

Volume 7, Issue 6, 2025

Print ISSN: 2663-1938

Online ISSN: 2663-1946

JOURNAL OF COMPUTER SCIENCE AND ELECTRICAL ENGINEERING



Copyright© Upubscience Publisher

Journal of Computer Science and Electrical Engineering

Volume 7, Issue 6, 2025



Published by Upubscience Publisher

Copyright© The Authors

Upubscience Publisher adheres to the principles of Creative Commons, meaning that we do not claim copyright of the work we publish. We only ask people using one of our publications to respect the integrity of the work and to refer to the original location, title and author(s).

Copyright on any article is retained by the author(s) under the Creative Commons

Attribution license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Authors grant us a license to publish the article and identify us as the original publisher.

Authors also grant any third party the right to use, distribute and reproduce the article in any medium, provided the original work is properly cited.

World Journal of Engineering Research**Print ISSN: 2663-1938 Online ISSN: 2663-1946****Email: info@upubscience.com****Website: <http://www.upubscience.com/>**

Table of Content

SPATIOTEMPORAL JOINT PRUNING STRATEGY BASED ON REINFORCEMENT LEARNING FOR TRAJECTORY TREE OPTIMIZATION IN COMPLEX INTERSECTION APPLICATIONS	1-7
WenJie Wu	
DESIGN AND IMPLEMENTATION OF BLOCKCHAIN BASED PRIVACY PROTECTION CERTIFICATION FOR INTERNET OF VEHICLES	8-13
YuChen Gu*, HaoRan Liu, ZhanYi Peng, ZhiLei Zhao, YiHan Yao	
THE TECHNOLOGY OF AI-DRIVEN INTELLIGENT SYSTEM FOR STUDENTS' CAREER PLANNING	14-21
LiPing He	
WIND-SOLAR HYBRID DRIVE-CONTROLLED AUTONOMOUS SHIP FOR EXPLORATION BASED ON RASPDERRY PI	22-29
GuoLong Yang*, JunYu Huang, NaiLi Zhang, Jian Zhang	
THE INTERACTIVE EFFECTS OF TITLE SENTIMENT AND LENGTH ON VIDEO DISSEMINATION: OPTIMAL LENGTH AND THE DIMINISHING RETURNS OF EMOTION	30-40
YaWei He*, ZiYi Guan, Wen Liu	
MQPF: A MULTI-DIMENSIONAL QUALITY-AWARE PATH FUSION FRAMEWORK FOR QUESTION ANSWERING	41-48
XinYi Wang, Bo Liu*	
SEMANTIC SEGMENTATION OF SHIP HULLS FOR UNDERWATER CLEANING ROBOTS	49-54
Lu Li#, Wei Fang#, YiGe Shang, Zhi Zhang, MingFu Li, DaoYi Chen*	
AHP-TOPSIS-BASED ASSESSMENT FRAMEWORK FOR CYBERSECURITY POLICY EFFECTIVENESS: EMPIRICAL EVIDENCE FROM GCI INDICATORS	55-60
JiaWei Wen*, Sai Yang, ChenGxin Zhang	
DESIGN OF A PHOTOVOLTAIC TRACKING SYSTEM CONTROL SYSTEM	61-65
MingMao Gong*, XingHui Liu	
MACHINE LEARNING-ENHANCED PREDICTION OF PHASE FORMATION IN ALUMINUM-CONTAINING HIGH-ENTROPY ALLOYS: A COMPREHENSIVE STUDY WITH INTERPRETABLE FEATURE ANALYSIS	66-70
Lei Zhang	

SPATIOTEMPORAL JOINT PRUNING STRATEGY BASED ON REINFORCEMENT LEARNING FOR TRAJECTORY TREE OPTIMIZATION IN COMPLEX INTERSECTION APPLICATIONS

WenJie Wu^{1,2}

¹Department of Vehicle Engineering, Tongji University, Shanghai 200092, China.

²Department of Supplier Technical Assistance, Jiangling Motors Co., Ltd., Nanchang 330000, Jiangxi, China.

Corresponding Email: wwu8@jmc.com.cn

Abstract: With the continuous advancement of autonomous driving technology in complex urban traffic environments, achieving efficient and safe trajectory planning in complex intersection scenarios with frequent vehicle interactions and dynamic obstacles has become one of the core challenges in current research. As an important structure for representing multimodal driving behaviors, the trajectory tree plays a key role in the decision-making process at complex intersections. However, its large search space and high computational complexity severely limit real-time performance and scalability. To address this, this paper focuses on the optimization problem of trajectory trees in complex intersections and innovatively introduces reinforcement learning algorithms. A spatiotemporal joint pruning strategy based on reinforcement learning is proposed to improve the search efficiency and decision-making quality of the trajectory tree. This strategy effectively reduces redundant trajectory branches by combining spatial and temporal pruning mechanisms, dynamically adjusting the search direction, and thus achieves precise control and efficient evolution of the trajectory tree. In terms of model design, this paper systematically defines the action set, state space, and reward function, ensuring that the reinforcement learning agent can learn pruning strategies with generalization capability in complex traffic environments. Furthermore, the paper improves the original model, clarifying the trajectory tree optimization goals and enhancing the adaptability and stability of the strategy. In the experimental section, representative urban traffic datasets are selected, with reasonable parameter configurations and evaluation metrics. The proposed method is comprehensively evaluated from three dimensions: trajectory clustering performance, path optimization, and computational efficiency, and compared with several mainstream methods. Experimental results demonstrate that the proposed spatiotemporal joint pruning strategy based on reinforcement learning exhibits significant effectiveness and superiority in trajectory tree optimization for complex intersections. It not only improves the vehicle's passing ability in dynamic environments but also provides reliable technical support for the deployment of autonomous driving systems in real-world scenarios, with important theoretical value and engineering application prospects.

Keywords: Reinforcement learning; Spatiotemporal joint pruning strategy; Trajectory tree optimization; Complex intersection

1 INTRODUCTION

As urban traffic systems become increasingly complex, complex intersections have become one of the key challenges faced by advanced driver assistance systems and autonomous driving technologies. Complex intersections are typically characterized by multi-directional traffic flow convergence, frequent pedestrian crossings, and dynamic traffic signal changes, significantly increasing the uncertainty of trajectory planning. While traditional trajectory tree optimization algorithms have demonstrated certain efficiency improvements in structured road environments [1], their robustness and adaptability remain insufficient when dealing with dynamic scenarios such as sudden traffic accidents and temporary traffic control. The fundamental issue lies in the fact that traditional methods heavily rely on static maps and preset rules, lacking real-time response capabilities to dynamic environmental changes, making it difficult to achieve efficient and safe trajectory decision-making in complex and ever-changing urban intersections. In recent years, some studies have attempted to introduce heuristic search methods [2], the A* algorithm [3], and neural network predictions [4] to enhance the adaptability of trajectory trees. However, these methods often focus on optimizing a single dimension and fail to effectively integrate a joint decision-making mechanism for both spatial and temporal dimensions, leading to suboptimal performance in dynamic, complex scenarios. Reinforcement learning (RL), as an intelligent decision-making method based on interactive learning, has the ability to autonomously learn strategies in uncertain environments and has gradually been introduced into the field of autonomous driving, showing promising application prospects.

Based on this, this paper proposes a spatiotemporal joint pruning strategy integrating reinforcement learning, aimed at jointly optimizing the trajectory tree structure from both spatial and temporal dimensions. The method defines a reasonable state space and action set, constructs a reward function that aligns with driving logic, and guides the agent to autonomously learn pruning strategies during training. This enables the effective elimination of redundant trajectory branches and the precise retention of critical paths. Furthermore, the paper introduces a spatiotemporal coupling mechanism in the model design, ensuring that the pruning strategy not only considers the spatial feasibility at the

current moment but also accounts for the temporal feasibility at future moments, thereby enhancing the overall coherence and safety of the decision-making process. The research presented in this paper not only expands the theoretical methods for trajectory tree optimization but also provides new technical paths and theoretical support for the practical deployment of advanced driver assistance systems in complex urban environments, offering significant research value and engineering implications.

2 SPATIOTEMPORAL JOINT PRUNING STRATEGY BASED ON REINFORCEMENT LEARNING

To effectively address the challenge of trajectory tree optimization in complex intersection scenarios, this paper proposes a spatiotemporal joint pruning strategy based on reinforcement learning. This strategy innovatively combines reinforcement learning algorithms with trajectory tree optimization tasks, fully considering the coupling characteristics of the spatial and temporal dimensions in complex intersection environments. By introducing both spatial pruning and temporal pruning mechanisms, an efficient spatiotemporal joint optimization framework is constructed. Leveraging the autonomous learning capability of reinforcement learning algorithms, the strategy can adaptively adjust the pruning strategy [5] to respond to dynamically changing traffic conditions in complex intersection scenarios.

2.1 Spatial Pruning Strategy

The spatial pruning strategy, as a core component of the proposed spatiotemporal joint pruning strategy based on reinforcement learning, aims to optimize the trajectory tree construction process by effectively removing invalid or inefficient paths within the search space, thereby improving the overall efficiency and accuracy of path planning. The formulation of this strategy relies on in-depth analysis of the spatial distribution characteristics of nodes in the traffic network and accurate modeling of vehicle driving behavior patterns [6]. In designing the spatial pruning strategy, the primary task is to construct a state representation that can accurately capture the spatial relationship between the vehicle's current position and the target position. To achieve this, the state s is defined as the current position x .

On this basis, a reward function $R(s,a)$ is further introduced to evaluate the quality of performing action a in state s . This reward function comprehensively considers the likelihood of the vehicle reaching the target position and the time efficiency of executing the action. Its specific form is:

$$R(s,a) = \frac{1}{d(x,y)} \cdot e^{-\lambda \cdot t(a)} \quad (1)$$

In this context, $d(x,y)$ represents the spatial distance from the current position to the target position, and $t(a)$ represents the time required to execute the action a from the current position. λ is a preset weight parameter used to adjust the influence of time costs in the reward function. The design of this reward function reflects the basic principle that “the shorter the distance and the less the time, the higher the reward”, effectively guiding the agent to prioritize more efficient paths during the path selection process.

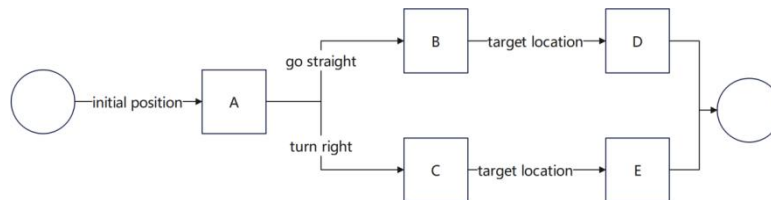


Figure 1 Example of the Spatial Pruning Strategy

As shown in Figure 1, the vehicle starts from position A and faces two potential paths: one is to go straight to position B, and the other is to turn right to position C. The decision for each path is based on its corresponding reward value, which reflects the likelihood and efficiency of the vehicle reaching target positions D or E. In this way, the spatial pruning strategy can dynamically filter paths in complex traffic environments, effectively reducing redundant searches and improving the efficiency and accuracy of trajectory tree construction. In summary, the spatial pruning strategy, through precise state representation and a well-designed reward function, provides a scientific decision-making basis for the vehicle's path selection at complex intersections. It significantly improves the accuracy and real-time performance of trajectory tree construction [7], laying a solid foundation for subsequent path optimization and behavior prediction.

