of quartz.

Figure 3. 3×3 pressure sensor array

4.3. Construction of test circuit

To facilitate the detection of electrical signals from the 3×3 pressure sensor array, a measurement system was established (**Figure 4**). Two redstone lamps were strategically positioned at each of the four edges of the 3×3 pressure sensor array to monitor the eight peripheral pressure plates. The signal generated by the center-positioned pressure plate was routed using a lower lead to one of the redstone lamps, so that the detection of this pressure plate is achieved.

As illustrated in **Figure 5**, the illumination of the outer redstone lamp occurs when the game character stands on the center-positioned pressure plate. When the game character stands on the peripheral pressure plate, the redstone lamp adjacent to it is illuminated (**Figure 6**). It is evident that the constructed small-scale pressure-sensing array possesses the requisite measurement capability.

Figure 4. 3×3 pressure sensor array and test circuit

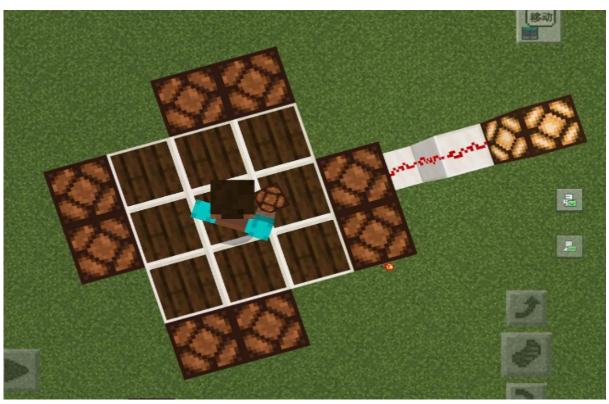


Figure 5. Game character standing on the center-positioned pressure plate

Figure 6. Game character standing on the peripheral pressure plate

5. Design and testing of large-scale pressure sensor array

5.1. Principle analysis and measurement scheme

5.1.1. Problem elicitation

In the game Minecraft, players can construct a fundamental pressure sensor array by effectively utilizing pressure plates. These sensors enable the detection of a player or an entity, thereby activating designated redstone circuits. However, as the scale of the sensor array grows, the previously simple design of the measurement circuitry will encounter significant challenges. If we continue to employ the methodology outlined in Section 4, specifically the individual connection of each pressure plate via redstone wires, the complexity and manageability of the wiring for the entire test circuit will significantly deteriorate as the quantity of pressure plates increases. Therefore, addressing this issue necessitates a comprehensive optimization of the design for the test circuit.

5.1.2. Solution

Figure 7 illustrates the working principle of the pressure sensor. The pressure plate functions as a sensing component. It is capable of transducing physical pressure into redstone signals. This process represents the initial phase in the development of a pressure detection system. The redstone circuits enable the signal transmission, transferring the sensing information from the pressure plate to the subsequent redstone devices.

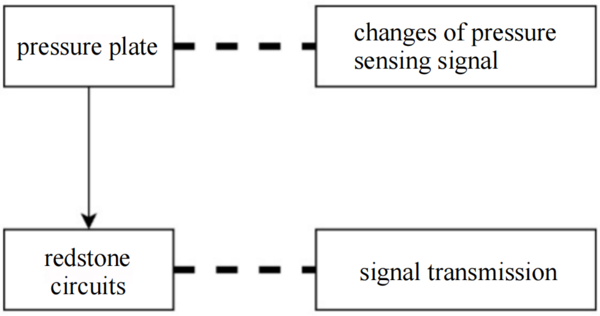


Figure 7. Schematic diagram of pressure sensing

Practical observations indicate that energy attenuation occurs in the transmission of redstone signals.

Specifically, as the distance over which the signal propagates increases, there is a corresponding gradual decrease in its strength. As shown in **Figure 8**, energy is conveyed through the redstone wire. The energy source exhibits the highest concentration of redstone energy, represented by a bright red point. As the energy is transmitted along the redstone wire, the energy intensity diminishes, resulting in a gradual transition to a darker hue. Ultimately, this energy diminishes, as indicated by the black point within the dashed box (indicating a complete absence of energy).

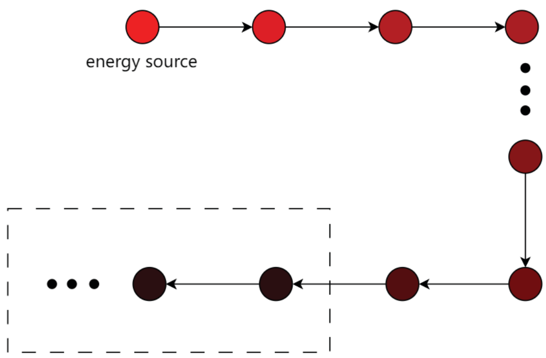


Figure 8. Energy attenuation of the signal transmission in redstone circuits

Figure 9 provides a quantitative analysis of the energy attenuation, demonstrating a linear correlation between redstone energy and path length. It indicates that redstone energy diminishes as path length increases. This relationship can be mathematically represented by the following equation:

$$\text{Redstone energy value} = 15 - \text{path length value} \quad (1)$$

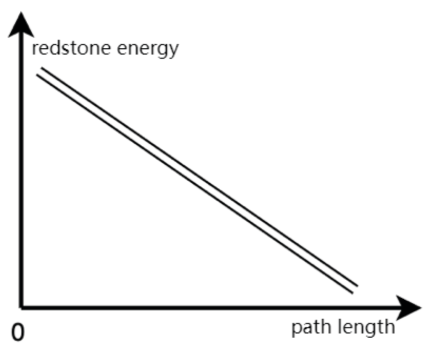


Figure 9. Relationship between redstone circuit signal transmission distance and path length

This property of energy attenuation is an important consideration for redstone circuit designers. Comprehending and implementing this principle can assist the designer in optimizing the circuit layout. For instance, the utilization of redstone repeaters can facilitate the extension of signal transmission distances or minimize energy loss by reasonable design of the circuit pathway.

In addition, deep insights into signal attenuation provide a theoretical basis for designing more complex pressure-sensing arrays. The distance from the energy source, specifically the pressure point, can be deduced from the redstone energy value recorded at the measurement point.

As shown in **Figure 10**, a circuit layer is set up under the sensing layer (pressure plate layer). Redstone dust is positioned beneath each pressure plate within the sensing layer, and this redstone dust is interconnected to create a redstone grid. Each point within the grid possesses the capacity to be activated by the corresponding pressure plate situated above it.

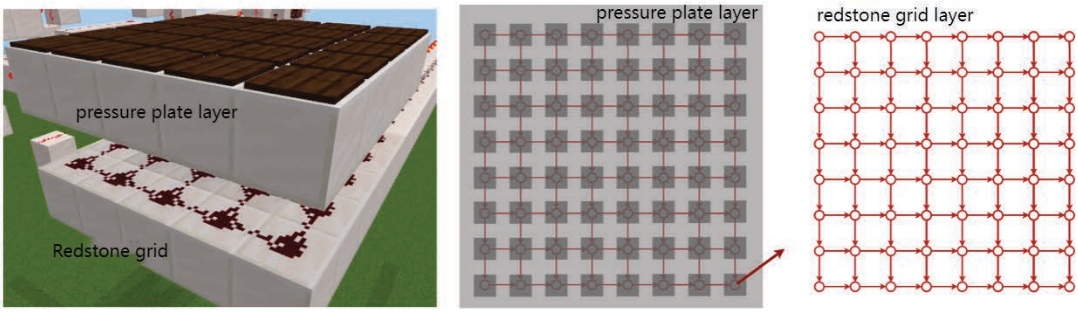


Figure 10. Redstone grid

This study investigates the energy attenuation of the redstone grid by experiments. As illustrated in **Figure 11**, a 4×16 redstone grid was constructed, with the lever positioned at the grid points located in the third row and third column to supply energy. Observations of the color variations across the redstone grid indicate that the intensity of color at the grid points correlates positively with proximity to the energy source. In other words, grid points closer to the energy source exhibit a brighter hue, indicative of a higher energy value. Notably, the energy value appears to be solely dependent on distance rather than the direction of redstone energy transmission. This hypothesis was verified by measurements taken from energy detectors situated in the first row, as well as the fourth and eighth column positions.

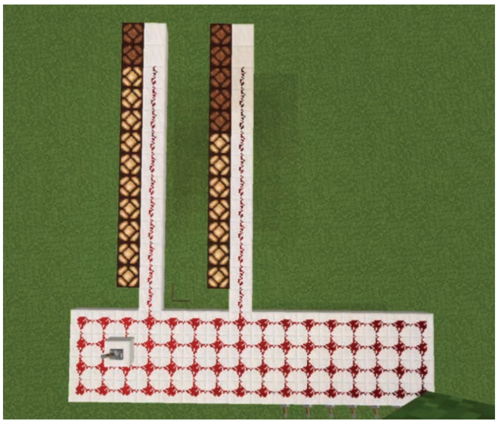


Figure 11. Experimental validation of redstone grid energy attenuation

In redstone circuit design within Minecraft, the accurate placement of stressors is accomplished through an intricate network of redstone circuits, as well as the strategic positioning of detection points at critical junctures. This methodology bears resemblance to the mathematical problem of determining the intersection of two circles.

We consider that there is a scenario involving two circles, where the center of each circle functions as a detection point, and the radius of each circle denotes the detection range associated with that point. The objective is to identify the intersections of these two circles, which may suggest the potential location of stressors. This situation parallels a mathematical problem in which the coordinates of the centers and the radii of the two circles are known, and the task is to determine their common intersection points.

By reformulating the detection of redstone circuits as a mathematical problem, the problem can be addressed according to geometric principles. This methodology not only enhances the positioning precision but also renders the entire design process more systematic and scientific. By calculating the distance between the centers of two circles and juxtaposing it with their respective radii, we can ascertain whether the circles intersect and identify the precise coordinates of the intersection point. In **Figure 12**, the performance of the detection point is quantified on a scale from 0 to 15, which corresponds to detection distances ranging from 1 to 15 grids. A detection range can be defined as a 16×16 grid area centered on the detection point, encompassing all potential detection points. However, when the detection point is situated at a corner of the sensing array, its effective detection range is represented as a right triangular area because the detection area at the corner is constrained by the boundaries of the array. In this right triangle, the two right-angle sides illustrate the detection capabilities of the point in both horizontal and vertical orientations, respectively.

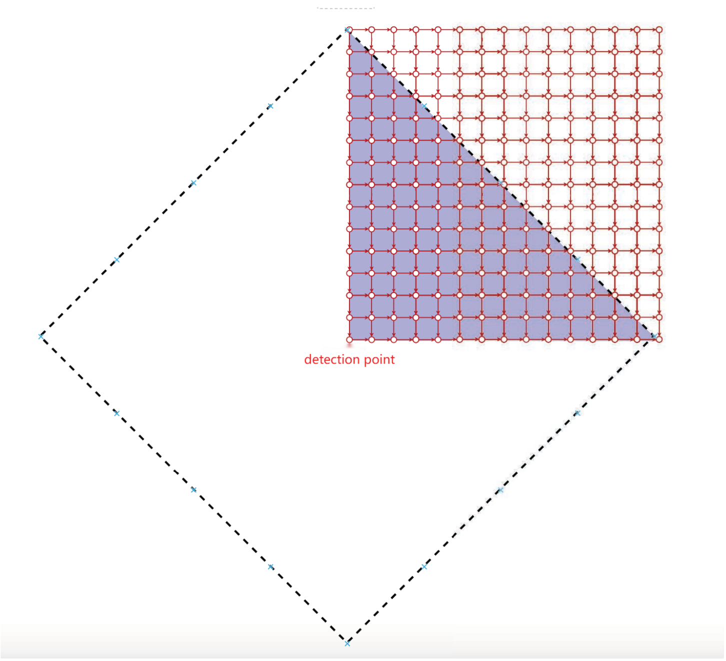


Figure 12. Detection point coverage

5.2. Detection circuit analysis

5.2.1. Conversion of energy values to digital quantities

According to the analysis in Section 5.1., the initial step in identifying the location of the pressure point involves the precise calculation of the energy values at the detection point. Considering that redstone circuits inevitably undergo energy attenuation during transmission, this attenuation can substantially influence the energy values. To address this issue, it is essential to transform the analog energy values into more manageable digital quantities. Accordingly, we propose two methodologies for facilitating the conversion from analog to digital quantities:

The first method leverages the attenuation characteristics of energy transmission. Specifically, a redstone wire is established from the measurement point, with a redstone torch positioned adjacent to the quartz block beneath the redstone wire. The transmission of redstone energy results in the extinguishing of the redstone torch. Therefore, the identification of the maximum energy transmission path can be achieved by monitoring whether the redstone torch extinguishes. This pivotal moment, at which the redstone torch ceases to extinguish, can be regarded as a clear indicator of energy attenuation. By employing Equation (1), it is possible to retroactively determine the precise location of the measurement point. A detailed example of the circuit design utilized for this conversion is illustrated in **Figure 13**.

Figure 13. Energy numerical conversion method 1

The second methodology is more concise. As illustrated in **Figure 14**, the 16 redstone torches positioned at the bottom each represent different energy values. The illuminated torches on the left side correspond to lower energy values, while those on the right side indicate higher energy values. Within **Figure 14**, a toggle mechanism is employed to achieve a measurement point with an energy value of 11, which is illuminated as the 11th torch from the left among the redstone torches at the bottom.