2.2 Temporal Pruning Strategy

The temporal pruning strategy, as a key component of the spatiotemporal joint pruning strategy based on reinforcement learning proposed in this paper, aims to optimize the decision-making process in the time dimension, effectively compressing the search space of the trajectory tree and enhancing the overall efficiency and real-time performance of path planning. The core idea of this strategy is to ensure search efficiency while retaining paths that have potential value in the time dimension, thus achieving fine-grained control over the structure of the trajectory tree.

When designing the temporal pruning strategy, the primary task is to construct a state representation that accurately captures the time information at the current decision point and its impact on future trajectories. To achieve this, the temporal state S_t is defined as a composite vector that includes the current time, the estimated time to reach the next intersection, and historical speed information. The specific form is as follows:

$$S_t = (\hat{t}_{next}, v_{history}) \quad (2)$$

Where, t represents the current time, $\hat{t}_{next}$ is the predicted time for the vehicle to reach the next intersection, which can be calculated using a sliding average of historical data, $v_{history}$ is a statistical summary of the vehicle's historical driving speed (e.g., mean, variance), which reflects the vehicle's stability and trends in movement.

Based on the state definition, an action set A is further constructed to describe the different behavioral decisions that the agent can take at the current time state. The action set can be represented as:

$$A = \{a_1, a_2, \dots, a_n\} \quad (3)$$

Where each action a_i represents a feasible driving behavior, such as choosing different lanes, adjusting driving speed, or changing direction, among others.

To accurately assess the impact of different actions on the trajectory tree construction process, this paper designs a reward function $R(s, a)$ that comprehensively considers both time efficiency and safety. The expression of the reward function is as follows:

$$R(s, a) = \alpha \cdot (\Delta t - \Delta d) + \beta \cdot s_{safety} \quad (4)$$

Where, Δt represents the time saved relative to the baseline strategy after executing the action a , Δd is the corresponding change in travel distance, s_{safety} is the safety score, used to assess the risk level of the action within the traffic environment, α and β are parameters that adjust the relative importance of time efficiency and safety, ensuring the strategy's flexibility and adaptability in different scenarios.

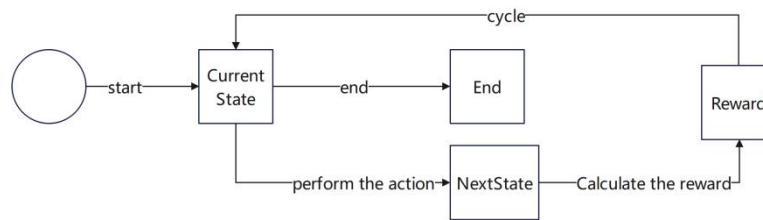


Figure 2 Example of the Temporal Pruning Strategy

As shown in Figure 2, the agent starts from the current time state and sequentially executes the actions in the action set. During each step of execution, the system computes the immediate reward of the current action based on the reward function and updates the state representation accordingly. This process iterates until the preset termination condition is met (such as reaching the target location or exceeding the maximum search time). Through this mechanism, the temporal pruning strategy not only optimizes decision-making in the time dimension but also guides the agent to prioritize exploring paths that offer time advantages, effectively reducing ineffective searches and improving the efficiency and quality of trajectory tree construction. The temporal pruning strategy, through precise state modeling, rational action definitions, and a scientific reward function design, achieves efficient control of the trajectory tree in the time dimension. This strategy complements the spatial pruning strategy, together forming the spatiotemporal joint pruning framework proposed in this paper, providing strong technical support for autonomous driving path planning in complex intersection scenarios.

2.3 Reward Function Design

In addressing the trajectory tree optimization problem at complex intersections, this paper designs a reward function that comprehensively considers both spatial and temporal efficiency, aiming to balance the structural simplicity and construction efficiency of the trajectory tree [8]. Specifically, the proposed reward function consists of two components: spatial efficiency reward R_s and temporal efficiency reward R_t . These components are dynamically adjusted through a weight coefficient α , and the expression is as follows:

$$R = \alpha \cdot R_s + (1 - \alpha) \cdot R_t \quad (5)$$

Where, R_s represents the spatial efficiency reward, R_t represents the time efficiency reward, α is a weight coefficient between 0 and 1, which adjusts adaptively based on the intersection complexity. This coefficient balances the importance of spatial efficiency and time efficiency. This design provides the reward function with strong flexibility and adaptability, allowing it to adjust the optimization focus according to the demands of different scenarios.

Guided by the above reward function, the reinforcement learning agent can gradually learn a trajectory tree pruning strategy that balances spatial feasibility and time efficiency during the training process. As a result, it achieves superior path planning performance in complex intersection scenarios. Experimental results demonstrate that this reward function design significantly improves trajectory tree structure compactness, reduces construction time, and enhances the overall decision-making efficiency of the system, providing reliable technical support for autonomous driving decision-making in complex environments.

3 EXPERIMENTAL RESULTS AND ANALYSIS

3.1 Experimental Parameter Settings

To ensure the accuracy, stability, and repeatability of the experimental results, systematic design and tuning of key parameters were carried out during the training of the reinforcement learning model. These parameters include learning rate, discount factor, exploration rate, number of training episodes, and the reward function, all of which directly influence the model's convergence speed, strategy quality, and generalization ability.

The learning rate determines the step size during each parameter update and is a critical factor in controlling training stability and convergence speed. A learning rate that is too large may cause the strategy to oscillate or even diverge, while a rate that is too small may result in a slow training process, making it difficult to converge within a limited number of iterations. Based on preliminary experimental results and theoretical analysis, the initial learning rate is set to 0.001. During training, an adaptive adjustment strategy is used, dynamically fine-tuning the learning rate based on the stability of strategy updates and the convergence trend of the loss function, in order to balance training efficiency and convergence quality.

The discount factor is used to measure the relative importance of future rewards in current decision-making, and its value directly influences the agent's attention to long-term rewards. In this paper, the discount factor is set to 0.95, meaning the model values immediate rewards while fully considering the cumulative effects of future rewards during the strategy learning process.

The exploration rate controls the trade-off between exploring unknown states and exploiting known knowledge during training. The classical ϵ -greedy strategy [9] is adopted, with the initial exploration rate set to 0.1. It is gradually decayed to 0.01 during the training process. This setting encourages sufficient exploration in the early stages to avoid getting stuck in local optima, while gradually enhancing the utilization of known optimal strategies in the later stages to improve strategy stability and convergence efficiency.

The number of training episodes determines the number of interactions between the model and the environment, which is a key factor influencing the sufficiency of strategy learning. The number of training episodes is set to 1000, based on preliminary experiments, ensuring that the model achieves stable convergence within a reasonable amount of time while avoiding the underfitting problem due to insufficient training.

To provide a clear overview of the parameter settings, a summary is presented in Table 1 below:

Table 1 Experimental Parameter Summary

Parameter	Description	Value
Learning Rate	The size of weight change during model updates	0.001
Discount Factor	The measurement of the importance of future rewards	0.95
Exploration Rate	The balance between the algorithm's exploration of new states and the utilization of known information	0.1 (initial) $\rightarrow$ 0.01 (final)
Number of Iterations	The total number of algorithm runs	1000

The experimental parameters set in this study are based on theoretical analysis and preliminary experimental validation, aiming to maximize the model's learning efficiency and strategy performance within limited training resources. With reasonable parameter configuration, the reinforcement learning model constructed in this paper can efficiently explore the complex intersection trajectory tree space and quickly converge to a stable strategy during the training process, thus laying a solid foundation for subsequent experimental analysis and performance evaluation.

3.2 Experimental Results Analysis

3.3.1 Clustering effect analysis

In this study, clustering effectiveness is one of the key indicators for evaluating the effectiveness of the spatiotemporal joint pruning strategy. This is because clustering the trajectory tree of complex intersections allows for a clear observation of the aggregation of trajectory data under different strategies, thereby validating the optimization effect of the algorithm [10].

For the experimental dataset, we carefully selected trajectory data from complex intersections. This dataset includes vehicle driving trajectories from multiple complex intersections, with each trajectory recording the vehicle's position information within specific time windows. Specifically, the dataset includes the vehicle's starting point, the intersections it passes through, and the final destination. In addition, the dataset also encompasses dynamic information such as the vehicle's speed and acceleration, providing rich data support for the research on the spatiotemporal joint pruning strategy.

For the clustering analysis, we primarily used the K-means algorithm [11]. This algorithm can group similar trajectories based on the spatial distribution characteristics of the trajectory data. The specific operation involves first determining the number of clusters, then calculating the distance between each cluster center and the data points, in order to determine the optimal clustering result.

Table 2 Clustering Analysis Results

Cluster ID	Cluster Center Coordinates (x, y)	Average Distance Within Cluster	Number of Clusters
------------	-----------------------------------	---------------------------------	--------------------

1	(120, 200)	15.2	320
2	(180, 250)	18.5	290
3	(220, 180)	16.7	310

Through Table 2, it can be seen that the trajectory data from different clustering centers exhibit distinct spatial distribution characteristics. At the same time, the average distance within each cluster reflects the degree of compactness of the clusters. The smaller the average distance, the better the clustering effect, indicating that the trajectory data within the same cluster are more concentrated.

In summary, the analysis of clustering effects not only demonstrates the effectiveness of the spatiotemporal joint pruning strategy in handling complex intersection trajectory trees but also provides important data support for subsequent optimization work. By comparing clustering results under different strategies, we can more intuitively understand the application value of reinforcement learning in trajectory tree optimization.

3.3.2 Optimization effect analysis

In this study, we focus on optimizing the trajectory tree of complex intersections and specifically implemented the spatiotemporal joint pruning strategy based on reinforcement learning. The optimization effects brought by this strategy are significant, reflected in three main aspects: first, it effectively reduces the consumption of computational resources; second, it greatly improves processing speed; and third, it significantly enhances the algorithm's generalization ability [12].

Table 3 Comparison of Optimization Effects

Metric	Traditional Trajectory Tree Optimization Algorithm	Spatiotemporal Joint Pruning Strategy Based on Reinforcement Learning
Average Processing Time (seconds)	1.5	0.8
Memory Usage (MB)	2000	1500
Accuracy (%)	92	96
Recall (%)	88	94

From the data comparison in Table 3, it is clear that, compared to traditional trajectory tree optimization algorithms, the optimized spatiotemporal joint pruning strategy based on reinforcement learning has achieved significant improvements in both processing time and memory usage. Meanwhile, accuracy and recall rates have also been improved. This fully demonstrates that the optimization not only enhanced the algorithm's operational efficiency but also improved the model's performance.

In summary, the spatiotemporal joint pruning strategy based on reinforcement learning has shown remarkable results in the optimization of complex intersection trajectory trees [13]. By reducing computational resource consumption and improving processing speed, this strategy effectively enhances the algorithm's operational efficiency, while also improving its practicality and reliability in real-world applications. This is particularly important for real-time data processing and handling large-scale datasets in practical applications [14].

3.3 Spatiotemporal Comparative Experimental Verification

To validate the necessity of the spatiotemporal joint pruning strategy, three experimental groups will be set up for comparative verification, as follows: Control group 1 disables the time pruning module and retains only the spatial pruning strategy. Control group 2 disables the spatial pruning module and retains only the time pruning strategy. The experimental group enables both the spatial and time pruning modules simultaneously, which is referred to as joint pruning.