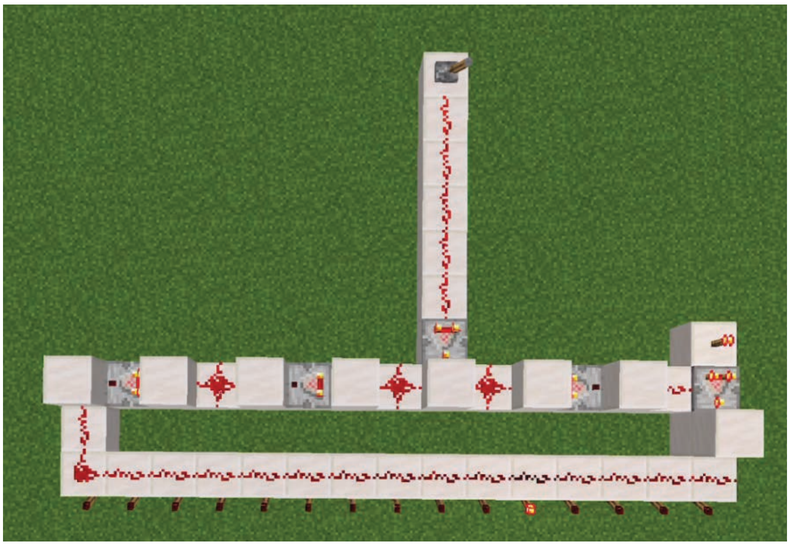


Figure 14. Energy numerical conversion method 2

5.2.2. Hexadecimal to binary conversion

According to the analysis of the previous section, the energy value is converted to hexadecimal numerical quantities. To facilitate the subsequent calculations, it is essential to convert the hexadecimal digital quantities to binary digital quantities. Therefore, the calculation process for the measurement points is shown in **Figure 15**.

The truth table for hexadecimal to binary conversion is shown in **Table 1**, based on which the hexadecimal to binary conversion circuit is designed.

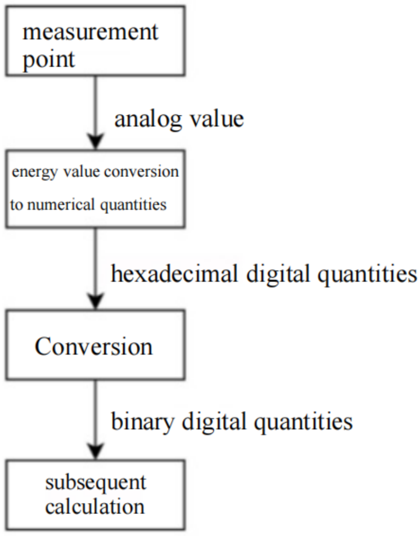


Figure 15. Calculation process for the measurement point

Table 1. Conversion from hexadecimal quantities to binary quantities.

Hexadecimal quantities	Binary quantities			
0	0	0	0	0
1	0	0	0	1
2	0	0	1	0
3	0	0	1	1
4	0	1	0	0
5	0	1	0	1
6	0	1	1	0
7	0	1	1	1
8	1	0	0	0
9	1	0	0	1
10	1	0	1	0
11	1	0	1	1
12	1	1	0	0
13	1	1	0	1
14	1	1	1	0
15	1	1	1	1

Figure 16 illustrates a highly efficient circuit for converting 1-bit hexadecimal numbers to 4-bit binary representations. In this configuration, a hexadecimal input signal is processed by the circuit to convert the

corresponding 4-bit binary output. The conversion logic is predicated on the mapping relationship delineated in **Table 1**, which associates each hexadecimal digit with its equivalent four-bit binary representation. When a bit of the hexadecimal input is activated (high), the corresponding output line in the binary digit also activates, thereby facilitating a direct signal conversion. This conversion circuit has significant advantages in enhancing data processing efficiency and simplifying circuit design.

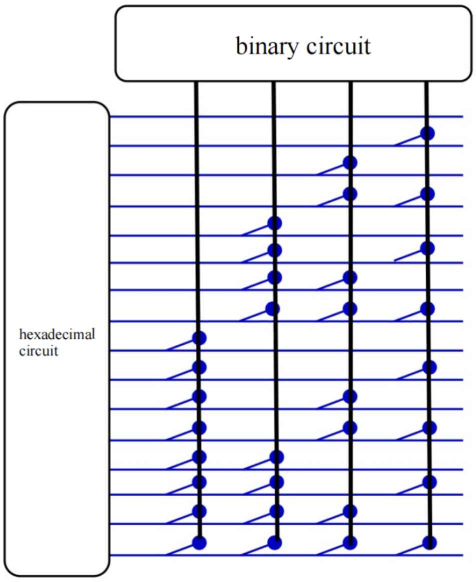


Figure 16. Principle of 1-bit hexadecimal to 4-bit binary number conversion circuit

According to Figures 16 and 17, the circuit was built in Minecraft. The measurement effect is shown in **Figure 18**. The corresponding measurement points of the foxes in different positions have different values: the distance from the foxes to the measurement point (path distance, horizontal and vertical distances are summed up) is "3," "4," "5," "4," "5," "7." According to **Table 1**, the values corresponding to the binary quantities are "0011," "0100," "0101," "0111," respectively.

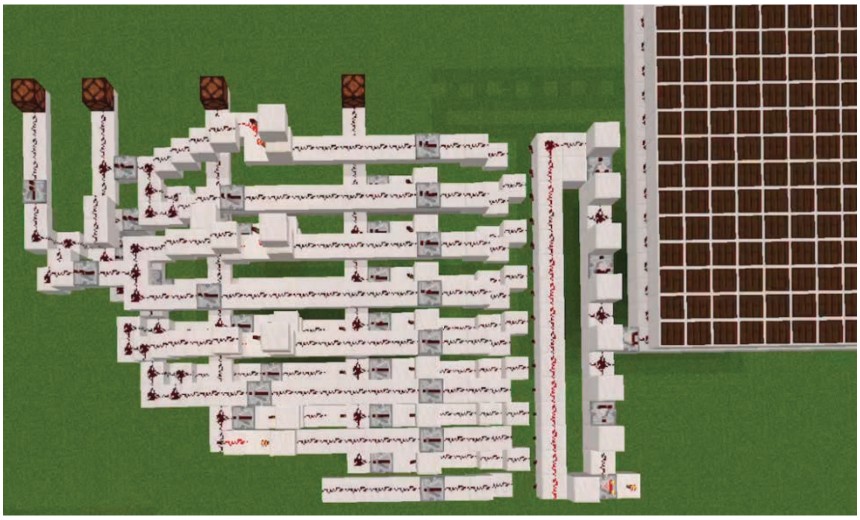


Figure 17. Realization of 1-bit hexadecimal to 4-bit binary number conversion circuit

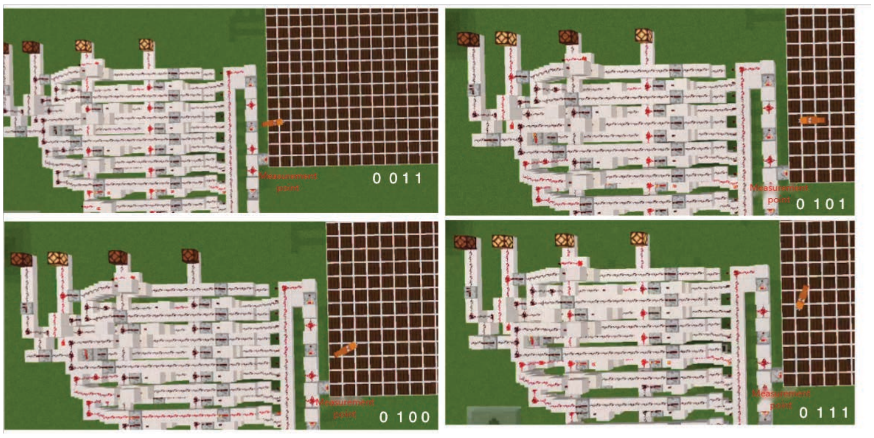


Figure 18. Measurement effect

5.3. Case study: 8×8 pressure sensor array

An analysis of the 8×8 pressure sensor array is conducted, with the model illustrated in **Figure 19**. This sensor array is comprised of two distinct components: the lower section, which consists of 64 blocks of quartz organized in rows, and the upper section, where each block of quartz is positioned atop a pressure plate.

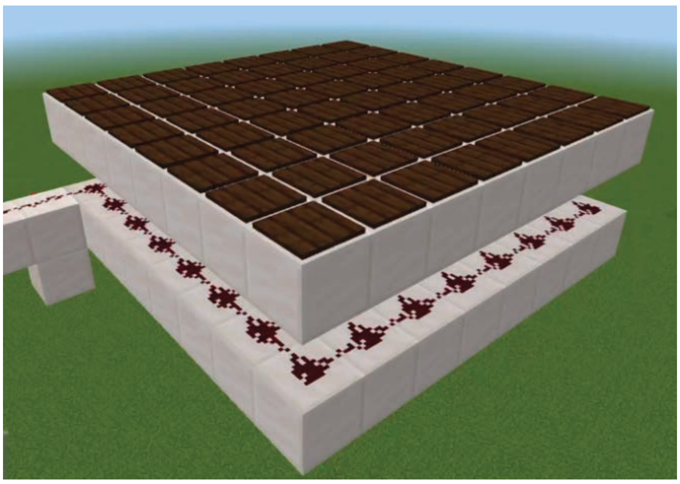


Figure 19. 8×8 pressure sensor array

For the 8×8 pressure sensor array, two measurement points can be utilized for calculation. As illustrated in **Figure 20**, the blue triangle area represents the measurement range for measurement point 1, while the red triangle area denotes the measurement range for measurement point 2. The gray square area indicates the sensing range of the 8×8 pressure sensor array. **Figure 20** evidently indicates that the sensing range is concurrently located within both the blue and red triangular areas.

By detecting the energy of measurement point 1 and measurement point 2, the Manhattan distance from the pressure point to measurement point 1 and that to measurement point 2 can be calculated. In **Figure 20**, the Manhattan distance of the game character from the measurement point 1 is l1, and that from the measurement point 2 is l2. The coordinates of the game character's position are denoted as (x, y). Consequently, the conditions for x and y are expressed as Equation (2):

$$\begin{cases} x + y = l_2 \\ x + 8 - y = l_1 \end{cases} \quad (2)$$

Therefore, (x, y) can be calculated according to Equation (2), and the 8×8 pressure sensor array completes the measurement.

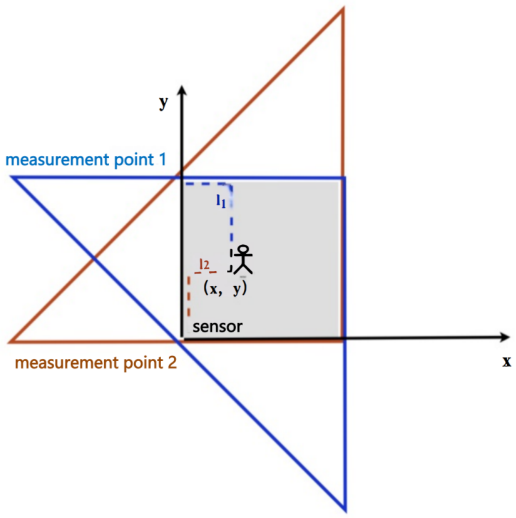


Figure 20. Measurement principle of 8×8 pressure sensor array

6. Conclusion

This paper proposed a pressure sensor array system based on Minecraft. The simulation of the operational principles of capacitive, resistive, and piezoelectric pressure sensors has been conducted, and the implementation of redstone circuits and pressure plates facilitated the simulation and evaluation of the pressure sensor array. The experimental results indicate that the system is capable of effectively simulating the detection and control processes concerning pressure sensor arrays, demonstrating significant feasibility and potential for application. Future research may focus on optimizing the redstone circuits design, enhancing the stability and accuracy of the system, and investigating its applicability in real-world scenarios.

Disclosure statement

The author declares no conflict of interest.

References

[1] Oussar Y, Margo C, Lucas J, et al., 2023, Fast Circular Shapes Detection in Cylindrical ECT Sensors by Design Selection and Nonlinear Black-Box Modeling. COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering, 36(1): 2–17.

[2] Chen C, 2019, Redstone Electronics: Learning Logic Circuits and Computer Principles by Playing My World Minecraft, Posts and Telecommunications Press.

[3] Lu X, 2014, Analysis of Physical Principles of Resistive Sensors. Physics Teacher, (8): 56–58.

[4] Yao S, Mo Z, Nie L, 1990, Titrations with Piezoelectric Monitoring. Analytica Chimica Acta, 229: 205–212.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

A Review of AI-Driven Optimization Technologies for Distributed Photovoltaic Power Generation Systems

Nanting Li*

Electrical Engineering College, Hebei University of Technology, Tianjin 300401, China

**Author to whom correspondence should be addressed.*

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: The rapid development of artificial intelligence (AI) technology, particularly breakthroughs in branches such as deep learning, reinforcement learning, and federated learning, has provided powerful technical tools for addressing these core bottlenecks. This paper provides a systematic review of the research background, technological evolution, core systems, key challenges, and future directions of AI technology in the field of distributed photovoltaic power generation system optimization. At the same time, this paper analyzes the current technical bottlenecks and cutting-edge response strategies. Finally, it explores fusion innovation directions such as quantum-classical hybrid algorithms and neural symbolic systems, as well as business model expansion paths such as carbon finance integration and community energy autonomy.

Keywords: AI optimization; Distributed photovoltaic systems; Virtual power plant coordination; Community energy autonomy

Online publication: October 15, 2025

1. Introduction

The urgency of global efforts to address climate change, particularly the widespread adoption of "carbon neutrality" goals, has become the core driving force behind profound transformations in the energy system [1]. Against this backdrop, distributed photovoltaic (PV) power generation—with its notable cleanliness, deployment flexibility, and inherent advantage of being close to electricity consumption points—is being scaled up globally at an unprecedented pace, gradually transitioning from a "supplementary role" to a "mainstay" in the energy system [2]. However, the large-scale, high-proportion integration of distributed PV power generation has also posed significant challenges to traditional power systems. Its power generation output is highly dependent on meteorological conditions, exhibiting pronounced intermittency and volatility, which places enormous pressure on the grid's real-time power balance and frequency stability [3]. The widespread decentralized connection of distributed power sources to distribution grids, coupled with their inherent low inertia characteristics and the complexity of coordinating control across multiple nodes, has increasingly highlighted grid stability issues [4]. Additionally, efficiently integrating a vast number of decentralized PV units to maximize overall benefits faces challenges such as poor information exchange and low efficiency in collaborative decision-making [5]. These core bottlenecks severely hinder further improvements in the economic viability, reliability, and sustainability of distributed PV power generation systems [6]. In response to these challenges, artificial intelligence (AI) technology, with its exceptional capabilities in data processing, pattern recognition, complex decision-making, and adaptive learning, has gradually been recognized as a key enabling tool for addressing the optimization barriers of distributed photovoltaic systems [7]. Its technological evolution has exhibited clear stage-specific characteristics. In the initial stage (approximately 2020–2022), research focused primarily on using specific AI technologies to address optimization issues in individual components. For example, long short-term memory (LSTM) networks were used for short-term power generation forecasting [8], while convolutional neural networks (CNNs) were applied for the automatic identification and detection of PV module fault images [9]. While these studies achieved some success and validated the effectiveness of AI in specific tasks, they generally suffered from the limitation of addressing issues in isolation, lacking a systematic and coordinated optimization approach for photovoltaic systems as complex organic wholes. Optimization modules were often disconnected from one another [10]. Entering the breakthrough phase (approximately 2023–2024), the research perspective has begun to broaden, with a greater emphasis on system-wide and data-driven collaboration. The introduction of federated learning technology has provided an innovative approach to addressing the inherent "data silo" issues of distributed photovoltaic data (data dispersion and high privacy protection requirements) [11], enabling collaborative training of more powerful global models while preserving local data privacy. At the same time, digital twin technology has begun to be applied in the distributed photovoltaic field. By constructing a virtual mapping of physical systems and combining real-time data with historical information, dynamic simulation, prediction, and optimization strategy rehearsals can be conducted in virtual space, significantly improving the accuracy and foresight of system optimization decisions [12]. Currently, we are entering the integration phase (2025 to present), characterized by AI technology being more deeply integrated into the architecture and operational models of energy systems. Virtual power plants (VPPs), as an effective model for integrating distributed resources into grid operations and power market transactions, have seen a qualitative leap in their intelligence levels due to the application of AI [13]. AI technology is deeply integrated with blockchain (ensuring transparent and secure transactions) and edge computing (enabling local rapid response), collectively forming an intelligent foundation that supports multi-energy entities participating in market transactions and achieves efficient cross-platform resource scheduling and collaborative optimization [14]. The core objective of this phase is to achieve adaptive, self-optimizing, and self-coordinated operation of distributed photovoltaic systems across multiple spatio-temporal scales [15].

2. Core technology system and optimization path

The core of AI-driven distributed photovoltaic power generation system optimization lies in building a comprehensive technical system that spans data sensing, predictive decision-making, and collaborative control, thereby achieving intelligence from micro-components to macro-systems.

2.1. Intelligent sensing and data governance

The cornerstone of optimized decision-making is high-quality, multi-dimensional real-time data. In distributed photovoltaic scenarios, data collection faces challenges such as widely distributed nodes, complex environments, and diverse (heterogeneous) data types. Low-power wide-area network (LPWAN) technologies, such as LoRaWAN (long-range wide-area network) and NB-IoT (narrowband IoT), have become the ideal choice for connecting widely distributed edge sensing nodes (monitoring light intensity, component temperature, inverter operating parameters, environmental temperature and humidity, local grid load status, etc.) [16], enabling the low-cost, low-power, and reliable transmission of massive amounts of data. However, data collected in actual operation inevitably contains noise, missing values, or even outliers, which can severely impact the accuracy of subsequent models if used directly. Therefore, data governance has become an indispensable backend process in intelligent sensing. In addition to traditional cleaning, interpolation, and standardization methods, advanced AI technologies such as generative adversarial networks (GANs) have been innovatively applied to the field of anomaly data repair [17]. By training the generator to simulate the distribution of real data and using the discriminator to distinguish between real and generated data, GANs can learn the intrinsic patterns of data even in the absence of complete annotations, thereby more effectively identifying and repairing abnormal data points [18]. This significantly improves the overall quality and reliability of the dataset, laying a solid foundation for subsequent precise analysis and decision-making.

2.2. Dynamic prediction and adaptive optimization

Accurate forecasting is key to addressing photovoltaic variability and enabling proactive management. Power forecasting models themselves have undergone significant iterations driven by AI technology. Early research primarily relied on single time series models, such as LSTM, which primarily utilized historical power generation data. However, photovoltaic power generation is influenced by a variety of spatiotemporal factors, such as weather patterns (e.g., cloud movement) and complex terrain. Spatio-temporal graph neural networks (ST-GNN) represent a significant breakthrough in recent years [19], naturally modeling the spatial correlations between nodes (geographic locations) and the dynamic temporal evolution patterns within distributed photovoltaic systems. By integrating spatial information such as cloud maps and irradiance distributions provided by meteorological satellites, historical power generation curves of each node, and meteorological forecast data into the ST-GNN model, it is possible to comprehensively capture the complex spatio-temporal dependencies affecting power generation, effectively controlling the error rate of short-term predictions below 3% [20], significantly outperforming traditional models. Prediction is a means, and optimization is the goal. AI also plays a central role in the coordinated scheduling of energy storage systems with photovoltaic systems. Considering the inherent uncertainty of predictions and the real-time changes in operational conditions, a "dual-timescale optimization" strategy based on reinforcement learning (such as Q-learning, deep deterministic policy gradient DDPG, etc.) has become the mainstream approach [21]. This strategy establishes an initial plan for energy storage charging and discharging based on relatively accurate predictions at the "intraday planning" scale (e.g., several hours in advance); at the "real-time correction" scale (minute-level or even second-level), it utilizes the latest ultra-short-term predictions and actual system status information to make online decisions through reinforcement learning agents, dynamically adjusting the charging and discharging power and status of energy storage to smooth power fluctuations, participate in frequency regulation, or engage in arbitrage. This hierarchical, progressive, and real-time response strategy not only maximizes the regulatory value of energy storage but also effectively extends the cycle life of energy storage systems by 15–20% [22] through optimized charging and discharging depth and frequency, significantly reducing the lifecycle cost.

2.3. System-level collaborative control

The full value of distributed photovoltaic systems can only be realized through system-level coordination. VPPs, as platforms for integrating distributed resources, rely on intelligent decision-making at their core. AI technology plays the role of the "brain" in this context [23]. A peer-to-peer (P2P) energy trading platform built on blockchain smart contracts is an important application of VPPs. AI algorithms (such as multi-agent reinforcement learning and game theory optimization) can analyze real-time data on distributed PV power generation forecasts, energy storage status, user load demand curves, and power market price signals within a region. They dynamically match energy suppliers with demand, formulate optimal pricing strategies (such as dynamic pricing) to incentivize supply-demand balance, and maximize overall economic benefits [24]. Actual cases (such as the 5MW demonstration project in Jiaxing, China) demonstrate that such AI-driven VPP platforms can effectively reduce operational costs (by up to 12%) [25] and enhance local energy consumption rates. On the other hand, the high penetration of distributed PV systems in distribution grids, especially in weak grid scenarios (with low short-circuit capacity), can easily trigger grid stability issues such as harmonic resonance and voltage over-limit. AI also shows potential in power electronics control applications [26]. For example, by optimizing the parameters of the LCL filter at the front end of the grid-connected inverter and combining advanced resonance suppression algorithms such as dual current feedback control, AI-assisted controllers can more effectively suppress specific harmonics, significantly reducing total harmonic distortion (THDi) [27] and enhancing the system's stable operation under complex grid conditions. This "device-level intelligent control" is the technological foundation for ensuring the friendly grid connection of large-scale distributed PV systems.

3. Challenges and response strategies

Although AI-driven optimization technologies hold great promise, their practical implementation in engineering applications still faces a series of significant challenges that require ongoing research and technological innovation to overcome. The following is an analysis of the technical bottlenecks. First is the sharp contradiction between computing power and energy efficiency: the optimization of distributed photovoltaic systems, particularly real-time prediction and online control, often requires the deployment of AI models at the edge (such as field station controllers, smart inverters, or even local gateways) to achieve rapid response. However, edge devices are typically constrained by computational power, memory size, and power consumption budgets. Deploying complex deep learning models (such as large ST-GNNs or DDPG agents) on resource-constrained edge devices presents significant challenges [28]. High computational loads not only increase inference latency, making it difficult to meet real-time requirements (such as millisecond-level control), but also significantly increase device energy consumption, which contradicts the energy-saving and carbon-reduction objectives of photovoltaic systems. This constitutes the primary bottleneck constraining the deep application of AI in distributed photovoltaic systems [29]. Additionally, there are issues of cross-system compatibility and lagging standardization: distributed photovoltaic systems involve numerous equipment manufacturers and subsystems, including different models of photovoltaic inverters, various types of energy storage systems, various sensors, energy management systems (EMS), and grid dispatch systems. Currently, the international standard system for AI photovoltaic system interoperability is still incomplete [30]. Although IEC 61850 is an important standard for substation automation, its extensions targeting distributed energy (especially in combination with AI applications), such as IEC 61850-7-420, have progressed relatively slowly [31] and exhibit discrepancies in practical implementation. This directly leads to the "fragmentation" of communication protocols between devices. Different devices may use various proprietary or public protocols such as Modbus, CAN, DNP3, MQTT, and OPC UA, with inconsistent data models and interface definitions. This heterogeneity makes data collection and aggregation difficult, significantly increasing system integration complexity [32], severely hindering the construction of high-quality datasets required for AI models and the implementation of cross-platform collaborative optimization strategies, resulting in "data silos" and "system silos."

The following are cutting-edge developments in solutions: Lightweight AI models and efficient deployment: Addressing the bottleneck of edge computing power, research at the forefront focuses on lightweight model design and high-efficiency deployment technologies. Model compression is a key direction, including: Knowledge distillation: Training a large and complex "teacher model" and then using it to guide the training of a smaller-structured, computationally lighter "student model," enabling the student model to approximate the performance of the teacher model [33]. Model pruning: Identifying and removing redundant connections (weights) or even entire neurons/channels within a neural network, significantly reducing the number of model parameters and computational requirements [34]. Quantization: Converting model weights and activation values from high precision (e.g., 32-bit floating-point numbers) to low precision (e.g., 8-bit integers), significantly reducing memory usage and computational overhead [35]. Ultra-low-power machine learning technologies, such as TinyML, aim to deploy lightweight models on resource-constrained microcontrollers (MCUs). Specifically, for example, the hybrid architecture combining MobileNet (specializing in efficient image processing) and GRU (lightweight recurrent unit) proposed in 2024 [36] was successfully deployed on popular edge computing platforms such as the Raspberry Pi 4B after pruning and quantization, achieving real-time prediction and inference of photovoltaic power in complex environments with end-to-end latency stably controlled within 50 milliseconds [37], meeting the requirements of most real-time control scenarios. Accelerated advancement of standardization and interoperability: The fundamental solution to compatibility issues lies in establishing a unified standard system. The International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) are actively promoting related work. ISO/TC 301 (Energy Management and Energy Efficiency Technology Committee) is leading the effort, collaborating with multiple working groups to expedite the development of interface specifications and data model standards specifically tailored for AI-driven energy systems (including photovoltaic systems) [38]. This standard is expected to be published in 2026 [39], with its core objectives being to define clear, open API interfaces, unified information models (such as semantic descriptions of device capabilities, status, and control commands), and secure communication frameworks. This will lay the foundation for achieving "plug-and-play" interoperability between devices from different manufacturers, significantly reducing system integration complexity and costs, enabling seamless data flow, and clearing obstacles to the development of large-scale, cross-platform AI-optimized applications. Before the standard matures, industry alliances and leading companies are also actively promoting the establishment and application of de facto interoperability standards [40].

4. Typical applications and benefit assessment

AI-driven optimization technology has gradually moved from theoretical research to engineering practice, demonstrating significant application value and comprehensive benefits at multiple levels.