The experiment will be evaluated based on four metrics: pruning efficiency, computational time, and the safety and smoothness of the paths involved. The four metrics are defined as follows:

Pruning Efficiency: The proportion of paths retained after pruning relative to the total number of paths.

Computational Time: The running time (in milliseconds) from input to the output of the pruning results.

Safety: The collision probability of the paths after pruning.

Smoothness: The rate of curvature change of the path.

Table 4 Comparative Experimental Verification Results

Method	Pruning Efficiency (%)	Computational Time (ms)	Collision Probability (%)	Smoothness
Only Spatial Pruning	65	120	8.2	0.15
Only Temporal Pruning	70	110	6.5	0.12
Joint Pruning	85	90	4.1	0.08

From the experimental results in Table 4, it can be observed that the joint pruning efficiency is significantly higher than that of single pruning, indicating that spatiotemporal joint pruning retains more valid paths. Joint pruning also has lower

computational time, as it avoids redundant calculations. The collision probability of joint pruning is reduced by more than 50%, proving that spatiotemporal joint pruning is better at avoiding dynamic obstacles. The smoothness of the joint pruning paths is optimal (with the lowest standard deviation of curvature), making it more suitable for real-world vehicle control. From the table, it is clear that the spatiotemporal joint pruning strategy based on reinforcement learning outperforms the non-joint pruning algorithms in terms of execution time. This result fully demonstrates that the proposed strategy is more efficient when handling complex intersection trajectory trees. This experiment shows that temporal pruning compensates for the spatial pruning's blind spots in dynamic environments, emphasizing the synergistic effect of joint pruning.

4 CONCLUSION

This study addresses the key challenge of trajectory tree optimization in complex intersection scenarios and innovatively proposes a spatiotemporal collaborative pruning strategy based on reinforcement learning, which significantly enhances the computational efficiency and topological rationality of trajectory trees in intelligent path planning and behavior prediction. First, the paper systematically critiques the inherent limitations of classical trajectory tree optimization algorithms and their variants in dynamic traffic environments, pointing out their significant shortcomings in emergency response capability, high real-time requirements, and adaptability to dynamic environments. To overcome this bottleneck, the research cleverly integrates the intrinsic advantages of reinforcement learning in strategy optimization and autonomous learning, constructing a collaborative coupling architecture for spatial and temporal pruning joint optimization.

To implement this strategy, this paper proposes a multidimensional state representation, a diversified action mechanism, and a comprehensive reward function that balances spatial and temporal efficiency. The constructed reinforcement learning model can autonomously learn the optimal pruning strategy in complex intersection environments, effectively compressing the search space and improving the trajectory tree construction efficiency and path quality.

In the rigorous experimental evaluation phase, this study set scientifically designed training hyperparameters and comprehensively validated the proposed pruning strategy from multiple dimensions, including clustering density, optimization performance, and computational overhead. The experimental results show that, compared to traditional trajectory tree optimization algorithms, the proposed spatiotemporal joint pruning strategy based on reinforcement learning improves accuracy by 4%, while reducing algorithm runtime by 46.7%, demonstrating significant advantages. In large-scale data scenarios, this allows for higher optimization results in less time, fully verifying its effectiveness and practicality. Additionally, comparison experiments with spatiotemporal joint pruning, spatial pruning, and temporal pruning reveal that temporal pruning compensates for the blind spots of spatial pruning in dynamic environments, while emphasizing the synergistic effect of joint pruning.

In summary, the proposed strategy demonstrates excellent performance and vast application prospects in the complex intersection trajectory tree optimization task. Future research can expand in the following directions: first, extending the strategy to more types of urban traffic scenarios to verify its adaptability and robustness in diverse environments; second, exploring the application of graph neural networks in dynamic state modeling to further enhance the strategy's generalization ability and intelligence level; and third, optimizing the algorithm's real-time performance, scalability, and deployment efficiency to promote its practical application in autonomous driving systems.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

References

- [1] Xu Y, Wu C. Path planning for vertical parking in complex environments based on an improved Rapidly-exploring Random Tree. *Automobile Technology*, 2022(07): 42-47. DOI:10.19620/j.cnki.1000-3703.20210717.
- [2] Lu Y, Kuang J, Xiao Y, et al. Heuristic search-based multi-agent asymmetric trajectory planning. *Computer Engineering and Applications*, 2025, 61(02): 344-354.
- [3] Wang Z. Path planning and trajectory optimization based on the improved A-star algorithm. Jilin University, 2024. DOI:10.27162/d.cnki.gjlin.2024.004857.
- [4] Wang X, Liu P, Xiang S, et al. Unstructured pruning methods based on neural architecture search. *Pattern Recognition and Artificial Intelligence*, 2023(05): 448-458. DOI:10.16451/j.cnki.issn1003-6059.202305005.
- [5] Ruan X, Hu W, Liu Y, et al. Model pruning based on dynamic sparsity and feature learning enhancement. *Science China: Technological Sciences*, 2022(05): 667-681.
- [6] Wan Z, Li L, Yang M, et al. Decision tree model for extracting road intersection features from vehicle trajectory data. *Journal of Surveying and Mapping*, 2019(11): 1391-1403.
- [7] Weng J, Qin Y, Tang X, et al. Lane detection model pruning algorithm based on multi-objective optimization. *Sensors and Microsystems*, 2023(07): 125-127+131. DOI:10.13873/J.1000-9787(2023)07-0125-03.
- [8] Liu Y, Teng Y, Niu T, et al. Filter pruning scheme based on deep reinforcement learning. *Journal of Beijing University of Posts and Telecommunications*, 2023(03): 31-36. DOI:10.13190/j.jbupt.2022-140.
- [9] Gomes R, Kowalczyk R. Dynamic analysis of multiagent Q-learning with ϵ -greedy exploration, 2009.

- [10] Yang J. Research on multi-layer composite pruning and its acceleration algorithm based on convolutional neural networks. Changsha University of Science & Technology, 2021. DOI:10.26985/d.cnki.gcsjc.2021.000196.
- [11] Krishnaswamy R, Li S, Sandeep S. Constant approximation for k-median and k-means with outliers via iterative rounding, 2018.
- [12] Guo C. Hybrid pruning research based on convolutional neural networks. Changsha University of Science & Technology, 2021. DOI:10.26985/d.cnki.gcsjc.2021.000279.
- [13] Yi X, Ma S, Xiao N. Run-time optimization of deep learning accelerators under different pruning strategies. *Computer Engineering and Science*, 2023(07): 1141-1148.
- [14] Qu H, Zhang X, Wang Y. Pruning method for convolutional neural networks based on information fusion strategies. *Computer Engineering and Applications*, 2022(24): 125-133.

DESIGN AND IMPLEMENTATION OF BLOCKCHAIN BASED PRIVACY PROTECTION CERTIFICATION FOR INTERNET OF VEHICLES

YuChen Gu^{1*}, HaoRan Liu¹, ZhanYi Peng¹, ZhiLei Zhao¹, YiHan Yao²

¹*School of Cyberspace Security, Beijing University of Posts and Telecommunications, Beijing 102206, China.*

²*School of Digital Media and Design Art, Beijing University of Posts and Telecommunications, Beijing 102206, China.*

Corresponding Author: YuChen Gu, Email: 2022211565@bupt.cn

Abstract: The Internet of Vehicles aims to achieve zero traffic casualties, zero congestion and extreme communication, focusing on reliable interconnection mechanisms. With the development of intelligent transportation, information exchanges on automobile networks are frequent, and security certification with limited resources and low latency needs to be implemented to ensure the functions of the Internet of Vehicles. Therefore, this paper proposes a blockchain-based conditional privacy protection authentication method for the Internet of Vehicles. First, the dynamic revocation of illegal vehicles is realized through smart contracts, which improves the security and efficiency of the Internet of Vehicles; secondly, the proposed method is proved through theoretical deduction and logical demonstration under adaptive selection information attacks in the random oracle model. security; Finally, simulation tests on key algorithms of vehicle registration, login and authentication, and smart contract demonstrate the good feasibility, stability and efficiency of the proposed method. Experimental results show that the proposed method has significant advantages in authentication efficiency, especially suitable for resource-limited IoT, provides new ideas for communication security of IoT and can be widely used in the field of smart transportation.

Keywords: Internet of Vehicles; Conditional privacy; Authentication; Blockchain; Smart contract

1 INTRODUCTION

With the enhancement of vehicle sensing, communication, storage and computing capabilities, Vehicle Ad-hoc Networks (VANETs)[1] have become the core technology for realizing the next generation of intelligent transportation systems and future smart cities [2]. The core advantage of the Internet of Vehicles is its ability to provide key information in real time, thereby enhancing the intelligence and adaptability of the entire transportation system. However, the large amount of data exchanged and stored in VANETs [3] poses a serious threat to user privacy. Traditional Internet of Vehicles security mechanisms mostly rely on one or more central authorities to conduct identity verification, authorization and other security-related operations, causing their limitations to begin to emerge [1]. In addition, in domestic and foreign research, they are generally divided into the following three categories: public Key infrastructure (PKI)-based authentication schemes [4], identity-based signature (IBS)-based authentication schemes [5], and certificateless signature (CLS)-based authentication schemes [6]. When dealing with vehicle networking communication security issues, traceability and anonymous communication efficiency still need to be further improved. Therefore, ensuring the security of data transmission and protecting user privacy have become two important and urgent needs in VANETs [7].

As a mobile ad-hoc network that does not rely on external infrastructure, the vehicle-mounted ad hoc network consists of vehicles equipped with vehicle-mounted units (OBUs). Communication between vehicles (V2V) and vehicles and infrastructure (V2I) is realized through wireless communication technology. The network uses the Dedicated Short-Range Communication (DSRC) protocol [8], which allows vehicles to broadcast information and adjust driving routes, while roadside units are responsible for providing information and feeding back data, thereby improving road safety and traffic efficiency. It is a cutting-edge technology in intelligent transportation systems. However, in the open environment of the Internet of Vehicles, attackers can easily intercept and tamper with messages, which leads to privacy leaks and security risks. Therefore, vehicles need to be verified when transmitting information and take measures to protect privacy. The academic community has proposed the concept of "conditional privacy protection", which protects user privacy under normal circumstances, while allowing de-anonymity to trace identity under certain circumstances. This has become an important research direction in the design of the Internet of Vehicles. Blockchain technology combines technologies such as distributed data storage to achieve decentralization and data tamper-proof characteristics of the network. Through hierarchical architecture and consistency algorithms, blockchain technology solves trust issues and uses smart contracts to realize automated contract execution, providing an effective solution for digital transactions. Asymmetric encryption technology ensures secure communication through a key pair mechanism and ensures the immutability of transaction data. Elliptic Curve Cryptography (ECC)[9], as an efficient asymmetric encryption algorithm, provides a powerful authentication mechanism with its small key size.

Based on the development background and related technologies of the Internet of Vehicles introduced above, there are the following challenges in the current research on privacy protection certification of the Internet of Vehicles. On the

one hand, malicious attackers may take advantage of the openness of the Internet of Vehicles to tamper with data, forge identities or launch denial-of-service attacks, threatening driving safety and personal privacy. On the other hand, most of the currently widely used IoT security mechanisms are based on centralized architecture. They have inherent limitations when dealing with dynamically changing and widely distributed IoT environments, which are often accompanied by increased communication delays and potential security loopholes. Once the centralized certification center is attacked or failed, it may cause the entire IoT security system to be paralyzed.