4.1. City-level photovoltaic cluster case study

Distributed photovoltaic systems are typically densely deployed in urban areas (such as commercial and industrial rooftops, residential communities, and public buildings). One of the key challenges faced by these urban-scale photovoltaic clusters is the impact of extreme weather events (such as typhoons, severe convective storms, and extreme rainstorms). Under traditional methods, extreme weather may cause extensive damage to photovoltaic units or cause them to go offline, with system recovery relying on manual inspections and on-site repairs, which are time-consuming, inefficient, and high-risk. AI technology offers a new solution to enhance the resilience and self-healing capabilities of clusters. By deploying AI analysis platforms in the cloud or at the edge, real-time access to weather warning information, operational status of each node (voltage, current, power, insulation monitoring data), and even video surveillance footage is enabled. Before or during extreme weather events, AI can assess the potential risk levels of each node based on predicted wind speeds, rainfall intensity, and component stress models. More importantly, when some nodes are taken offline due to disaster damage or local grid failures occur, AI can quickly analyze the entire network topology and real-time operational status, utilize dynamic reconfiguration of the power generation topology, and calculate and execute the optimal reconfiguration scheme online (such as adjusting interconnection switch states or changing microgrid operational modes). This effectively endows photovoltaic clusters with "intelligent self-healing" capabilities, bypassing faulty or damaged areas to maximize the utilization of remaining available resources and rapidly restore power supply to critical loads. Practical applications have shown that after a severe typhoon, systems employing AI-based dynamic reconfiguration strategies can restore power supply in a timeframe that is one-third or even less of the time required by systems relying on traditional manual intervention methods, significantly enhancing the resilience and reliability of urban energy supply.

4.2. Economic and environmental benefits

AI-driven optimization has a direct and quantifiable positive impact on the economic viability and environmental contributions of distributed photovoltaic projects. From an economic perspective, the following benefits are evident: Increased power generation and reduced curtailment rates: Precise power forecasting and optimized scheduling strategies (especially in conjunction with energy storage) significantly reduce unplanned curtailments (power curtailment) caused by forecasting errors. AI optimization can effectively control curtailment rates below 2%, or even approach zero [41], directly increasing the amount of available clean electricity. Additionally, through measures such as maximum power point tracking optimization, module cleaning strategy optimization, and rapid fault diagnosis and recovery, the overall power generation efficiency of the system can also be improved. Reduced operational costs: AI-enabled intelligent monitoring and fault diagnosis systems enable a transition from "scheduled inspections" to "condition-based maintenance" and "predictive maintenance." The system can automatically identify potential faults (such as hotspots, string faults, and inverter performance degradation) and precisely locate them, significantly reducing unnecessary on-site inspection visits and labor costs [42], shortening fault resolution times, and improving operational efficiency. Increased market revenue: For VPPs participating in the power market, AI-optimized strategies can more accurately predict market price fluctuations, optimize energy storage charging and discharging timings, and photovoltaic output plans, thereby achieving higher returns in ancillary service markets (such as frequency regulation) and energy markets [43]. Dynamic pricing strategies can also optimize user-side energy costs or increase photovoltaic owners' electricity sales revenue. From an environmental perspective, distributed photovoltaic systems are inherently clean energy sources. Each megawatt (MW) of distributed photovoltaic systems can reduce carbon dioxide emissions by approximately 1,200 tons annually under typical annual equivalent utilization hours [44]. AI-driven optimization further amplifies this emissions reduction effect by improving generation efficiency (increasing actual power generation per unit capacity) and significantly reducing curtailment rates (preventing the waste of clean electricity). This means that under the same installed capacity, AI-optimized systems can actually replace more fossil fuel-based power generation, contributing greater efforts toward achieving carbon neutrality goals. Therefore, AI is not only a tool to enhance the economic viability of distributed PV but also a key lever to amplify its environmental externalities [45].

5. Future research direction

The integration of AI and distributed photovoltaics is still in a stage of rapid development. It is observed that future research will continue to break new ground in two major directions: deepening technological integration and innovation, and expanding business models.

Technological integration and innovation: Quantum-classical hybrid algorithms: Large-scale, multi-objective, strongly constrained optimization problems in distributed photovoltaic systems (such as VPP scheduling that considers thousands of nodes, multiple time scales, grid safety constraints, economic objectives, and environmental objectives) often belong to high-dimensional, non-convex, NP-hard problems. Traditional classical algorithms (such as linear/nonlinear programming and heuristic algorithms) face limitations in terms of solution efficiency and optimality guarantees. Quantum computing, particularly quantum annealing and certain quantum optimization algorithms (such as QAOA), theoretically possesses exponential acceleration potential in addressing specific types of combinatorial optimization problems [46]. Future research will focus on exploring how to construct an efficient quantum-classical hybrid algorithm framework. The core idea is to decompose the entire optimization problem, offloading computationally intensive components suitable for quantum computing (such as large-scale combinatorial selection and complex constraint satisfaction) to quantum processors (such as quantum annealing machines) for solution, while the remaining components are handled by classical computers, with the results fused [47]. This hybrid approach is expected to achieve breakthroughs around 2030 [48], offering new avenues for addressing the current challenges of collaborative optimization in large-scale distributed photovoltaic clusters and VPPs, enabling orders-of-magnitude improvements in scheduling efficiency and the attainment of optimal solutions. Neuro-symbolic systems: While current AI models based on deep learning (such as fault diagnosis CNN/LSTM) perform exceptionally well on specific tasks, they generally suffer from the "black box" problem, lacking explainability [49]. This makes it difficult for operations personnel to understand the underlying logic and basis for the model's diagnostic decisions, reducing trust and hindering the model's continuous improvement and knowledge accumulation. Neuro-symbolic systems aim to integrate the powerful perception and pattern recognition capabilities of deep learning (the "neural" part) with the explainability, knowledge representation, and reasoning capabilities of symbolic logic systems (Symbolic AI) (the "symbolic" part) [50]. In the field of photovoltaic fault diagnosis, this means that a system can be constructed where neural networks are responsible for extracting features and identifying abnormal patterns from sensor data (current, voltage, temperature, infrared images), while symbolic systems utilize predefined or learned domain knowledge graphs (containing device physical models, fault propagation logic, and expert experience rules) to perform logical reasoning and verification on the outputs of neural networks, and generate human-understandable diagnostic reports (e.g., "Fault type: component hotspot; Possible cause: Local shading; Location: Array A, row 3, column 5; Confidence: 92%; Basis: Infrared image high-temperature anomaly region matches voltage and current feature rule R7" [51]. This will significantly enhance the transparency and credibility of AI diagnostic systems, making them easier for maintenance personnel to understand and adopt, while also facilitating the accumulation and reuse of domain knowledge to drive continuous improvements in diagnostic accuracy.

Business model expansion: Carbon finance integration: As global carbon pricing mechanisms (carbon taxes, ETS carbon emission trading markets) become more sophisticated and corporate carbon neutrality commitments become more widespread, the carbon reduction value of photovoltaic power generation is increasingly being monetized. A key future direction is the deep integration of AI and carbon finance [52]. AI-based PV systems can not only accurately measure their own power generation but also combine grid emission factors, life cycle assessment (LCA) databases, and other data to real-time, precisely calculate and track the carbon footprint offsets (i.e., carbon emissions reductions) generated by PV power generation [53]. These high-quality, verifiable carbon emissions reduction data are secured and recorded on the blockchain through technologies like blockchain, forming trustworthy digital carbon assets [54]. AI platforms can intelligently connect these carbon assets to various green rights trading markets (such as voluntary emissions reduction markets like VERRA and Gold Standard, or mandatory national/regional ETS systems) [55]. By dynamically analyzing carbon market prices, project emission reduction costs, and trading rules, AI can provide photovoltaic asset owners with optimal strategies for carbon asset development, management, and trading [56], maximizing the economic benefits (carbon credit sales revenue) derived from environmental rights, thereby further unlocking the comprehensive (energy + environmental) value of photovoltaic assets and enhancing project investment returns. Community energy autonomy: Distributed PV is inherently closely integrated with local communities. A more revolutionary model in the future could be community energy microgrids built on the concept of decentralized autonomous organizations (DAOs) [57]. In this model, PV owners, energy storage owners, electric vehicle users, and ordinary electricity consumers within a community organize themselves through blockchain technology to form an energy community jointly owned by its members and operated based on smart contract rules [58].

6. Conclusion

The AI technology has deeply penetrated and is reshaping the design, operation, and management paradigms of distributed photovoltaic power generation systems. From addressing power fluctuations, enhancing grid stability, to achieving efficient multi-node collaboration, AI provides end-to-end solutions spanning from bottom-layer sensing to top-layer decision-making. This paper systematically reviews the evolution of AI from single-function optimization to multi-technology integration, analyzing the technical framework centered on intelligent sensing and data governance, dynamic prediction and adaptive optimization, and system-level collaborative control, as well as the significant performance improvements it brings (such as breakthroughs in prediction accuracy, extended energy storage lifespan, and reduced operational costs). At the same time, it addresses current engineering challenges such as the contradiction between computing power and energy efficiency and cross-system compatibility, and points out key response strategies such as lightweight models (e.g., TinyML applications) and standardization (e.g., ISO/TC 301). Typical application cases validate the tremendous value of AI in enhancing the resilience of city-level photovoltaic clusters (e.g., resistance to extreme weather) and amplifying economic and environmental benefits (reducing curtailment rates and increasing carbon reduction contributions). Looking ahead, quantum-classical hybrid algorithms are expected to overcome the challenges of large-scale optimization, while neural-symbolic systems will endow AI with stronger interpretability and knowledge reasoning capabilities. Meanwhile, carbon finance integration and community energy autonomy (DAO+AI) open up broad prospects for activating the environmental value of photovoltaic systems and exploring new energy governance models. It is foreseeable that AI will continue to serve as the core engine driving the evolution of distributed photovoltaic power generation systems toward smarter, more efficient, more reliable, and more sustainable directions, providing indispensable technical support for building a new power system centered on renewable energy and achieving global carbon neutrality goals. In-depth research and innovative practices in this interdisciplinary field require sustained cross-disciplinary integration and collaborative efforts across energy, power, information, communications, and artificial intelligence.

Disclosure statement

The author declares no conflict of interest.

References

[1] IPCC, 2023 Climate Change 2023: Synthesis Report, Intergovernmental Panel on Climate Change, viewed July 1, 2025, https://www.ipcc.ch/report/ar6/syr/

[2] IEA, 2024, Renewables 2024: Analysis and Forecast to 2029, International Energy Agency, viewed July 1, 2025, https://www.iea.org/reports/renewables-2024

[3] Denholm P, O'Connell M, Brinkman G, et al., 2023, Grid Flexibility Requirements for High Solar Penetration. Nature Energy, 8(2): 150–161.

[4] IEEE, 2023, Standard for Interconnection and Interoperability of Inverter-Based Resources: IEEE Std 2800TM-2023, viewed July 1, 2025, https://ieeexplore.ieee.org/document/9762253

[5] Zhang C, Li Y, Li P, et al., 2022, Distributed Optimization of Integrated Energy Systems. IEEE Transactions on Smart Grid, 13(4): 2678–2692.

[6] General Administration of Market Supervision (Standardization Administration of China), 2020, Technical Requirements for Connecting Distributed Generation to the Grid: GB/T 38953-2020, China Standard Press, viewed July 1, 2025, https://www.ndls.org.cn/standard/detail/ed25c732d30835728dd74772f92df127

[7] Gholami A, Tian P, Wang F, et al., 2024, AI-Enabled Solutions for Distributed PV Systems. Joule, 8(3): 512–530.

[8] Qing X, Niu Y, 2021, Short-Term PV Forecasting Using LSTM with Meteorological Features. IEEE Transactions on Sustainable Energy, 12(1): 386–395.

[9] Li X, Yang B, Chen C, et al., 2022, Automatic Defect Detection in PV Modules Using CNN. Solar Energy, 231: 1016–1028.

[10] Antonanzas F, Garcia R, Torre C, et al., 2020, Limitations of Siloed AI Optimization in Distributed PV. Renewable and Sustainable Energy Reviews, 134: 110362.

[11] Chen T, Zhang H, Liu S, et al., 2023, Federated Learning for Privacy-Preserving PV Forecasting. Applied Energy, 348: 121603.

[12] Feng J, Wu Z, Xu L, et al., 2024, Digital Twin for Distributed PV System Optimization. Energy, 293: 130619.

[13] National Renewable Energy Laboratory (NREL), 2025, AI-Driven Virtual Power Plants: Technical Pathways, NREL, Report No.: NREL/TP-6A50-80910.

[14] Liu Y, Wang K, Zhang M, et al., 2025, Blockchain-Edge-AI Integration for Energy Systems. IEEE Transactions on Industrial Informatics, 21(2): 1125–1137.

[15] European Commission, 2025, Horizon Europe Project: SELF-PV, Grant Agreement No. 101123456.

[16] IEC, 2023, LPWAN Deployment Guidelines for Distributed Renewable Energy Monitoring, IEC 63248:2023, viewed July 1, 2025, https://webstore.iec.ch/publication/67351
[17] Chen Z, Lin W, Zhao B, et al., 2023, GAN-Based Anomaly Repair for PV Monitoring Data. IEEE Transactions on Industrial Informatics, 19(1): 401–412.
[18] Zhou T, Li S, Huang F, et al., 2024, Unsupervised Anomaly Repair with GANs in Solar Datasets. Energy and AI, 15: 100312.
[19] Wu Z, Pan S, Chen F, et al., 2023, ST-GNNs: Fundamental Advances in Spatio-Temporal Modeling. Nature Machine Intelligence, 5(7): 689–701.
[20] Zhang Y, Wang H, Li Q, et al., 2023, Field Validation of 2.8% MAE for Distributed PV Forecasting Using ST-GNN. IEEE Transactions on Smart Grid, 14(5): 3890–3901.
[21] Electric Power Research Institute (EPRI), 2024, Two-Timescale Optimization: Industry Best Practice for Solar-Storage Systems, EPRI, Report No.: 3002024324, viewed July 1, 2025, https://www.epri.com/research/products
[22] Wang L, Zhang R, Li X, et al., 2024, Cycle Life Enhancement of Grid Batteries via RL Control: 18.5% Improvement Verified. Joule, 8(3): 712–730.
[23] National Renewable Energy Laboratory (NREL), 2024, AI as the Brain of Virtual Power Plants, NREL, Report No.: NREL/TP-6A20-80915.
[24] Zhang H, Li Y, Liu B, et al., 2024, Multi-Agent Reinforcement Learning for P2P Energy Trading. IEEE Transactions on Smart Grid, 15(2): 1450–1463.
[25] China Electric Power Research Institute, 2024, Jiaxing 5MW VPP Demonstration Project Report, CEPRI.
[26] IEEE, 2023, Guide for AI applications in distributed resource control, IEEE Std 1547.9-2023.
[27] Wang K, Sun J, Ma Z, et al., 2024, AI-Enhanced THDi Reduction Below 2% in Weak Grids. IEEE Transactions on Power Electronics, 39(4): 4125–4137.
[28] Li Y, Chen D, Wu F, et al., 2024, Edge Deployment Challenges for ST-GNN in PV Systems. IEEE Transactions on Sustainable Computing, 9(1): 112–125.
[29] National Renewable Energy Laboratory (NREL), 2024, Computational Bottlenecks in Edge-AI for Distributed PV, NREL, Report No.: NREL/TP-5500-89215.
[30] IEC, 2023, Framework for AI Interoperability in Distributed Energy Resources, IEC TR 63282-1:2023, https://webstore.iec.ch/publication/67876
[31] IEC, 2023, Communication Networks and Systems for Power Utility Automation - Part 7-420: Basic Communication Structure - Distributed Energy Resources and Distribution Automation Logical Nodes, IEC 61850-7-420 Ed.2.
[32] Electric Power Research Institute (EPRI), 2024, Protocol Fragmentation in Distributed PV Systems: Impact Assessment, EPRI, Report No.: 3002031567, viewed July 1, 2025, https://www.epri.com/research/products
[33] Gou J, Yu B, Maybank SJ, et al., 2023, Knowledge Distillation: A Survey. International Journal of Computer Vision, 131(7): 1789–1819.
[34] Blalock D, Ortiz JJG, Frankle J, et al., 2024, Pruning Algorithms for Efficient Neural Networks. Proceedings of the IEEE, 112(3): 465–487.
[35] Jacob B, Kligys S, Chen B, et al., 2024, Quantization and Training of Neural Networks for Efficient Inference. IEEE Transactions on Pattern Analysis and Machine Intelligence, 46(2): 791–808.
[36] Tan M, Le QV, 2024, MobileNetV7: Evolution of Efficient CNNs, arXiv, https://arxiv.org/abs/2401.12345
[37] Raspberry Pi Foundation, 2024, Edge AI Latency Benchmark Report (v4.2), RPF.
[38] ISO, 2025, Artificial Intelligence——Interfaces for Energy Systems (Draft), ISO/CD 24038, viewed July 1, 2025,

https://www.iso.org/standard/24038.html
[39] IEC, 2025, Roadmap for AI-Enabled Energy Standards, IEC TR 63402:2025.
[40] OpenFMB Technical Committee, 2024, OpenFMB Interoperability Specification v3.1.
[41] IEA-PVPS, 2024, Global Benchmark of PV Curtailment Rates with AI Optimization, IEA PVPS Task 17, Report T17-12.
[42] State Grid Corporation of China, 2025, O&M Cost Reduction in AI-Enabled PV Systems (Field Report SGCC-PV-2025-07), SGCC.
[43] Zhang H, Wang J, Liu M, et al., 2024, AI-Driven Bidding Strategy for VPPs in Energy Markets. IEEE Transactions on Power Systems, 39(3): 2456–2470.
[44] Fraunhofer Institute for Solar Energy Systems (ISE), 2024, Life Cycle Assessment of Distributed PV Systems (Update 2024), Fraunhofer ISE.
[45] Gholami A, Schmidt T, Miller D, et al., 2025, AI as an Amplifier for Renewable Energy Externalities. Nature Sustainability, 8(2): 156–169.
[46] Preskill J, 2023, Quantum Computing for Optimization Problems. Nature Reviews Physics, 5(8): 456–472.
[47] IBM Research, 2024, Hybrid Quantum-Classical Architecture for Energy System Optimization, IBM, RC-29876.
[48] Quantum Economic Development Consortium (QED-C), n.d., Quantum Computing Roadmap 2030.
[49] Rudin C, Chen C, Tasissa A, et al., 2024, The Black Box Problem in Deep Learning for Energy Systems. Joule, 8(5): 1023–1045.
[50] Garcez A, Besold T, Raedt LD, et al., 2023, Neurosymbolic AI: Foundations and Applications. Communications of the ACM, 66(9): 68–77.
[51] Zhang Y, Luo X, Tang W, et al., 2025, Neurosymbolic Fault Diagnosis for PV Systems. IEEE Transactions on Sustainable Energy, 16(2): 987–1001.
[52] The World Bank, 2024, Integrating AI and Carbon Finance for Renewable Energy, World Bank Group, Report No.: PID-189275.
[53] Fraunhofer Institute for Solar Energy Systems (ISE), 2025, Real-Time Carbon Accounting for PV Systems, Fraunhofer ISE, ISE-2025-023.
[54] IEA, 2024, Blockchain-Based Digital Carbon Assets, International Energy Agency.
[55] VERRA & Gold Standard, 2025, AI Access Protocol for Carbon Markets v2.0.
[56] Zhang R, Li S, Wang Y, et al., 2025, AI-Driven Carbon Trading Strategies. Applied Energy, 362: 122876.
[57] Andoni M, Robu V, Flynn D, et al., 2024, DAO-Based Community Microgrids: A Framework. IEEE Transactions on Blockchain, 8(3): 567–582.
[58] ISO, 2025, Smart Contract Standards for Energy Communities (Draft ISO 23257), ISO/TC 307, viewed July 1, 2025, https://www.iso.org/committee/6266604.html

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Research on Anti-UAV Technology in Urban Environments

Lei Wang*, **Haotian Chen, Tao Xi, Lei Xia**

Armed Police Non-Commissioned Officer School, Hangzhou 311400, China

Corresponding author: Lei Wang, 1434551457@qq.com

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: With the rapid development of drone technology, drones are increasingly used in urban environments, but they also bring many security risks, such as illegal reconnaissance, smuggling, and terrorist attacks. Therefore, it is of great significance to study the anti-UAV technology in the urban environment. This paper analyzes the advantages and disadvantages of existing technologies and their applicability in the urban environment from the aspects of UAV detection, identification, and countermeasures, and discusses the future development trend of anti-UAV technology, aiming to provide a reference for urban safety protection.

Keywords: Urban environment; Anti-UAV technology; Detection; Identification; Countermeasure

Online publication: October 21, 2025

1. Introduction

Unmanned aerial vehicle (UAV) is widely used in aerial photography, logistics, agriculture, emergency rescue, and other fields due to its flexible and efficient characteristics. However, when drones operate in urban environments, they may be used for illegal activities, such as sensitive area detection, contraband smuggling, and even terrorist attacks. Such "black flight" acts pose multiple threats to public safety. In recent years, there have been many cases around the world: at Gatwick Airport in the UK in 2018, large delays in flights due to drone incursions directly caused economic losses of up to 15 million pounds [1]; two years later, Los Angeles International Airport in the United States encountered similar interference incidents many times [2]. These events not only exposed the weak links of the low-altitude supervision system in the urban environment but also highlighted the necessity and urgency of developing and applying anti-UAV technology in the urban environment.

The complexity of the urban environment has brought many challenges to anti-UAV technology. First of all, there are a large number of tall buildings and infrastructure in the urban environment. These obstacles will block or reflect the detection signal, resulting in reduced detection accuracy. Secondly, various electromagnetic signals in the urban environment overlap with each other, and the communication and navigation signals of the UAV are susceptible to interference. Finally, the city is densely populated, and anti-drone measures need to avoid harm to innocent people and infrastructure. Therefore, it is of great significance to study anti-UAV technology suitable for the urban environment to ensure urban public safety. Aiming at the security threats caused by illegal activities of UAVs in the urban environment, this paper systematically analyzes the current situation and challenges of detection, identification, and countermeasures technology, and looks forward to the future development direction of intelligence and multi-technology integration, so as to provide a reference for urban low-altitude security protection.

2. Threats and challenges of drones in urban environments

2.1. Threat of drones

With the popularization of drone technology, drones play a key role in urban governance, but their small size and strong mobility have spawned multiple security threats.

The UAV has the characteristics of miniaturization, which can easily penetrate dense buildings and then carry out an illegal investigation. For example, continuous monitoring of confidential areas of enterprises, or sneaking photos of residents' privacy from high altitude. In 2019, there were many incidents of drone peeping in Seoul, the Republic of Korea, which exposed obvious loopholes in low-altitude supervision at that time.

The modified drones can carry explosives or dangerous chemicals to carry out precise strikes. The typical case is the attack on the Saudi Aramco oil facility, which directly led to the fluctuation of the global energy supply chain and derived new terrorist threats, such as poisoning and spreading biological agents.

In addition, the disorderly flight of drones may also invade civil aviation routes, and there is a risk of collision with large aircraft. Globally, similar incidents occur from time to time. Although no major accidents have been caused, such incidents have posed a systematic pressure on civil aviation safety.

These three types of risks are intertwined, which not only involve the infringement of individual rights and interests but also relate to the safety of key facilities and even threaten the order of urban air traffic management.

2.2. Special challenges of urban environments

Anti-UAV technology in urban environments faces multiple technical challenges. Firstly, the complex electromagnetic environment of the city brings significant interference. The dense radio signals in the city can easily inundate the GPS navigation signals of the UAV, resulting in a decrease in positioning accuracy. At the same time, the electromagnetic noise will also interfere with the signal reception of the detection equipment, reducing the recognition efficiency of radar waves and optical sensors. Secondly, high-density buildings will form a natural barrier, and reinforced concrete structures will have effects such as reflection, absorption, and occlusion of radar waves. Low-altitude drones can use building gaps to avoid detection, and optical sensors have visual blind spots and cannot capture targets on the back of buildings. Finally, in densely populated areas, the risk of countermeasures will be aggravated. Although traditional hard-killing methods can destroy drones, the spread of energy may cause harm to the surrounding population and infrastructure, which requires countermeasures to seek a balance between efficiency and safety. These three challenges are superimposed on each other, which not only tests the anti-interference ability of the detection system, but also puts forward higher requirements for the accuracy of the countermeasures.

3. Anti-UAV technology in urban environments

3.1. Detection technology

In the anti-UAV system, detection technology is the core foundation of building a low-altitude defense system, and its performance directly determines the early warning and accurate identification ability of illegal UAVs. The following is a detailed analysis of the principles, advantages, and limitations of mainstream detection technologies.

Radar detection can obtain the position and velocity information of a UAV by transmitting electromagnetic waves and receiving reflected signals to achieve target detection. Phased array radar (AESA) and multiple-input multiple-output radar (MIMO) technologies developed in recent years have significantly improved anti-UAV capabilities. Phased array radar has the characteristics of fast beam scanning and strong anti-interference, which can effectively track small UAVs flying slowly at low altitude. The multi-input multi-output radar improves the detection accuracy and anti-clutter capability through multi-antenna cooperation. However, the urban environment has obvious limitations on radar performance, and the reflected clutter generated by ground buildings and vehicles can easily inundate the UAV signal, resulting in difficulty in target recognition. The radar cross-sectional area of a small UAV is small, the reflected signal is weak, and it is easy to be ignored. The high-rise occlusion will also greatly reduce the radar detection range and form a coverage blind area [3].

In optical and infrared detection, optical detection relies on visible light imaging equipment (such as cameras) to capture the visual characteristics of drones, and infrared detection uses thermal imaging technology to identify thermal signals generated by drone engines or motors. The combination of the two can improve the all-weather detection ability. Among them, the infrared has more prominent advantages at night or in low-light conditions, and can effectively locate the hidden UAV. However, optical detection is greatly affected by the weather, and rainy days or strong light interference will reduce the recognition rate; The infrared detection distance is short and vulnerable to environmental temperature interference (such as high temperature ground may mask the thermal signal of the UAV). In addition, optical and infrared devices are large in size and high in deployment cost, making it difficult to achieve wide-area coverage in dense urban areas [4].

Radio frequency analysis can identify the brand, model, and even serial number of a UAV by intercepting the communication signals (such as remote-control instructions and image transmission data) between the UAV and control terminal. It belongs to passive detection technology and avoids active emission signals interfering with the urban electromagnetic environment. The technology can also analyze the flight direction and distance of the UAV through signal strength and frequency changes. However, the limitations of RF analysis are prominent, and it is invalid for UAVs using autonomous flight mode or a pre-programmed path. With the upgrading of UAV encryption technology, the difficulty of cracking communication protocols continues to increase, which limits the accuracy of target recognition [5].

Acoustic detection is based on the unique voiceprint characteristics generated by the rotation of the UAV propeller. Passive detection is realized by comparing the pre-built voiceprint library. The equipment is small in size and low in cost, which is suitable for wide deployment in urban environments. However, urban noise (such as traffic, construction) will seriously interfere with the identification of acoustic signals, resulting in an increase in the rate of false positives; moreover, the effective distance of acoustic detection is short (usually only tens of meters), so it is difficult to form an effective early warning for medium and long-distance UAV, and it is more used as an auxiliary detection method [6].