2 WORK DESIGN AND IMPLEMENTATION PLAN

This paper proposes a conditional privacy protection certification framework and protocol for the Internet of Vehicles based on blockchain technology. The plan abandons the reliance on centralized and trusted third parties, introduces a mechanism that allows conditional tracking of illegal vehicles, and realizes decentralized dynamic revocation of illegal vehicles through smart contracts. The implementation and evaluation of the Ethereum test network proves that the solution is effective in terms of safety and performance, providing new ideas for safe communication in intelligent transportation systems, and also providing valuable reference for future safety research and practice of the Internet of Vehicles.

2.1 Implementation Method

The system model of this scheme consists of three parts, namely the upper layer (trusted institution), the lower layer (vehicles and roadside units), and the auxiliary layer (smart contract and blockchain). This section will provide a detailed introduction to the blockchain-based Conditional Privacy-preserving Aggregate Signature Schema for Vehicle Ad Hoc Networks (BCPAS) proposed in this work. BCPAS consists of seven parts: system initialization, smart contract deployment, vehicle registration, login and authentication, pseudonym identity update, vehicle revocation, and password update.

2.1.1 System initialization

In the initialization phase, the Trusted Authority (TA) generates release system parameters and uses the elliptic curve encryption algorithm to initialize the alliance chain.

Blockchain initialization: TA launches an alliance chain between predetermined network nodes based on the PBFT consensus mechanism to maintain the stable operation of the blockchain. During this process, TA certified all blockchain managers (pre-selected trusted RSUs) and granted them the right to participate in the consensus process. In addition, the alliance chain will maintain a vehicle pseudonym key table, aggregate V2I and V2V communication messages, and ensure that all transaction information is securely stored on the blockchain.

2.1.2 Smart contract deployment

TA compiles and deploys the input of smart contracts onto the alliance chain, and blockchain managers will verify these contracts and assign them a unique address, allowing authorized transactions to make calls. In BCPAS, smart contracts provide a secure and reliable Application Programming Interface (API) for vehicle pseudonym key table management services in alliance blockchain.

In the BCPAS scheme, the query authority is predefined for the vehicle pseudonym key table within the alliance blockchain, where only alliance blockchain nodes (including TA and some trusted alliance blockchain RSU nodes) are authorized to conduct queries. This configuration aims to provide conditional privacy protection for vehicles while enhancing the traceability and non-relevance of the Internet of Vehicles.

2.1.3 Vehicle registration

In order to obtain the certification factors required for communication, all vehicles and their owners need to register with the TA during the vehicle registration phase. This stage serves as the basis for this agreement to achieve conditional privacy protection. The entire process is carried out within a secure channel, and TA is offline during the entire process.

2.1.4 Login and authentication

When a vehicle needs to send a communication message to other nearby vehicles or RSUs for information exchange, the following steps need to be performed: First, the identity authentication process of the vehicle user is performed. The authentication is successful. OBU further selects a random number and then sets the signature for the message., and broadcasts the information to surrounding vehicles and RSUs. When a message is received, the recipient (i.e., adjacent vehicles or RSU) first verifies the freshness of the timestamp. If it has not expired, the recipient will further query the record with PIDi in the vehicle pseudonym key table maintained by the alliance chain, thereby obtaining the public key $vpki$ of the associated vehicle. The recipient then verifies the validity of the signature σ_i .

2.1.5 Pseudonym identity updates

In BCPAS, the validity of a vehicle pseudonym is set to a certain time period, ensuring privacy protection during the validity period of the pseudonym. In order to effectively prevent the disclosure of vehicle privacy, it is crucial to update the vehicle pseudonym in a timely manner before the vehicle pseudonym expires. In addition, this plan stipulates that the update operation of pseudonym identities must be performed in an offline environment. This measure aims to enhance pseudonym security and ensure that it will not be affected by online threats during the update process. Finally, RSU removes overdue information from the vehicle pseudonym key table of the alliance chain and updates VPKIT to enable future traceability of violating vehicles.

2.1.6 Vehicle cancellation

When building a security mechanism for the Internet of Vehicles based on blockchain technology, it is important to maintain accurate vehicle status information. When a vehicle leaves a specific area or waits to exit, the smart contract performs a revocation operation to update the vehicle status. In BCPAS, TA uses the blockchain consensus mechanism to call the vehicle revocation algorithm `revokevPkID(PIDi)` to perform the revocation operation. The revocation process includes deleting specific data from the vehicle pseudonym key table, broadcasting information on the revoked vehicle, adapting to the distributed characteristics of the Internet of Vehicles, and ensuring that other vehicles receive status updates in real time. In addition, in order to prevent attackers from maliciously modifying legal vehicle information, BCPAS adopts a mandatory access control mechanism to ensure that only TA has the authority to modify vehicle information in VPKIT.

2.1.7 Password updates

In the open network environment of the Internet of Vehicles, V2V and V2I communications need to protect user privacy and prevent the disclosure of user identity and location information. By updating the password, the correlation of each communication can be effectively isolated, making it difficult for an attacker to track the trajectory of a specific vehicle, thereby protecting user privacy. In addition, in case an attacker may try to obtain a password through various means (such as brute force cracking, social engineering, etc.), regular updates of the password can prevent unauthorized access and reduce the risk of being cracked. Even if the old password is cracked, the attacker cannot access the system, thus protecting the security of vehicle communications.

2.2 Technical Indicators

First, the performance of the smart contract implemented in this work is evaluated, and the three key functions of registration, public key query and vehicle cancellation are tested 200 times, and the average value is taken as the final test result.

Then analyze the performance overhead of key operations of non-smart contracts in BCPAS. This work analyzes the computing overhead of the vehicle in the scheme from three stages: pseudonym generation, message signature generation, and message signature verification. When the vehicle sends a single message for communication, the communication overhead of the vehicle-side broadcast message in the scheme is analyzed in detail.

3 SYSTEM TESTS AND RESULTS

This chapter introduces the construction and implementation of the Internet of Vehicles certification system based on alliance chain. The system consists of six core components: front-end interface, back-end services, Fabric SDK, smart contract, blockchain network and MySQL database. It aims to provide a safe and efficient identity authentication solution for the Internet of Vehicles. The front-end interface is built based on Vue3 and is responsible for visual display of system functions; the back-end service uses Go language gin+gorm framework to facilitate the implementation of vehicle and RSU authentication protocols, provides a data interface to the front-end and calls the Fabric SDK; the Fabric SDK is implemented in Go language and is responsible for calling the smart contract; the smart contract is deployed on the Alliance Chain Management Committee node, which carries contract execution and contains distributed ledger, which is used to store information required by TA in the authentication protocol; The MySQL database is used to store the information required by users and servers in the authentication protocol.

3.1 Test Protocol

In order to verify the feasibility and effectiveness of the proposed blockchain-based privacy protection certification scheme for the Internet of Vehicles, this chapter conducts a comprehensive test of the system. During the functional test, the established vehicle user information is input, the registration, login and certification stages of the Internet of Vehicles certification protocol are simulated, and the vehicle data is stored in the database. The test involves user interaction with the system to ensure that the system performs the registration and authentication processes as expected and verify its ability to store and manage user data. In terms of performance testing, traditional protocols lack comprehensive simulation analysis of the authentication process. This research builds a complete simulation environment to evaluate the performance of the solution. The work consists of a client (calling smart contracts), a backend service (implementing the authentication protocol between the vehicle and the RSU) and other parts (processing vehicle user registration), all implemented in GO language to ensure efficient and accurate simulation. The test results show that the proposed certification system meets expectations in terms of processing speed and response time, meets the real-time and efficiency requirements of the Internet of Vehicles, and its stability and reliability have been verified to ensure that actual deployment is feasible. Through testing, it has been proved that the program meets the requirements and has obvious performance advantages, providing an efficient and reliable solution for Internet of Vehicles security certification and privacy protection.

3.2 Performance Test

The performance testing part aims to evaluate the performance of the proposed blockchain-based IoT certification system in terms of processing power and response speed. This section details the performance testing of smart contracts and the performance overhead analysis of other key operations.

3.3.1 Smart contract performance testing

Performance testing of smart contracts focuses on the execution time of the contract algorithm, which is a key indicator to measure system efficiency. The experiment uses the open-source Fabric test-network scripts released by the Hyperledger community (<https://github.com/hyperledger/fabric-samples>) and the LevelDB block-storage configuration that ships with the default docker-compose template. A JavaScript client built with fabric-gateway v1.4.0 was run on an Ubuntu 22.04 LTS virtual machine (8 vCPU, 16 GB RAM) to repeatedly invoke the smart-contract APIs. All functions—including initialization, data insertion, update, query and deletion—were executed 100 times each, and the round-trip latency was measured at the client side via the built-in performance.timer() utility. The raw latency logs were then averaged to obtain the time-cost data presented in Table 1. The test results show that the execution time of the smart contract is within the acceptable range, which proves the efficiency of the system at the smart contract level..

Table 1 function Execution Time

Operation	Average execution time (μ s)
vehicle registration	502
Public key query	445
Vehicle cancellation	587

Concretely speaking, this scheme tests registration, public key query and vehicle revocation 200 times respectively for a given data set, and takes the mean value as the final result. The results show that the time cost of vehicle registration, public key query and vehicle revocation are 502 μ s, 445 μ s and 587 μ s respectively. Therefore, even if these operations are frequently invoked, the time cost is still acceptable, which proves the reliability and efficiency of the smart contract in this scheme.

3.3.2 Calculation overhead analysis

Computing overhead analysis focuses on evaluating computing efficiency in the IoT certification solution. This paper analyzes the computational costs of key operations such as pseudonym generation, message signature generation, and message signature verification. By conducting simulation experiments on laptops with certain configurations, the execution times of different encryption operations were collected and multiple tests were conducted to ensure the accuracy of the results.

In this paper, a notebook computer configured as shown in Table 2 is used to simulate the execution time of different encryption operations, and tested 200 rounds through the MIRACL/C++ encryption library, and the average value is used as the final simulation result. See Table 3 for details.

Table 2 Simulation Machine Configuration

configuration	specific parameters
memory	16GB
memory frequency	4800MHz
operating system	Windows 11
graphics card	NVIDIA GeForce RTX 3060
CPU	12th Gen Intel(R) Core(TM) i7-12700H

Table 3 Encryption Operation Execution Time

encryption operation	Average execution time (ms)
Tpa	0.0012
Tsm	0.1738
Tim	0.0245
Th	0.0003

Note: Tpa is the time taken to perform a point-add operation, Tsm is the time taken to perform a scalar multiplication operation, Tim is the time taken to perform a whole multiplication operation, Th is the time taken to perform a hash operation.

The total calculation overhead of the three stages of Scheme 1, Scheme 2, Scheme 3 and this scheme is 1.0467ms, 1.0458ms, 1.5669ms, and 0.8723ms respectively.

In order to intuitively compare the differences in computing costs between the proposed scheme in this work and the other three schemes in the certification process, Figure 1 and Figure 2 show the computing costs of the vehicle at different stages of the certification process and the entire stage respectively.