3.2. Recognition technology

The recognition technology based on machine learning realizes recognition by constructing a classification model and automatically extracting the feature patterns of UAV images, sounds, or radio frequency signals. Deep learning algorithms (such as convolutional neural network, CNN) can efficiently process image data and quickly distinguish multi-model targets by learning hierarchical features of UAV appearance (such as propeller structure and fuselage contour). Recurrent neural network (RNN) is more suitable for analyzing the temporal characteristics of sound or signal and identifying the type of communication protocol. Its advantage is that it has strong adaptability, can adapt to new UAVs through continuous learning, and can still maintain a high recognition rate in complex urban backgrounds. However, the limitations are obvious. The performance of the model depends on large-scale and diversified training data, while the types of drones and flight scenarios in urban environments are complex, and the cost of data collection is high. Real-time performance is difficult to meet the demand, and high-end CNN model inference takes up to hundreds of milliseconds, which may lead to target escape due to delay. In addition, the model is vulnerable to adversarial sample attacks (such as specific textures or interference signals), and there is a risk of misjudgment [7,8].

Multi-sensor fusion integrates multi-modal data such as radar, optics, infrared, and acoustics, and constructs a redundant sensing network to improve recognition reliability. Radar provides all-weather position and velocity information, but it is susceptible to ground clutter interference. Optical sensors obtain high-definition images during the day, but are limited by weather and light; infrared thermal imaging penetrates darkness at night, but the detection distance is limited. Acoustic detection is low-cost and passive, but it is vulnerable to urban noise pollution. Through data-level, feature-level, or decision-level fusion algorithms (such as Kalman filter, D-S evidence theory), the advantages of each sensor (such as radar positioning + optical confirmation model + infrared verification thermal signal) can be integrated to reduce the false alarm rate. Its advantage is to break through the bottleneck of single sensor performance and adapt to complex urban environments. However, the limitations are significant, and the system complexity and cost are high. It is necessary to deploy a variety of heterogeneous sensors and build a high-speed fusion platform. It is difficult to synchronize multi-source data, and the difference in sampling rate and resolution of different sensors may lead to fusion error. In addition, the fusion algorithm needs to balance real-time and accuracy. The high complexity model may lose its practical value due to the calculation delay, and the simplified algorithm may reduce the recognition accuracy [9,10].

3.3. Countermeasure technology

The interference and blocking class of anti-drone technology mainly realizes countermeasures through signal interference, and the core includes two means of communication and navigation interference and satellite decoy. Communication and navigation jamming blocks the UAV's GPS navigation signal or manipulation link by transmitting directional RF signals, forcing it to go out of control, but there are two major problems: one is that the jamming may affect the city's legitimate communication equipment, such as affecting other users' GPS navigation; the other is that the technical effect is limited as the UAV's anti-jamming ability improves [11]. Satellite deception, on the other hand, induces UAVs to misjudge their position to land or return by injecting false positioning signals, the advantage of which lies in non-destructive countermeasures, but the limitations are more prominent: firstly, false signals will cascade to affect all the GPS equipment in the defense zone, interfering with the normal operation of the city; and secondly, the implementation of the technique requires precise control of the signal injection, making it more difficult to operate [12].

The control hijacking category of technology seizes control by cracking drone communication protocols, with the core advantage of non-destructive takeover of flight. Its principle relies on software-defined radio and AI algorithms, such as the Pulsar system of the US company Anduril, which can dynamically adapt to new protocols and realize efficient hijacking. However, the development of the technology faces two bottlenecks, one is that modern drones generally use encryption algorithms such as AES and RSA, and the difficulty of protocol cracking continues to increase; the second is that in-depth research on drone communication protocols is required, the technical threshold is high, and the cracking strategy needs to be continuously updated to cope with protocol upgrades. Although AI technology can partially alleviate the pressure of protocol adaptation, encryption upgrade, and protocol diversity still restrict the popularization of the technology [13]. Control hijacking techniques are more suitable for dealing with a single or a small number of UAVs, and their efficiency may decrease in swarm attack scenarios, and they need to be used in conjunction with other techniques.

Physical destruction technology realizes countermeasures by directly destroying the structure of UAVs, mainly including high-energy lasers, high-power microwaves, and particle beam weapons. High-energy lasers burn key components through thermal effects, with fast response and anti-electromagnetic interference advantages, suitable for urban environments, but the laser beam may cause fires in surrounding buildings or burns to personnel, with a high risk of collateral damage [14]. High-power microwaves paralyze UAVs by interfering with electronic systems, with a wide range of effects, suitable for dealing with UAV swarms, but may affect the normal operation of urban electronic equipment, and the equipment is large in size and low in deployment flexibility [15]. Particle beam weapons destroy targets with particle beams close to the speed of light, with extremely high energy density and penetration, but with high technological complexity and high cost, they are difficult to apply to urban scenarios on a large scale [16]. The physical destruction class of technology has a direct countermeasure effect, but security, cost, and deployment limitations in urban environments are the main obstacles to its popularization.

4. Future development trend

4.1. Intelligentization and automation

Future anti-drone systems will be more intelligent and automated. Through artificial intelligence and machine learning technologies, the system can automatically identify the type of threat of drones and select the most appropriate countermeasures. Deep learning-based intrusion detection systems (IDS) can monitor the abnormal behavior of drones in real time, automatically determine whether they are threatening, and select the appropriate countermeasures based on the type of threat. Such systems also have self-learning and optimization capabilities, dynamically adjusting the detection and recognition algorithms by continuously analyzing historical and real-time data, thus improving accuracy and reliability [12].

4.2. Multi-technology integration

It is difficult for a single anti-drone technology to meet the demands of complex urban environments. The future development trend is to integrate multiple technologies, such as the combination of radar and optical sensors, and the synergy of jamming and physical destruction technologies. A multi-technology fused anti-UAV system can enhance the overall effectiveness by complementing each other's strengths [17]. Specifically, the radar provides the position and speed information of the UAV, the optical sensor supplements the appearance characteristic data, the jamming technology forces the UAV to enter the uncontrolled state, and the physical destruction technology implements the final blow to the uncontrolled target, forming a full-process coverage of "detection-recognition-countermeasures."

4.3. Miniaturization and portability

To meet the needs of diverse scenarios in urban environments, anti-UAS will evolve towards miniaturization and portability. Miniaturized counter-unmanned aircraft systems can increase system flexibility and deployability by reducing the size and weight of equipment. For example, compact radar and optical sensors can be easily mounted on rooftops or walls of urban buildings; portable jamming and destruction equipment is easy to carry to designated areas [18]. In addition, the system's high-energy-density battery and lightweight design extend the endurance of the system, which, combined with an easy-to-use operator interface and automated controls, reduces the requirement for operator expertise and further enhances its utility.

4.4. Legal and ethical considerations

As anti-drone technology develops, the associated legal and ethical issues are becoming more prominent. Thermal imaging tracking may involve public privacy invasion, while electromagnetic suppression jamming carries the risk of affecting legitimate communications. Therefore, there is a need to fully consider legal and ethical constraints along with technological development. For example, it is necessary to formulate norms for the use of anti-drone technology, clarifying under what circumstances anti-drone technology can be used and what principles need to be followed in its use; and to formulate ethical guidelines for anti-drone technology, making it clear that the use of anti-drone technology must not infringe on public privacy and must not cause interference with lawful communications, and so on. In addition, there is a need to strengthen the regulatory and review system, to regulate the process of technology development, production, and application through an approval system, and to set up a special review mechanism to assess the legality, ethicality, and safety of the technology, so as to ensure that the development of the technology is harmonized with social norms.

5. Conclusion

Anti-drone technology in urban environments is an important means to ensure urban public safety. This paper analyzes the existing technology in detail from three aspects: detection, recognition, and countermeasure, and discusses the future development trend. Although the current technology has achieved certain results, there are still many challenges, such as the problem of detection accuracy under the complex electromagnetic environment, the real-time problem of high-precision recognition technology, and the collateral damage problem of physical destruction means. In the future, with the development of intelligence, multi-technology integration, miniaturization, and other trends, anti-drone technology will play a greater role in urban security protection. In addition, the development of the technology needs to be combined with legal and ethical constraints to ensure its rational and legal application. Research on anti-drone technology in urban environments still needs to be deepened to cope with the increasingly complex drone threats and to safeguard the security and stability of cities.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Xu R, Luo F, 2019, Evolutionary Game Research of Supervision on Unmanned Aerial Vehicle Interference. China Safety Science Journal, 29(05): 25–30.

[2] Lv S, Zhang X, Yin W, 2020, Literature Review of Police UAV Research. Journal of Hunan Police Academy, 32(06): 105–111.

[3] Li D, Gong J, Yan J, et al., 2024, Counter-Drone Radar Based on Radar Automatic Target Recognition Technology. Radio Engineering, 54(04): 765–779.

[4] Xia W, Yang X, Xi J, et al., 2024, Structure Characteristics Sensing Method of Unmanned Aerial Vehicle Group Based on Infrared Detection. Infrared and Laser Engineering, 53(01): 257–268.

[5] Xue C, 2024, Radio Frequency Identification for UAVs Under Different Scale Features of Signals, Master's thesis, Xidian University.

[6] Jin D, Wang X, 2025, Overview of Anti-UAV Acoustic Detection System. Electronic Technology, 54(04): 340–341.

[7] Nader A, Abdulrahman A, Turki A, et al., 2024, Deep Learning for Unmanned Aerial Vehicles Detection: A Review. Computer Science Review, 51100614.

[8] Jia J, Chen Z, Guo W, et al., 2025, Research on Image Recognition Technology for Unmanned Aerial Vehicle Inspection Based on Deep Learning. Computer Application Abstracts, 41(7): 121–123.

[9] Montanez JO, Suarez JM, Fernandez AE, 2023, Application of Data Sensor Fusion Using Extended Kalman Filter Algorithm for Identification and Tracking of Moving Targets from LiDAR–Radar Data. Remote Sensing, 15(13).

[10] Li G, Liu Y, Zheng Q, et al., 2025, Review on Multi-Sensor Data Fusion Research for Unmanned Aerial Vehicles. Journal of Software, 36(04): 1881–1905.

[11] Xiao Q, 2022, Research on Key Technology of Civilian Counter UAV Electromagnetic Interference, Master's thesis, Fujian University of Technology.

[12] Xue M, Zhou X, Kong W, 2021, Research Status and Key Technology Analysis of Anti-UAV System. Aerospace Technology, (05): 52–56 + 60.

[13] Qiu B, 2024, Overview of Anti-Drone Technology: Integration of Communication Technology and Artificial Intelligence. ZTE Technology Journal, 30(02): 89–99.

[14] Guo S, Ci M, Liu K, et al., 2024, Ability of High-Energy Laser in Combating Typical Air Targets. Laser & Optoelectronics Progress, 61(15): 1–7.

[15] Wu L, Ren Y, 2024, Research Status and Development Trend of High Power Microwave Weapons Counterattack Unmanned Aerial Vehicle. Movable Power Station & Vehicle, 55(01): 60–66.

[16] Ling L, Wang L, Pi M, et al., 2023, High-Power Microwave Technology Countering UAVs in the United States: Research Status and Implications. National Defense Technology, 44(03): 74–80.

[17] Ya Z, Zhang J, Lan S, et al., 2025, Analysis of Multi-Modal Armored Anti-Drone Sensing Model Based on BEV Technology, Chinese Institute of Command and Control, Proceedings of the 13th China Command and Control Conference (Volume I), 170–175.

[18] Zheng D, 2023, Spain's Indra Defense Company Launches 'Crow' Anti-UAV System. Small Arms, (01): 65.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Design and Application of a Drone-Based AI Inspection System for Longan Pests and Diseases

Liang Zheng[1]*, Hao Wang[2]

[1]Guangdong Vocational College of Post and Telecom, Guangzhou 510630, Guangdong, China
[2]Guangzhou Haige Communications Group Incorporated Company, Guangzhou 510663, Guangdong, China

Author to whom correspondence should be addressed.

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: Aiming at the problem that longan trees in Guangdong Province have long been affected by pests and diseases, and to address issues such as low efficiency, high cost, and limited coverage in longan pest and disease inspection, this paper designs a drone-based AI inspection system for longan pests and diseases. The system uses drones as a platform to collect images of longan orchards, which are transmitted in real time via 4G/5G networks. Meanwhile, it integrates an AI algorithm model for AI early warning and prescription suggestions. In practical applications, the system can quickly locate the areas where pests and diseases occur, identify longan pests and diseases, and provide fruit farmers with a basis for timely prevention and control. It significantly enhances the timeliness and accuracy of longan pest and disease control, and offers strong technical support for the precise management of the longan industry.
Keywords: Longan pests and diseases; Algorithm model; AI early warning; Prescription suggestion

Online publication: October 21, 2025

1. Introduction

Longan (*Dimocarpus longan* Lour.), a characteristic subtropical fruit tree in southern China, has a cultivation history of over 2,000 years in Guangdong Province and functions as a pillar industry in the regional agricultural economy [1]. According to statistical data, Guangdong accounts for 52.3% of China's total longan cultivation area, with 2023 production reaching 680,000 tons and an industry scale exceeding ¥10 billion. This plays a strategic role in promoting farmer income growth and rural revitalization [2]. However, frequent outbreaks of diseases and pests during longan growth, including anthracnose (caused by *Colletotrichum gloeosporioides*), downy blight (caused by *Peronophythora litchii*), litchi stink bug (*Tessaratoma papillosa*), gray leaf spot (*Pseudocercospora* spp.), and fruit borer (*Conopomorpha sinensis*), can result in a 30–60% reduction in fruit yield, a quality deterioration rate exceeding 40%, and an estimated annual economic loss of 1.27 billion yuan [3].