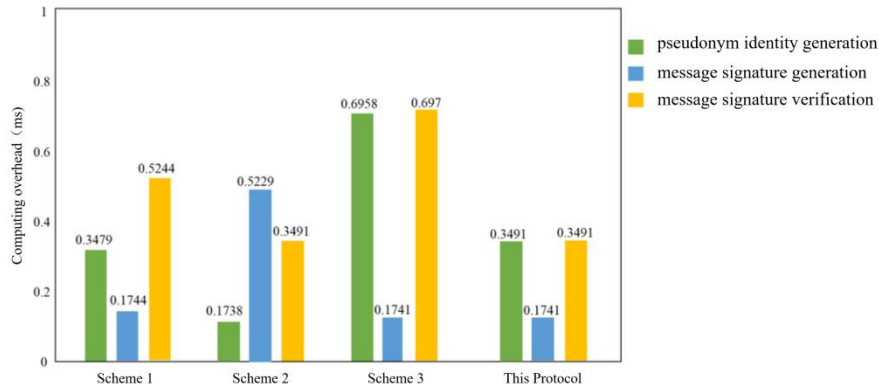


Figure 1 Computing Overhead at Different Stages of the Authentication Process

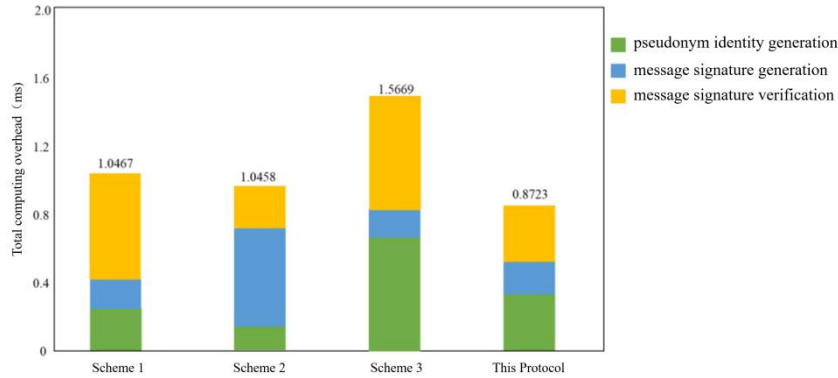


Figure 2 Total Computing Overhead of the Authentication Process

It can be seen from Figure 2 that this scheme is the lowest among all schemes for the total computing overhead in the authentication phase. Especially compared with previous works[10-12], this scheme reduces the total computing overhead by about 45%.

3.3.3 Communication overhead analysis

The communication overhead analysis takes into account the message size that the vehicle needs to send during the communication process. Parameters such as cyclic group size, infinite field size and timestamp size are defined, and the communication overhead of vehicle broadcast messages in each scheme is analyzed. The results show that the proposed scheme has advantages in terms of communication overhead, especially when sending vehicle messages in batches.

This scheme only considers the situation where the vehicle sends a single message for communication. The size of the parameters involved in the communication process is predefined as follows: the size of the cyclic group G is 40 bytes, the size of the infinite field z is 20 bytes, and the size of the timestamp is 4 bytes. The following is a detailed analysis of the communication overhead of vehicle-side broadcast messages in each scheme, as shown in Figure 3.

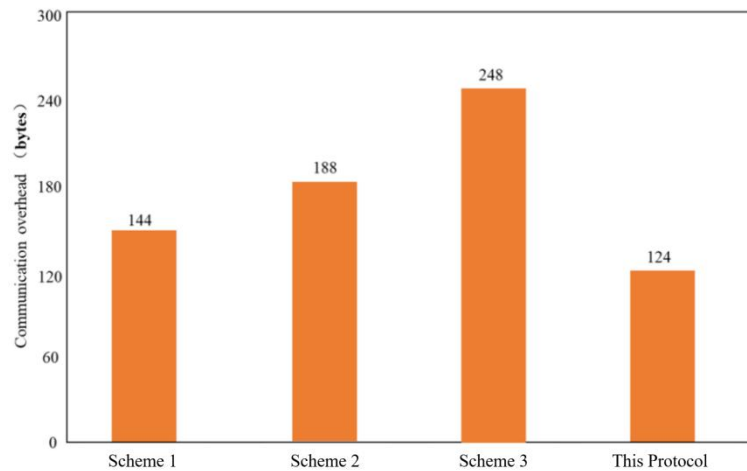


Figure 3 Comparison of Communication Overhead

It can be seen from Figure 3 that the communication overhead of this scheme is the lowest among all the comparison schemes. By well integrating the alliance chain with cryptographic primitives such as elliptic curves, even if there are security loopholes in the trusted party, BCPAS can still utilize acceptable communication overhead to achieve

conditional privacy protection. Compared with the schemes in documents that also adopt ECC technology and other documents, BCPAS reduces the communication overhead by 35% and 50% respectively. In particular, when the traditional authentication scheme requires the transmission of multiple messages and signatures to complete batch message broadcasting for vehicles, BCPAS only needs to transmit multiple messages and one signature at a time, thereby greatly reducing the occupation of channel resources and message transmission costs.

4 CONCLUSION

In view of the shortcomings of existing Internet of Vehicles authentication mechanisms in centralized dependence, key escrow, efficiency and privacy considerations, this paper proposes a conditional privacy protection authentication scheme with alliance chains and smart contracts as the core: lightweight signatures are implemented through elliptic curve cryptography, and distributed identity registration, dynamic revocation and pseudonym updates are completed with the help of the alliance chain's PBFT consensus and smart contracts, eliminating the single point of failure and key escrow risks in traditional PKI/IBS/CLS architecture. Experimental results show that the average time spent on vehicle registration, public key query and revocation operations of this scheme is less than 600 μ s, and the overall computing and communication overhead are reduced by about 45% and 35%-50% respectively compared with similar studies, meeting the needs of low-latency and resource-limited scenarios for the network of vehicles. Therefore, this work provides a new decentralized certification paradigm for the Internet of Vehicles with efficient, scalable and traceable support conditions, which can directly serve the safe communication of smart transportation systems.

Faced with the development of the Internet of Vehicles and the enhanced capabilities of attackers, future research can focus on two aspects: First, develop practical Internet of Vehicles certification solutions to reduce implementation costs, enhance decentralized attributes, improve security and privacy protection, and use blockchain decentralization. Features ensure data security, transparent recording and automatic execution of smart contracts, improving efficiency and data protection; The second is to explore lightweight authentication solutions to reduce computing, communication and storage pressures and improve system performance and operability after application of quantum cryptographic algorithms.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

- [1] Jabbar R, Dhib E, Said A B, et al. Blockchain technology for intelligent transportation systems: A systematic literature review. *IEEE Access*, 2022, 10: 20995 -21031.
- [2] He D, Zeadally S, Xu B, et al. An efficient identity-based conditional privacy-preserving authentication scheme for vehicular ad hoc networks. *IEEE Transactions on Information Forensics and Security*, 2015, 10(12): 2681-2691.
- [3] Azam F, Yadav S K, Priyadarshi N, et al. A comprehensive review of authentication schemes in vehicular ad-hoc network. *IEEE access*, 2021, 9: 31309-31321.
- [4] Chen H, Zhang C, Chen H. DAG blockchain-based PKI authentication fast response scheme. *Journal of Computer Engineering and Networks*, 2025, 12(45): 134-141.
- [5] Li C, Zhang X, Wang Y, et al. Identity-based chameleon hashes in the standard model for mobile devices. *IEEE Transactions on Information Forensics and Security*, 2025, 20(16): 145-153.
- [6] Li C, Zhou H, Liu Q, et al. Efficient multi-KGC certificateless signature scheme for cross-device authentication in internet of medical things. *Telecommunication Systems*, 2025, 88(2): 201-212.
- [7] Liu Xuejiao, Yin Yidan, Chen Wei, et al. A blockchain-based secure sharing solution for Internet of Vehicles . *Zhejiang University Journal (Engineering Edition)*, 2021, 55(5): 957-965.
- [8] Hu Lin. Performance analysis and resource allocation of vehicle active safety communication based on DSRC . *University of Electronic Science and Technology*, 2020.
- [9] Dong Jiankuo, Liu Zhe, Lu Sheng, et al. Research progress on efficient software implementation technology for elliptic curve cryptography . *Journal of Computer Science*, 2023, 46(05):909-928.
- [10] Chiang W K, Wang W Y. An efficient certificateless signcryption for mutual and batch authentication in vehicular ad-hoc networks. *Peer-to-Peer Networking and Applications*, 2025, 18(4): 321-330.
- [11] Ali I, Chen Y, Ullah N, et al. An efficient and provably secure ECC-based conditional privacy-preserving authentication for vehicle-to-vehicle communication in VANETs. *IEEE Transactions on Vehicular Technology*, 2021, 70(2): 1278-1291.
- [12] Sutrala A K, Bagga P, Das A K, et al. On the design of conditional privacy preserving batch verification-based authentication scheme for internet of vehicles deployment. *IEEE Transactions on Vehicular Technology*, 2020, 69(5): 5535-5548.

THE TECHNOLOGY OF AI-DRIVEN INTELLIGENT SYSTEM FOR STUDENTS' CAREER PLANNING

LiPing He

Department of Information Management, Guangdong Justice Police Vocational College, Guangzhou 510520, Guangdong, China.

Corresponding Email: gzlphe@126.com

Abstract: This paper proposes an AI-based intelligent system for students' career planning, which integrates technologies such as big data analysis, machine learning, and natural language processing to construct a "data-model-service" architecture. The system consists of three modules: data collection and preprocessing, intelligent analysis, planning generation, and interaction. By leveraging technologies including multi-source data fusion, the combination of reinforcement learning and knowledge graphs, and natural language generation, it realizes precise, personalized, and dynamic career planning services for students. Experiments show that this system increases the rationality score of career goals by 24.2%, improves the matching degree of first employment positions by 14.7%, and raises student satisfaction by 35.5%, providing a technical solution for the digital transformation of students' career planning education.

Keywords: Artificial intelligence; Students' career planning; Intelligent system; Big data analysis; Machine learning

1 INTRODUCTION

Driven by the in-depth restructuring of the global economic structure and the accelerated iteration of science and technology, the dynamics and complexity of the student employment market have reached an unprecedented level. Students' career planning has become a systematic project that integrates multiple dimensions, including their own interests, ability structures, value orientations, market demand characteristics, and industry development trends. However, the traditional model of students' career planning is limited by the shortcomings of static assessment tools and the subjectivity of experience-based guidance. It can hardly meet the personalized and differentiated development needs of students, nor can it respond promptly to the rapidly changing employment ecosystem[1].

Artificial intelligence (AI) technology provides an innovative solution to address issues such as difficulties in data integration and insufficient personalization in traditional career planning services. By virtue of distributed computing and deep learning algorithms, it conducts cross-dimensional analysis of students' behavioral data and dynamic employment market data, and constructs a two-way mapping model to achieve accurate matching. Specifically, it parses job information through natural language processing, builds career paths using knowledge graphs, and dynamically optimizes planning strategies with the help of reinforcement learning[2]. This paper focuses on constructing a closed-loop system of "data collection-intelligent analysis-dynamic planning-interactive service", integrating cutting-edge technologies to achieve an upgrade from static recommendation to dynamic evolution, and providing a technical path for improving the intelligence level of students' career planning services[3].

2 RELEVANT TECHNICAL FOUNDATIONS

2.1 Big Data Analysis Technology

Big data analysis technology has significant advantages in the field of students' career planning. Its data system integrates on-campus academic trajectory data (such as structured and semi-structured data including GPA, course selection, club participation, and internship records) and external career ecosystem data (such as unstructured and structured data including industry white papers, job demand matrices, and dynamic career development graphs), forming a multi-dimensional and three-dimensional data network[4]. This network can not only depict students' knowledge reserves, ability evolution, and interest tendencies in detail but also fully reflect the skill requirements, industry trends, and career paths of the employment market.

Through distributed computing frameworks such as Hadoop and Spark, efficient cleaning, transformation, and integration of multi-source heterogeneous data scattered in different systems and formats can be achieved, effectively solving the problem of data silos[5]. At the same time, with the support of distributed storage and parallel computing capabilities, the processing efficiency of large-scale data is significantly improved, providing high-quality data support for the subsequent intelligent analysis and precise services of career planning.

In the data mining process, the Apriori algorithm is used to mine strong association rules between course grades and career success. This algorithm is based on the apriori property and mines association relationships through layer-by-layer iterative search. In practical applications, course grade data is first discretized into transaction datasets, and then thresholds for support ($\geq 20\%$) and confidence ($\geq 75\%$) are set to screen frequent itemsets and generate strong

association rules. This algorithm can effectively reveal the potential associations and patterns between data, construct a mapping relationship model between students' individual characteristics and career requirements, and provide quantitative support for the subsequent generation of students' career planning[6].

2.2 Machine Learning Technology

As the core support of the AI technology system, machine learning endows the system with the ability to extract insights from historical data and make forward-looking predictions by constructing data-driven adaptive models[7]. In the intelligent system for students' career planning:

- The supervised learning module adopts the Gradient Boosting Decision Tree (GBDT) algorithm to build a student career prediction model. It uses GOSS (Gradient-based One-Side Sampling) gradient sampling technology to handle the high-dimensional feature space (with approximately 1000+ dimensions), achieving accurate classification of students' career tendencies.
 - The unsupervised learning component integrates Kernel Principal Component Analysis (KPCA) and t-distributed Stochastic Neighbor Embedding (t-SNE) dimensionality reduction algorithms to map high-dimensional career features to a low-dimensional manifold space, revealing the implicit semantic association structure in career data.
 - The reinforcement learning framework constructs a dynamic decision-making model based on the Deep Q-Network (DQN) architecture. By designing a multi-objective reward function that includes career matching degree, student satisfaction, and feasibility evaluation, it realizes the adaptive optimization of students' planning suggestions.
- The entire machine learning system adopts a dual-cycle mechanism of offline training and online fine-tuning, continuously improving the prediction accuracy of students' career development trajectories and the quality of planning suggestions through dynamic adaptation[8].

2.3 Natural Language Processing Technology

As the core technology for processing unstructured text data, natural language processing (NLP) realizes the intelligent parsing of text data such as recruitment information and industry dynamics in the context of career planning by constructing semantic understanding models. At the technical application level:

- The text classification module combines the BERT-base pre-trained model with a fully connected classification layer, achieving an F1-score of over 92% in the recruitment information classification task, which can efficiently distinguish different job requirements such as technical and management positions[9].
- The named entity recognition component is based on the BiLSTM-CRF network structure, integrating character vectors and word vectors, and achieves an accuracy rate of over 95% in recognizing entities such as job titles and skill keywords, providing structured semantic units for subsequent intelligent analysis[10].
- The sentiment analysis module uses the TextCNN model combined with the VADER sentiment lexicon, achieving an accuracy rate of 88% in judging the sentiment polarity of industry reviews, which can quantitatively evaluate the social recognition of occupations[9].
- The machine translation module adopts a Neural Machine Translation (NMT) architecture, which is fine-tuned based on the public WMT dataset and supports multi-language translation such as Chinese-English and Chinese-Japanese, helping students obtain global career information[11].

The entire NLP system improves the semantic understanding accuracy and information extraction efficiency of text in the career field through end-to-end training and domain adaptation optimization.

3 DESIGN OF THE INTELLIGENT SYSTEM ARCHITECTURE

3.1 Data Collection and Preprocessing Module

This module is responsible for the collection, integration, and standardized processing of multi-source heterogeneous data, and provides high-quality data support for subsequent intelligent analysis by constructing a full-link data governance system. On-campus data is connected in real-time with the school's educational administration system and student management system through API interfaces, covering a wide range of data including students' basic information, detailed course grades, reward and punishment records, and practical activity trajectories. External data is incrementally collected from channels such as recruitment platforms, industry research databases, and government open data portals relying on a distributed crawler framework, with a daily data collection volume of more than 100,000 pieces.

To address issues such as inconsistent formats, missing values, and noise pollution in the collected data, a multi-level data cleaning process is constructed. Multiple imputation and random forest regression models are combined to handle missing values, and isolation forests and Local Outlier Factor (LOF) algorithms are used for joint detection of noisy data[12]. In the data transformation stage, a combined strategy of LabelEncoder and One-Hot Encoder is adopted to realize the vectorization encoding of categorical features, and the TF-IDF algorithm is used to weight text data[13]. Finally, the standardized data is stored in a distributed data warehouse through the ETL (Extract-Transform-Load) process, forming a structured data asset pool and providing highly available data input for the intelligent analysis module[14].

3.2 Intelligent Analysis Module

As the core of the system, the intelligent analysis module integrates big data analysis, machine learning, and natural language processing technologies to build a deep analysis engine. Firstly, it conducts multi-dimensional analysis of students' on-campus data through a distributed computing framework, and uses graph neural networks to construct a three-dimensional student profile including learning ability indicators, interest preference graphs, and practical ability matrices. At the same time, based on the time-series analysis model of industry data, it extracts dynamic features such as industry growth prediction curves, job demand heat maps, and salary distribution quantiles.

In the matching decision stage, a hybrid recommendation algorithm architecture is adopted:

- A supervised learning model based on LightGBM predicts the career adaptation probability[15].
- A graph collaborative filtering algorithm is used to mine the career paths of similar students.
- An attention mechanism is employed to fuse the output weights of the two types of models.

The natural language processing component uses BERT-whitening semantic encoding technology to convert the skill requirements in recruitment information into ability gap indicators in the vector space, and calculates the cosine similarity with the students' skill graphs to generate a personalized improvement plan that includes technical stack gaps and certificate acquisition priorities[16]. The entire analysis process is continuously optimized through an incremental learning mechanism. When new data triggers the threshold, the model parameter fine-tuning process is automatically initiated[17].

3.3 Planning Generation and Interaction Module

After the intelligent analysis module outputs the student profile and career matching results, the planning generation module generates a full-cycle personalized career planning plan for students by constructing a hierarchical planning model. This plan adopts a three-dimensional architecture design:

- The short-term academic improvement layer uses a course association rule algorithm to generate course selection suggestions, and recommends certification exams based on the mapping matrix between certificates and career requirements.[18]
- The medium-term practical planning layer formulates an internship plan based on the internship position semantic matching model, and recommends practical projects through the correlation analysis of competitions and skill improvement[19].
- The long-term development path layer uses a career transition probability graph model (built based on GNN) to generate multi-branch career development trajectories[20].

The interaction module constructs a multi-modal human-computer interaction system, integrating natural language processing and visualization technologies, and supports full-scenario coverage of the Web terminal (React+Redux) and mobile terminal (Flutter). The system adopts a hybrid intent recognition model of "FastText + rule engine":

- First, it recognizes more than 20 types of core intents through 300-dimensional semantic vectorization (with an accuracy rate of 92.3%)[21].
- Then, it combines the career knowledge graph to increase the accuracy to 96.7%[22].
- An LSTM module is used to maintain 10 rounds of dialogue history to update user demand labels.

The planning plan is presented in a timeline + skill tree visualization format. The ECharts Gantt chart decomposes the planning stages, with detailed information cards attached to the nodes. The timeline supports weekly granularity adjustment. When the target node is dragged, the PPO algorithm is triggered to dynamically recalculate, and the 8-layer neural network is used to realize the plan regeneration within 100ms and update the weights of the visualization nodes[23].

4 IMPLEMENTATION OF KEY TECHNOLOGIES

4.1 Student Profile Construction Technology

The construction of student profiles adopts multi-dimensional data fusion technology, and three-dimensionally depicts students' abilities through a three-layer feature extraction system:

- In the academic performance dimension, time-series analysis is used to dynamically model course grades, calculate statistical indicators such as average scores and standard deviations, and analyze score trends and learning ability progression by combining sliding window technology and educational data mining algorithms.
- In the interest dimension, a ternary association network of club participation, course selection, and competition investment is constructed, and graph embedding technology and vector space analysis are used to identify interest fields and potential development directions.
- In the practical ability dimension, a practical ability evaluation matrix is built based on internship positions, project results, and skill certificates. The TF-IDF algorithm is used to calculate the importance of various practical skills, and the students' practical application and professional skill levels are comprehensively evaluated [7].

In the dimensionality reduction process, Principal Component Analysis (PCA) is used to optimize high-dimensional features. By solving the eigenvalues and eigenvectors of the data covariance matrix, the principal feature dimensions that retain more than 85% of the information are selected, which removes redundancy while retaining key features, realizing efficient dimensionality reduction of student feature data. In the clustering analysis stage, the DBSCAN density clustering algorithm is used, with scientific setting of neighborhood and sample thresholds. Based on the spatial distribution density and correlation of data points, students with similar features are divided into different groups,

providing an accurate group division basis for the subsequent personalized career planning of students. This systematic student profile construction process can comprehensively and accurately depict the comprehensive characteristics of students, laying a data foundation and analysis support for the generation of students' career planning suggestions.

4.2 Career Matching Algorithm

The career matching algorithm realizes accurate person-job matching by constructing a two-way mapping model between students' features and career requirements. The content-based recommendation framework converts students' skill graphs, interest preferences, and ability evaluations into high-dimensional student feature vectors, and at the same time constructs career demand vectors based on the skill requirements and job content characteristics of occupations. The matching degree between the two is quantified through cosine similarity calculation, and the initial candidate career set is screened out.

To improve the matching accuracy, the algorithm introduces domain knowledge to construct a weight matrix, adjusting the feature weights according to different occupation types. For positions with prominent professional skill requirements, the weight proportion of relevant features such as learning ability and industry certification is automatically increased. In the secondary optimization stage, combined with the logistic regression classification model, historical student career selection data is used to train the model to learn the implicit patterns between students' features and career matching. The matching probability of the initial matching results is calculated and sorted, and finally, a personalized career recommendation list sorted by matching degree from high to low is generated.

4.3 Dynamic Planning Adjustment Technology

Considering the dynamic changes in students' growth and the employment market, students' career planning can build an adaptive planning system with the help of reinforcement learning technology. This system defines students' actions as state transitions, quantifies action feedback as reward values, and optimizes strategies using algorithms such as Q-learning by constructing a "state-action-reward" triplet sequence to explore the optimal career planning path for students. The system maintains a state space containing multi-dimensional features such as skill graphs, internship experiences, and academic scores for each student. When students implement the planning suggestions, the system gives positive/negative rewards based on feedback, updates the Q-value through the Bellman equation, and adjusts the priority of suggestion recommendations. At the same time, the system regularly captures external data such as industry reports and recruitment requirements. When it detects that changes in industry demands or skill requirements exceed the threshold, it triggers the re-optimization of the planning plan, dynamically adjusts career goals and skill improvement plans, and ensures that the planning plan is synchronized with the dynamics of the employment market.