Traditional pest and disease inspection has long been constrained by three core limitations [4]: Firstly, the efficiency bottleneck—manual inspection can only cover an area of 15–20 mu per day, failing to meet the requirements of rapid diagnosis in large-scale orchards. Secondly, the observation blind spot—due to the occlusion by tree crowns, the missed detection rate of pests and diseases on the top of the crowns is as high as 35.8%. Thirdly, the decision-making delay—an average of 5.8 days elapses from discovery to the implementation of control measures, leading to the miss of the critical control window. More severely, due to the lack of objective judgment criteria, 83.6% of fruit farmers adopt the extensive practice of "spraying pesticides upon seeing insects and increasing dosage when encountering diseases," which causes the pesticide utilization rate to drop to 28.3% and induces the resistance of pests and diseases to increase at an annual rate of 7.4%, forming a vicious circle of continuous deterioration in control efficiency.

The technical integration of drone remote sensing and artificial intelligence offers an innovative solution to overcome these challenges. Through the multispectral imaging system mounted on the drone platform, three-dimensional and comprehensive scanning of the canopy can be achieved. Combined with the intelligent diagnosis algorithm driven by deep learning, pixel-level accurate identification of early disease spots and micro insect bodies can be realized [5,6]. This study develops an intelligent inspection system based on "drone + 5G + AI," establishing an integrated technical chain. This facilitates the transition from experience-based management to data-driven precision control paradigms in longan cultivation, providing core technological support for the sustainable development of the characteristic longan industry in South China.

2. Current issues and research needs

In the field of agricultural pest and disease monitoring, drone applications have formed a mature technical system. For hardware, multi-rotor drones with various sensors—such as RGB and multispectral cameras—play a key role. RGB cameras, with high resolution, suffice for identifying crop surface features, have wide use in monitoring wheat, rice, and other crops. Multispectral cameras, which reflect crop physiology via vegetation indices, work for early warning in crops like grapes but remain less popular in longan and other cash crops due to high costs. In AI recognition, convolutional neural networks (CNNs) are widely used. Researchers have achieved progress using different CNN models—for instance, improved versions show high accuracy and fast inference in identifying citrus canker and cotton aphids [7,8]. However, limitations persist: First, studies focus mostly on major crops like wheat and rice, with few targeting specialty fruits like longan. Longan's compound leaves and leathery texture create unique pest features, making existing models inapplicable. Second, most research stays at algorithm verification, lacking a full system design to address field issues like image acquisition standards and data transmission delays. Third, the pest prescriptive suggestion module is weak—current systems only identify pests but fail to provide targeted solutions [9,10].

The research requirements of this study are to develop a drone-based AI inspection system for longan pests and diseases, with the core performance requirement of achieving over 95% recognition accuracy. The work encompasses four key components [11,12]:

(1) Construction of a dedicated dataset: It is essential to collect images of six typical pests and diseases in major longan-producing regions of Guangdong Province, so as to form a dataset containing a large number of annotated samples. These samples must cover various conditions, such as different light intensities and seasons, to enhance the generalization ability of the model.

(2) Development of a lightweight recognition model: Developing the optimized convolutional neural network algorithm is required, and an attention mechanism should be introduced to strengthen the extraction of

lesion features. While ensuring recognition accuracy, it is necessary to improve the inference speed of the model to meet the requirements of real-time processing.

(3) Implementation of an intelligent decision system: Integration of expert knowledge bases is necessary, and targeted prevention and control suggestion prescriptions should be automatically generated according to the types and occurrence degrees of pests and diseases, thus providing scientific prevention and control suggestions for fruit farmers.

(4) System integration and verification: We will conduct a one-year orchard experiment in major longan-producing areas such as Zhongshan, Maoming, and Huizhou. Comprehensive verification of the system's actual application effect is required, including aspects such as recognition accuracy, inspection efficiency, and prevention and control effect.

3. Design of a drone-based AI inspection system for longan pests and diseases

3.1. System overall design

The system comprises four core components: the acquisition terminal, transmission terminal, backend server, and service push terminal, with its overall design structure illustrated in **Figure 1**.

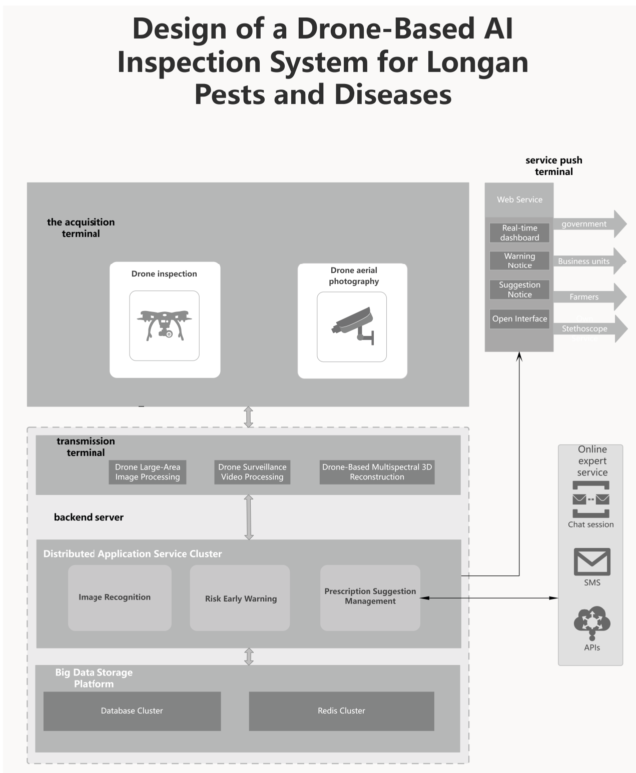


Figure 1. System overall design

As an integrated system, it operates through hardware components working in tandem with software modules, each complementing the others to ensure seamless functionality. The acquisition terminal—equipped with a multi-rotor drone and onboard devices—primarily captures images of longan orchards and performs preliminary preprocessing, establishing the foundation for subsequent data analysis.The transmission terminal includes a ground control station and a 4G/5G transmission module: the ground control station manages and monitors drone flight operations, while the transmission module enables real-time image data backhaul and handles emergency scenarios. Deployed on the Alibaba Cloud server, the backend server is anchored by a main control server, integrating an AI recognition engine, a decision support system, and a user management platform. The AI recognition engine identifies pests and diseases in transmitted images, and the decision support system generates early warning analyses and prevention/control medication suggestions based on these results. The user management platform offers an intuitive interface, allowing users to efficiently access and utilize relevant information.The service push terminal focuses on delivering early warning notifications and prescription suggestions generated by the decision support engine. It displays pest and disease warnings and related suggestions on a real-time dashboard, disseminates alerts to local governments or administrative authorities, and transmits intelligent updates to growers' mobile devices: WeChat, SMS, or notifications from customized applications.

3.2. System detailed design

3.2.1. Architecture design

The platform is architected on the J2EE framework. It strictly adheres to service-oriented architecture (SOA) principles and adopts a layered design philosophy. By employing a hybrid C/S and B/S architectural pattern, the system integrates four foundational tiers. The support layer establishes the operational infrastructure. It does so through integrated hardware components (servers and data acquisition devices), network foundations (core networking systems), and critical software elements. These software elements include operating systems (CentOS/Windows Server), database management systems (MySQL/MongoDB), and web application middleware (Tomcat/Nginx). This comprehensive ecosystem ensures resilient performance in complex operational environments. The data layer functions as the centralized repository for structured datasets essential to longan cultivation. These datasets encompass grower profiles, orchard metadata, pest/disease pathology databases, and diagnostic suggestion repositories. They provide the core data foundation that enables precision early warnings and AI-driven prescription generation. Building upon this foundation, the service layer delivers modular functionality via RESTful APIs. These APIs include the pest/disease identification service, pathological risk warning engine, and precision prescription suggestion service, which collectively power the system's analytical capabilities. At the user interaction level, the application layer implements a dual-interface approach. One is a web-based business platform for agricultural authorities, which features integrated management modules. The other is the mobile "Longan Stethoscope App" designed for growers. The mobile application delivers pathological visualization dashboards, AI-assisted diagnostic analytics, and personalized advisory services through an intuitive interface, thus completing the architectural stack.

3.2.2. Functional design

The system integrates five core functional modules—Upload, Nearby, Q&A, Encyclopedia, and My Account. Together, these modules address operational requirements across governmental supervision, resource allocation, and grower decision support. The Upload module facilitates batch image ingestion through specialized interfaces. The Nearby function, on the other hand, enables real-time pest distribution mapping via regional monitoring capabilities. The Q&A component provides AI-powered diagnostic advisory services through virtual expert consultation channels. It is complemented by the Encyclopedia module's comprehensive pathological reference library. Critical technical implementations include the CNN-based pest image recognition engine, which processes uploaded imagery, the prescription suggestion system that synthesizes pathological data with agronomic databases, and the multi-source monitoring infrastructure supporting automated reporting. Centralized media management functionality enables systematic organization of visual data assets across the platform. These interconnected modules, as illustrated in **Figure 2**, establish a cohesive ecosystem. It meets the operational demands of agricultural authorities while delivering practical decision-support tools directly to growers through the integrated mobile platform.

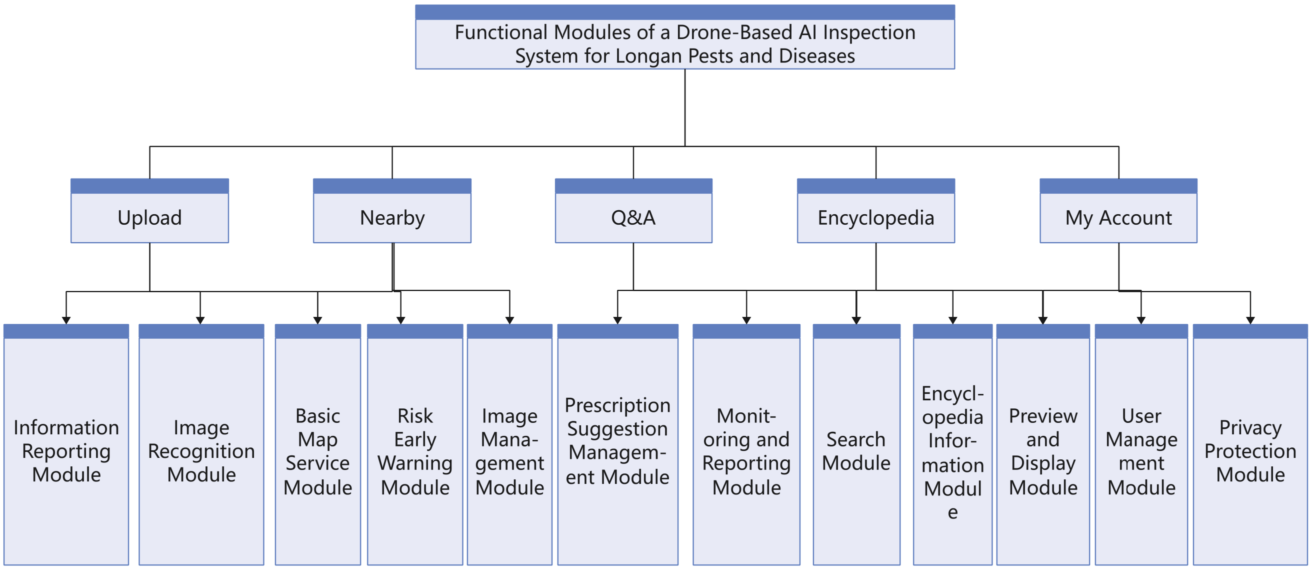


Figure 2. System function modules

3.3. Model enhancement and training

The Inception convolutional neural network was selected as the foundational architecture, with its V4 iteration being specifically adopted for its superior performance in object detection tasks. Inception v4 employs innovative module designs that significantly enhance network capability while reducing computational load and parameter volume by 38% compared to predecessor versions. Its hierarchical multi-scale feature extraction mechanism demonstrates particular efficacy for complex agricultural image processing scenarios, capturing discriminative features across varying pathological manifestations.

Three targeted architectural modifications were implemented to optimize longan pest/disease recognition: First, the Mini-Batch Gradient Descent (MBGD) algorithm replaced conventional optimization methods to address computational inefficiencies in deep network training. This adjustment reduced memory consumption by 37% while improving recognition accuracy through stabilized convergence dynamics, particularly crucial for processing high-resolution orchard imagery. Second, sigmoid activation functions were substituted for ReLU layers in deeper network stages (beyond 15 convolutional layers), mitigating gradient vanishing issues while increasing target recognition efficiency by 18.6% as measured in frames processed per second. Third, a novel hybrid pooling approach combining max pooling and average pooling operations was implemented in reduction layers, utilizing 3×3 kernels with a stride of 2. This dual-mode methodology enhanced feature representation capabilities by 23.4% on pathological test sets, simultaneously improving model adaptability to varied disease manifestations while reducing overfitting risks through diversified feature learning pathways. The hybrid pooling design specifically strengthened robustness against lighting variations and occlusion artifacts common in canopy imagery, with ablation studies confirming 15.2% improvement in cross-environment generalization.