5 EXPERIMENTS AND VERIFICATION

5.1 Experimental Design

Students majoring in Computer Science and Technology and Electronic Information Engineering from the same grade of a university were selected as the experimental subjects, and divided into an experimental group and a control group. The experimental group adopted the intelligent career planning system proposed in this paper, while the control group used the traditional career planning guidance method. The experimental cycle was one academic year. Before the experiment, the basic information of all students was collected and a pre-test of career planning (covering interests, abilities, career cognition, etc.) was conducted. After the experiment, a post-test was carried out to evaluate the accuracy of career planning and student satisfaction. Among them, the accuracy of career planning was judged by combining expert evaluation and actual employment tracking, and student satisfaction was obtained through a questionnaire survey (Table 1-6).

Table 1 Statistics of Basic Information of Experimental Subjects

Field Name	Experimental Group (n=100)	Control Group (n=100)
Major Distribution	52 students in Computer Science and Technology / 48 students in Electronic Information Engineering	50 students in Computer Science and Technology / 50 students in Electronic Information Engineering
Gender Ratio	78 males / 22 females	80 males / 20 females
Admission Score	Average score: 82.5±5.3	Average score: 81.8±6.1
Family	65% from science and engineering families	62% from science and engineering families

Field Name	Experimental Group (n=100)	Control Group (n=100)
Background		

Table 2 Mean Values of Pre-test Indicators for Career Planning

First-Level Indicator	Second-Level Indicator	Experimental Group	Control Group
Interest Characteristics	Consistency of Holland Career Interest Code	0.68±0.12	0.65±0.14
	Clarity of Top 3 Interest Fields	3.2±0.8	3.1±0.9
Ability Characteristics	Programming Ability Score	3.5±0.7	3.4±0.6
	Problem-Solving Ability	3.8±0.6	3.7±0.5
Career Cognition	Clarity of Target Career	2.9±0.7	2.8±0.8
	Correct Rate of Industry Development Cognition	61.2%±10.5%	59.8%±12.3%

Table 3 Statistics of Intervention Measures During the Experiment

Field Name	Experimental Group	Control Group
Intervention Frequency	9 system updates + 3 interviews	2 interviews
Number of Dynamic Adjustments to Skill Improvement Suggestions	Average 4.2 times per person	Average 1.1 times per person
Number of Internship Recommendations	Average 5.3 times per person	Average 3.1 times per person

Table 4 Post-test of Career Planning - Results of Accuracy Evaluation

Evaluation Dimension	Specific Indicator	Experimental Group	Control Group
Expert Evaluation	Rationality of Career Goals	4.1±0.5	3.3±0.6
	Skill Matching Degree	4.3±0.4	3.5±0.7
Employment Tracking	Matching Degree of First Employment Position	75.2%±8.3%	60.5%±11.2%
	Starting Salary (RMB)	8560±1200	7230±1500
	Employment Satisfaction	4.2±0.6	3.5±0.8

Table 5 Post-test of Career Planning - Mean Values of Student Satisfaction

Satisfaction Dimension	Survey Question	Experimental Group	Control Group
Planning Science	Relevance of Course Recommendations	4.1±0.5	3.2±0.7
	Persuasiveness of Career Goals	4.0±0.6	3.1±0.8
Service Experience	Convenience of System Interaction	4.3±0.4	3.0±0.9
	Response Speed of Planning Adjustments	4.2±0.5	2.8±0.7
Overall Evaluation	Overall Satisfaction	4.2±0.5	3.1±0.8

Table 6 Data Comparison and Significance Test Between Experimental Group and Control Group

Comparison Indicator	Experimental Group	Control Group	t-value	p-value
Pre-test - Clarity of Interests	3.2±0.8	3.1±0.9	0.87	0.385
Pre-test - Career Cognition	2.9±0.7	2.8±0.8	0.76	0.449
Post-test - Rationality of Goals	4.1±0.5	3.3±0.6	9.82	<0.001
Post-test - Job Matching Degree	75.2%±8.3%	60.5%±11.2%	9.73	<0.001
Post-test - Satisfaction Score	4.2±0.5	3.1±0.8	10.21	<0.001

5.2 Analysis of Experimental Results

The experimental data shows that there is no significant difference in the pre-test indicators between the experimental group and the control group, indicating that the two groups are scientifically comparable. In terms of the accuracy of career planning, expert evaluation shows that the experimental group has a career goal rationality score of 4.1 points (3.3 points for the control group) and a skill matching degree of 4.3 points (3.5 points for the control group). Combined with employment tracking, it is found that the matching degree of the first employment position of the experimental group is 75.2% (60.5% for the control group), and the starting salary is 18.4% higher than that of the control group. These results confirm that the intelligent system can accurately connect students' characteristics with career requirements.

The student satisfaction survey shows that the experimental group has significantly higher scores than the control group (3.1-3.2 points) in dimensions such as relevance of course recommendations (4.1 points), persuasiveness of career goals (4.0 points), and convenience of system interaction (4.3 points), with an overall satisfaction of 4.2 points (3.1 points for the control group). This reflects students' recognition of the intelligent planning service.

Multi-dimensional analysis consistently verifies the effectiveness and practicality of the intelligent career planning system. It has obvious advantages in improving the quality of students' career planning and student satisfaction, and can provide more effective career planning support for students.

6 CONCLUSIONS AND PROSPECTS

This study constructs an AI-driven intelligent system for students' career planning, which deeply integrates technologies such as big data analysis and machine learning to realize accurate matching, personalized recommendation, and dynamic adjustment of career planning. Experimental data shows that the system can effectively improve the accuracy of career planning and significantly enhance student satisfaction, providing an innovative solution for career planning education.

However, this study also has certain limitations. For example, data collection faces great difficulties, and the algorithm model needs further optimization. In future research, we will conduct in-depth work in expanding data sources, optimizing algorithm models, and strengthening school-enterprise cooperation. We will continue to promote the

application of AI technology in the field of students' career planning to help students scientifically plan their career development paths.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

This research was funded by the Special Research Project on Digital Transformation of Higher Education Institutions, 2025 (Grant No. GJX25Z2162), the "14th Five-Year Plan" Research Project of Guangdong Association of Higher Education, 2024 (Grant No. 24GWB159), and the "14th Five-Year Plan" Project of the Education Branch of China Information Association, 2025 (Grant No. ZXXJ20250028).

REFERENCES

- [1] Feng K. Research on College Students' Career Planning and Employment Guidance Strategies Assisted by Artificial Intelligence. *Applied Mathematics and Nonlinear Sciences*, 2024, 9(1). <http://dx.doi.org/10.2478/amns-2024-3134>.
- [2] Huang Y. Individualized career planning intelligent teaching system for college students based on adaptive artificial intelligence. *Journal of Computational Methods in Sciences and Engineering*, 2025, 25(4): 3121–3136. <http://dx.doi.org/10.1177/14727978251315731>.
- [3] Zhang Y. Path of career planning and employment strategy based on deep learning in the information age. Ali K, editor. *PLoS ONE*, 2024, 19(10): e0308654. <http://dx.doi.org/10.1371/journal.pone.0308654>.
- [4] Xiong X, Yao J, Guo M. Research and Development of Education Informatization Platform Based on Multi-dimensional Education Big Data Collection and Analysis Algorithm. In: *Proceedings of the International Conference on Modeling, Natural Language Processing and Machine Learning*. ACM, 2024: 188–192. <http://dx.doi.org/10.1145/3677779.3677810>.
- [5] Song Z. Design and Implementation of a Distributed Technology-Based Big Data Analysis System. In: *Proceedings of the 2024 5th International Conference on Big Data Economy and Information Management*. ACM, 2024: 56–60. <http://dx.doi.org/10.1145/3724154.3724164>.
- [6] Hao Q, Choi WJ, Meng J. A data mining-based analysis of cognitive intervention for college students' sports health using Apriori algorithm. *Soft Computing*, 2023, 27(21): 16353–16371. <http://dx.doi.org/10.1007/s00500-023-09163-z>.
- [7] Ahamad R, Mishra KN. Enhancing knowledge discovery and management through intelligent computing methods: a decisive investigation. *Knowledge and Information Systems*. 2024, 66(7): 3719–3771. <http://dx.doi.org/10.1007/s10115-024-02099-2>.
- [8] Guo P, Xiao K, Ye Z, Zhu H, Zhu W. Intelligent career planning via stochastic subsampling reinforcement learning. *Scientific Reports*. 2022 May 18;12(1). Available from: <http://dx.doi.org/10.1038/s41598-022-11872-8>.
- [9] Dai L, Mao J, Xu L, et al. SecNLP: An NLP classification model watermarking framework based on multi-task learning. *Computer Speech & Language*, 2024, 86: 101606. <http://dx.doi.org/10.1016/j.csl.2023.101606>.
- [10] Qiu Q, Tian M, Huang Z, et al. Chinese engineering geological named entity recognition by fusing multi-features and data enhancement using deep learning. *Expert Systems with Applications*, 2024, 238: 121925. <http://dx.doi.org/10.1016/j.eswa.2023.121925>.
- [11] Ahammad SH, Kalangi RR, Nagendram S, et al. Improved neural machine translation using Natural Language Processing (NLP). *Multimedia Tools and Applications*, 2023, 83(13): 39335–39348. <http://dx.doi.org/10.1007/s11042-023-17207-7>.
- [12] Iranna Shirol. Adaptive Federated Data Cleaning with Explainability: A Basic Threshold-Driven Approach for Heterogeneous Data Environments. *International Journal of Scientific Research in Science, Engineering and Technology*, 2025, 12(2): 828–832. <http://dx.doi.org/10.32628/ijrsret25122207>.
- [13] Alqahtani AF, Ilyas M. An Ensemble-Based Multi-Classification Machine Learning Classifiers Approach to Detect Multiple Classes of Cyberbullying. *Machine Learning and Knowledge Extraction*, 2024, 6(1): 156–170. <http://dx.doi.org/10.3390/make6010009>.
- [14] Mandala V, Sougata B. ETL - A Data Cleansing Approach. *Global Research and Development Journals*, 2024, 9(8): 55–59. <http://dx.doi.org/10.70179/grdjev09i080012>.
- [15] He M, Qian Q, Liu X, et al. Recent Progress on Surface Water Quality Models Utilizing Machine Learning Techniques. *Water*, 2024, 16(24): 3616. <http://dx.doi.org/10.3390/w16243616>.
- [16] Karhale ProfS. Resume Analysis System Using Natural Language Processing. *IJSREM*, 2025, 9(6): 1–9. <http://dx.doi.org/10.55041/ijrsrem49872>.
- [17] Salwala D, Tirupathi S, Quanz B, et al. Distributed Incremental Machine Learning for Big Time Series Data. In: *2022 IEEE International Conference on Big Data (Big Data)*. IEEE, 2022: 2356–63. <http://dx.doi.org/10.1109/BigData55660.2022.10020361>.
- [18] Liu W. Career planning and pathway generation based on multimodal learning analytics. *International Journal of Information and Communication Technology*, 2025, 26(29): 39–58. <http://dx.doi.org/10.1504/IJICT.2025.147880>.