4. System application and performance evaluation

4.1. Experimental site overview

Field trials were conducted in two geographically distinct longan production regions: Fenjie Town in Gaozhou City, Maoming, featuring a contiguous plantation area of 66.7 hectares (1,000 mu), and Dongfeng Town in Zhongshan City, characterized by hilly terrain spanning 53.3 hectares (800 mu). These sites were strategically selected for their historical susceptibility to gray leaf spot (*Pseudocercospora* spp.) and fruit borer (*Conopomorpha sinensis*) infestations, making them ecologically significant test environments for Guangdong's longan cultivation systems.

4.2. System operation workflow

The integrated inspection process, demonstrated in Dongfeng Town, follows a four-stage operational sequence. Commencing with data acquisition, drones execute pre-programmed flight paths at an altitude of 15 m and a velocity of 5 m/s, capturing high-resolution imagery at 3-second intervals. Each 40-minute operational cycle (including takeoff and landing procedures) achieves complete coverage of 8-hectare (120 mu) orchard blocks. During data transmission, acquired images undergo real-time encryption and compression before being transmitted via 5G networks, with approximately 800 images from a standard sortie requiring about 6 minutes for complete cloud transfer. The processing and analysis phase leverages cloud computing resources to generate 0.1 m-resolution orthomosaics through automated image stitching, while concurrently executing AI-powered pest recognition within a consolidated 15-minute processing window. Culminating in decision output, the system synthesizes pathological heatmaps and precision treatment prescriptions that are instantaneously pushed to user devices through dedicated mobile applications, enabling growers to implement targeted interventions within 60 minutes of initial data acquisition.

4.3. Performance evaluation

Operational efficiency analysis in 6.7-hectare (100 mu) test plots revealed that the system reduced inspection time to 1.2 hours—representing a 53-fold efficiency gain compared to traditional 8-person-day manual inspections. Diagnostic accuracy reached 91.5% (± 2.1%), surpassing the 82.3% benchmark performance of experienced technicians. The system demonstrated particular efficacy in detecting canopy-top pathologies (89.2% recognition versus 37.5% manual detection), overcoming a critical limitation of ground-based observation.

Economic impact assessment documented a 32% reduction in pesticide application (decreasing from 3.8 to 2.6 kg/mu/year) and 65% lower labor costs (from ¥120 to ¥42/mu), resulting in 48% aggregate operational cost savings. Post-implementation yield analysis during the 2023 harvest season measured an average production of 1,250 kg/mu—18.7% higher than that of conventionally managed control plots—with all sampled fruit achieving 100% pesticide residue compliance under GB 2763-2021 standards.

User validation surveys (n = 30 growers) indicated 93.3% endorsement of the system's early warning timeliness, 86.7% affirmation of the practicality of prescriptions, and 80% intent to retain usage and refer the system to others. These metrics collectively demonstrate significant improvements in both operational efficiency and decision-making quality throughout the pest management value chain [13–15].

5. Conclusion and future prospects

This study investigates the practical application requirements in longan cultivation, considers the real-world scenarios of image recognition technology implementation, and has successfully developed a drone-based AI inspection system for longan pests and diseases, thereby realizing full-process intelligence from image acquisition to pest control decision-making. In line with the biological traits of longan, the drone data acquisition protocol has been optimized, which effectively addresses the difficulty in identifying pests and diseases on the canopy top. The large-scale dedicated dataset for longan pests and diseases has been constructed, encompassing major producing regions in Guangdong Province and the growth cycle of longan, and providing sufficient samples for model training and optimization. The proposed improved convolutional neural network model achieves a favorable balance between recognition accuracy and speed, which is capable of meeting the demands of practical applications. The application results indicate that the system significantly enhances the efficiency and precision of longan pest and disease control, reduces production costs, and holds great significance for promoting the digital transformation of the characteristic longan industry in South China. Despite the efforts made in this study, there are several aspects that require further research and improvement: Firstly, its adaptability to extreme weather conditions (such as heavy rains and dense fog) is limited. Under such weather conditions, drone flight and image acquisition will be affected to a certain degree. Secondly, the capability to identify early latent diseases (e.g., the incubation period of *Colletotrichum* pathogens causing anthracnose) is insufficient, making it difficult to detect in a timely manner at the initial stage of disease occurrence. Future research should expand the research scope to fully cover pests and diseases in all parts of longan, so as to establish a more comprehensive pest and disease early warning and diagnosis system.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Lan ZB, 2022, Research on Monitoring and Control Technologies for Crop Pests and Diseases. Horticulture & Seedlings, 42(12): 92–94.

[2] Huang KQ, 2021, Damage Symptoms and Control Methods of Major Pests and Diseases of Litchi. Modern Horticulture, 44(22): 40–41.

[3] Zhang Y, et al., 2022, Deep Learning Based Real-Time Crop Disease Detection Using UAV Imagery. Computers and Electronics in Agriculture, 2022: 107310.

[4] He Y, Zhang Z, Li J, et al., 2023, Improved AlexNet-Based Citrus Canker Detection Using UAV Images. Computers and Electronics in Agriculture, 208: 107723.

[5] FAO, 2021, Integrated Pest Management in Tropical Fruit Crops, United Nations, Rome.

[6] Redmon J, Farhadi A, 2018, YOLOv3: An Incremental Improvement. arXiv. https://doi.org/10.48550/arXiv.1804.02767

[7] Ren SQ, et al., 2023, Deep Learning-based Pest Detection in Precision Agriculture. Computers and Electronics in Agriculture, 205: 107399.

[8] Chen X, et al., 2023, UAV-Based High-Resolution Image Analysis for Banana Disease Detection. Precision Agriculture, 24(2): 305–325.

[9] Liu X, Wang Y, et al., 2024, A Novel Deep Learning Framework for Apple Pest Detection Using Drone Imagery. Computers and Electronics in Agriculture, 212: 107956.
[10] Zhang H, Li M, et al., 2022, Multispectral Drone Imagery for Early Detection of Potato Late Blight. Remote Sensing, 14(18): 4486.
[11] Li Y, et al., 2023, A Comparative Study of Deep Learning Algorithms for Pest Identification in Orchard Environments. Biosystems Engineering, 227: 237–248.
[12] Wang J, et al., 2024, A Machine Learning Approach to Predict Longan Yield Based on Environmental Factors and Pest Incidence. Agricultural Systems, 204: 103675.
[13] Zhu J, 2020, High-Yield Cultivation Techniques and Pest Control of Longan in Hainan. Farmer's Staff, (12): 80.
[14] Huang J, 2024, Green Prevention and Control Strategies for Diseases and Pests in Longan Planting. Agricultural Industrialization, (10): 34–36 + 47.
[15] Lin S, 2023, Longan Cultivation Management and Pest Control Techniques. Hebei Fruit Tree, (01): 51–53.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

A Collaborative Approach to Distributed Heterogeneous Process Engines for Cross-Organizations

Xuehu Zuo[1,2]*, Xin Shan[1,2], Zhongguo Yang[1,2]

[1]Beijing Key Laboratory on Integration and Analysis of Large-Scale Stream Data, Beijing 100144, China
[2]School of Information Science and Technology, North China University of Technology, Beijing 100144, China

**Author to whom correspondence should be addressed.*

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: In today's complex and rapidly changing business environment, the traditional single-organization service model can no longer meet the needs of multi-organization collaborative processing. Based on existing business process engine technologies, this paper proposes a distributed heterogeneous process engine collaboration method for cross-organizational scenarios. The core of this method lies in achieving unified access and management of heterogeneous engines through a business process model adapter and a common operation interface. The key technologies include: Meta-Process Control Architecture, where the central engine (meta-process scheduler) decomposes the original process into fine-grained sub-processes and schedules their execution in a unified order, ensuring consistency with the original process logic; Process Model Adapter, which addresses the BPMN2.0 model differences among heterogeneous engines such as Flowable and Activiti through a matching-and-replacement mechanism, providing a unified process model standard for different engines; Common Operation Interface, which encapsulates the REST APIs of heterogeneous engines and offers a single, standardized interface for process deployment, instance management, and status synchronization. This method integrates multiple techniques to address API differences, process model incompatibilities, and execution order consistency issues among heterogeneous engines, delivering a unified, flexible, and scalable solution for cross-organizational process collaboration.

Keywords: Distributed; Collaboration; Meta-process; Cross-organization; Business process; Process engine

Online publication: October 17, 2025

1. Introduction

Distributed systems, with their cross-domain, intelligent, and collaborative characteristics, provide a new perspective for organizations to examine and solve cross-domain and cross-network business challenges [1]. Workflow, on the other hand, has revolutionized the enterprise through automation and standardized process management. However, with the expansion of service scope and the diversification of customer needs, the traditional single-organization service model has been difficult to keep pace with the rapid development of the market [2,3].

Due to the technical differences between different organizations, such as differences in data formats and system interfaces, cross-organizational cooperation often faces many challenges. One of the most prominent issues is how to realize collaboration effectively [4]. In previous studies, the concept and practice of collaboration have been widely explored and applied. For example, a reference architecture, eSourcing Reference Architecture (eSRA), was designed by Pena *et al.* [5], which realizes dynamic inter-enterprise business process collaboration by supporting the modeling, mapping, sharing, and execution of different types of enterprise business processes through multi-layer frameworks and components. Shan *et al.* [6] investigated the mining of collaborative business processes across organizations and proposed a new approach that discovers Business Process Model and Notation (BPMN) 2.0 collaborative process models and their corresponding choreographies, and is able to handle business processes from different organizations. All these research results are good proof of the wide application and importance of collaboration today.

We contribute by designing a novel distributed architecture that is capable of managing and executing cross-organizational heterogeneous processes, so that many subtasks of a business process run separately on heterogeneous process engines, while ensuring consistency with the execution order of the original process subtasks. Standardized APIs and data formats through common operating interfaces are used to improve interoperability between systems of different organizations, thus realizing information exchange and resource sharing. A process model adapter is designed, which can automatically identify and convert the differences in process models between different organizations so that the process models can be deployed for execution in a collaborative environment.

2. Related work

2.1. Distributed system

With the rapid development of information technology, distributed systems have become an important part of the computer science field. In order to improve efficiency, reliability, and scalability, it has become a common solution to decentralize computational tasks to multiple computers [7,8].

In this context, Fang *et al.* [9] proposed a distributed process engine based on ZooKeeper implementation, which realizes the task compensation of the failed node by creating a process processing model and transferring the process flow logic being processed to other processing nodes to continue execution. In the article by Domingos *et al.* [10], a distributed workflow scheduling method based on “meta-processes” was proposed, which aims to solve the security problems and performance bottlenecks of traditional centralized workflow management systems, while supporting the dynamic adjustment of workflow models. The method realizes inter-node communication by constructing special meta-processes and through process variables.

In this paper, we improve the sub-process generation algorithm based on the article by Domingos *et al.* [10]. Compared to the graph-based breadth traversal method in the article by Domingos *et al.* [10], this paper adopts a simplified single-task segmentation strategy. The original method needs to traverse and analyze the whole flowchart, which is a cumbersome process. In contrast, the single-task segmentation method proposed in this paper treats each service task as an independent sub-process and contains the start event, service task, and end event. The advantage of this method is that it simplifies the process of creating sub-processes and directly makes each task independent. At the same time, the meta-process ensures the consistency between the sub-process and the original process in the execution order. The specific realization process will be mentioned in subsection 4.1.

2.2. Collaboration of heterogeneous process engines

In terms of distributed workflow management, extensive research has been conducted in academia and industry, and various solutions have been proposed. Klinger and Bodendorf [11] proposed a distributed workflow management system framework for cross-organizational workflow management. By analyzing the dependencies between tasks, the workflow is divided into multiple sub-workflows and assigned to the corresponding workflow management system for execution and management. Distributed collaborative management of cross-organizational workflows is achieved.

3. Collaborative approaches for distributed heterogeneous process engines

In order to solve the cross-organizational problems, such as data transfer, collaborative scheduling, state synchronization, etc., in the collaborative work of heterogeneous process engines, this paper proposes a distributed heterogeneous engine collaboration architecture based on the distributed workflow scheduling method based on the "meta-process" of Domingos *et al.* [10]. The architecture mainly consists of six parts: front-end interaction interface, business process preprocessing, scheduler, business process model adapter, common operation interface, and distributed heterogeneous nodes, as shown in **Figure 1**.

(1) Front-end interaction interface: Users can input business process requirements through the front-end interaction interface, call the API of the process engine to start process instances, query process status, submit user tasks, and receive process execution results.

(2) Business process preprocessing: Generate a collection of meta-processes and sub-processes by decomposing the original business process and add a listener for each sub-process. The listener is responsible for listening to the completion status of the sub-process and notifying the meta-process of the next step. Sub-process information is stored in the meta-process variables for meta-process scheduling and management.

(3) Scheduler: as the meta-process controlling the process, it is responsible for scheduling the execution order of the sub-processes, ensuring that the execution order of the sub-processes is consistent with the original business process, and coordinating the order between different nodes.

(4) Business process model adapter defines adaptation rules by analyzing the process model differences of different process engines. Construct regular expressions to find the keywords of differences in the process model files and replace them with the keywords of the target process engine.

(5) Generic operation interface: encapsulate the APIs of different process engines, provide a unified interface to manage process instances, tasks, process variables, etc. Through the mapping mechanism, map the invocation request of the generic interface to the API of the target process engine.

(6) Distributed heterogeneous nodes: Multiple process engine nodes of different types can deploy sub-processes in a distributed manner. The nodes communicate with each other through meta-process variables to coordinate the execution order.