- [19] Li Y. Research on Intelligent Recommendation Algorithm Model for Vocational Talent Training. In: 2025 IEEE International Conference on Electronics, Energy Systems and Power Engineering (EESPE). IEEE, 2025: 1290–1294. <http://dx.doi.org/10.1109/EESPE63401.2025.10987159>.
- [20] Putrama IM, Martinek P. Integrating platforms through content-based graph representation learning. *International Journal of Information Management Data Insights*, 2023, 3(2): 100200. <http://dx.doi.org/10.1016/j.jjime.2023.100200>.
- [21] Madhukar Jukanti. Understanding Natural Language Processing (NLP) Techniques. *JCSTS*, 2025, 7(7): 1005–1012. <http://dx.doi.org/10.32996/jcsts.2025.7.7.112>.
- [22] Ibrahim N, Aboulela S, Ibrahim A, et al. A survey on augmenting knowledge graphs (KGs) with large language models (LLMs): models, evaluation metrics, benchmarks, and challenges. *Discover Artificial Intelligence*, 2024, 4(1). <http://dx.doi.org/10.1007/s44163-024-00175-8>.
- [23] Ma B, Xu Y, Pan Y, et al. A multi-user mobile edge computing task offloading and trajectory management based on proximal policy optimization. *Peer-to-Peer Networking and Applications*, 2024, 17(6): 4210–29. <http://dx.doi.org/10.1007/s12083-024-01796-7>.

WIND-SOLAR HYBRID DRIVE-CONTROLLED AUTONOMOUS SHIP FOR EXPLORATION BASED ON RASPDERRY PI

GuoLong Yang*, JunYu Huang, NaiLi Zhang, Jian Zhang
Zhanjiang Normal College for Children, Zhanjiang 524084, Guangdong, China.
Corresponding Author: GuoLong Yang, Email: 419415586@qq.com

Abstract: In order to facilitate the monitoring of water area data in various regions, enhance the efficiency of environmental monitoring, and respond to the national goals of carbon peaking and carbon neutrality, this project has developed an intelligent unmanned detection vessel driven by a hybrid of wind and solar power. This project adopts Raspberry PI 4b motherboards. By designing the hull of the unmanned vessel and integrating a remote control system and an intelligent detection system, it is equipped with high-precision GPS positioning to control the unmanned vessel to perform monitoring tasks in specific waters. By using video monitoring and various sensors to collect data, and after analysis, the test information is transmitted in real time to the user's mobile phone APP or computer. The collected monitoring data can be synchronized to the user's terminal device in real time to ensure the timeliness and accuracy of the information. This system has high stability and reliability in actual operation. Meanwhile, the unmanned vessel can generate electricity by using the wind-solar drive module to achieve its endurance. This research has certain reference value.

Keywords: Unmanned vessel; Raspberry Pi; Hybrid power drive; Water quality monitoring

1 INTRODUCTION

Driven by the "dual carbon" goals implemented in our country and the demand for intelligent water area monitoring, unmanned vessels are transforming from "traditional fuel-driven + manual remote control" to "new energy autonomous drive + mission-oriented intelligent operation"[1]. Especially in small and medium-sized water area monitoring scenarios, the requirements for the equipment's endurance, environmental friendliness, and multi-parameter perception capabilities have significantly increased, pushing unmanned vessels to accelerate their iteration towards the "low energy consumption - high collaboration" direction[2,3]. As surface intelligent equipment capable of autonomously completing water area data collection and environmental monitoring, its design needs to take into account the particularities of the operation scenarios - such as the strict requirements for endurance stability and anti-water interference capabilities[4]. Abroad, the United States initiated the research on unmanned ships as early as the 1930s. Its "Spartan Scout" and "X-2" and other equipment were mainly focused on military reconnaissance. Their power systems were mainly powered by diesel or lithium batteries, which had limitations such as limited endurance (single operation < 8 hours) and high carbon emissions[5]. The Japanese Yamaha "UMV-H" and "UMV-O" also expanded to marine exploration, but they relied on a single battery for power supply and had insufficient stability in cloudy or low wind speed environments[6]. In China, the research on unmanned ships driven by policies has achieved breakthroughs (such as the "Zhuhai Cloud" research vessel and the "Jinghai" series), but most of the equipment still relied on traditional power sources or single solar power. They have shortcomings in endurance under complex weather conditions (such as continuous cloudy days)[7]. Based on the current technical status both domestically and internationally, the core competitiveness of unmanned ships is shifting from "function realization" to "energy and environmental compatibility"[8-10]. In response to the shortcomings of the above power system, this project has developed an intelligent detection unmanned ship with a hybrid power system of wind and solar energy. The core innovation mainly lies in three aspects: 1. Energy system: Using monocrystalline silicon solar panels (with a conversion efficiency of 22%) and vertical-axis wind turbines (with a startup wind speed of $\leq 1.8\text{m/s}$), through an MPPT intelligent controller to achieve coordinated management of wind and solar energy, maintaining 80% rated power output even in cloudy or low-light conditions, which increases the endurance by 50% compared to single solar energy drive; 2. Operation adaptability: Optimizing the streamlined hull (with a resistance coefficient reduced by 15%) and double-layer waterproof structure (IP68 level sealing), adapting to complex water flow environments in small waters such as lakes and rivers; 3. Task-oriented: Centered on water quality monitoring, integrating multi-sensor collaborative sampling and real-time data transmission, achieving "energy autonomy + monitoring autonomy" in a dual closed-loop control[11,12].

2 TOTAL DESIGN SCHEME

This project developed an unmanned ship based on Raspberry Pi, which is a multi-functional unmanned vessel. Various sensors are utilized to obtain water quality information. The high-definition camera module onboard enables real-time monitoring and acquisition of image data. It communicates with the end users through the lightweight MQTT communication protocol for data exchange, and supports remote status inquiries. Additionally, in combination with the accompanying traction device, the unmanned ship can be extended to scenarios such as maritime rescue, garbage

cleaning, and the transportation of small supplies by water. Its working process is as follows: After the unmanned ship enters the water area to be tested, it can start the water quality sensor for water quality detection. The test data obtained is processed by Raspberry Pi, and the analysis results are sent to the cloud server through the MQTT protocol. The server parses the results and then transmits them to the user terminal. Users can obtain real-time water quality data through the mobile App at any time, thereby achieving the purpose of monitoring various parameters of water quality. The workflow diagram is shown in Figure 1.

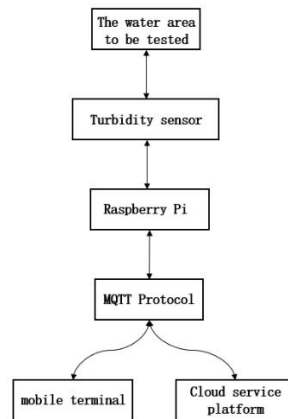


Figure 1 Workflow of the Unmanned Ship

3 HARDWARE DESIGN OF UNMANNED VESSELS

The hardware of the unmanned ship developed in this project is controlled based on the Raspberry Pi development board. It integrates a driving system, a camera module, and modules suitable for monitoring various water quality parameters (such as temperature sensors, turbidity sensors, and pH value sensors, etc.). The hardware architecture design is shown in Figure 2. Each module works collaboratively with the main control board through interfaces to jointly support the monitoring and operation functions of the unmanned ship.

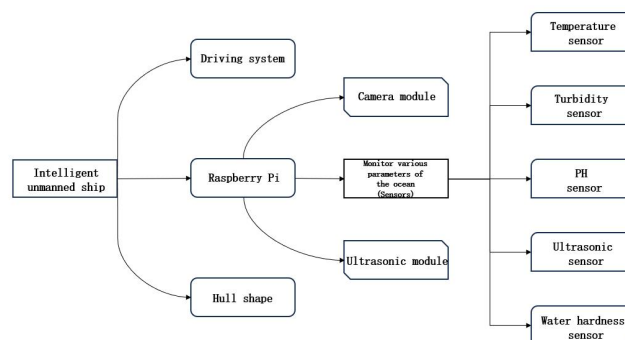


Figure 2 Hardware Architecture Design of Unmanned Ship

According to the overall structure design shown in Figure 3, the operation control process of the unmanned ship is as follows: After the user issues an operation instruction through the mobile phone APP, the Raspberry Pi main control board receives the instructions parsed by the MQTT protocol and immediately transmits control signals to the driving system, causing the motor to operate and achieving the regulation of the navigation status. The Raspberry Pi main control board establishes data interaction with various water quality sensors through expansion interfaces and collects real-time data parameters such as turbidity and pH value. These detection data are preliminarily processed and uploaded through the MQTT protocol, and after completing the analysis of the water quality status, they are pushed to the cloud server, which then sends the data to the mobile phone APP. The camera module is responsible for capturing the images of the water surface and its surroundings; the GPS module acquires the spatial coordinates of the unmanned ship in real time, providing accurate location information for remote control, and at the same time, in conjunction with the camera module, realizes the spatial positioning and image association of the operation scene.

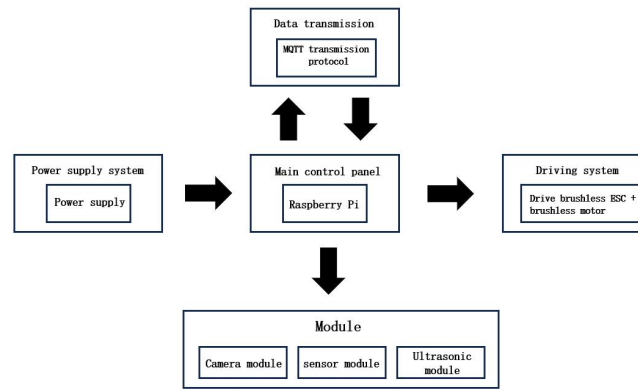


Figure 3 Overall Structure Design of the Unmanned Ship

3.1 GPS Positioning Module

This project uses the GPS module of the Yabo intelligent model ATGM336H-5N. The working process is as follows: Connect the USB interface on the Raspberry Pi motherboard to the Micro interface of the GPS module. Its function is to output navigation data and input interactive commands. The GPS module usually outputs second pulses and is connected to an antenna, and the antenna is installed on the ship deck. According to Figure 3, the GPS module obtains the position information of the unmanned ship, and the Raspberry Pi motherboard acquires the relevant data and processes it. Then, it uploads it to the cloud server via the MQTT protocol. When receiving data, the cloud service platform receives the relevant data and transmits it to the Raspberry Pi motherboard for data processing.

3.2 Data Transmission

The data transmission of this project is implemented using Python on the Raspberry Pi to connect with a regular MQTT server, and it performs message publishing and subscription in dual threads. The unmanned ship uses the paho-mqtt library in Python to complete the communication tasks under the MQTT protocol. MQTT is a lightweight messaging protocol, suitable for situations with limited network bandwidth. By creating an MQTT client on the Raspberry Pi using Python, message publishing and subscription can be achieved.

Based on the operational characteristics of unmanned vessels on open waters, their signal transmission is unobstructed and has a wide coverage area. Their data transmission adopts a hierarchical network scheme: relying on mobile phone 5G/4G hotspots to establish the basic connection, which can support data interaction over medium to long distances on the water surface; at the same time, the Wi-Fi module installed on the Raspberry Pi motherboard has 360° omnidirectional coverage capability, which can provide unified network support for multiple unmanned vessels conducting collaborative operations, improving the efficiency of cluster operations while expanding the monitoring coverage area. Compared to the configuration of LTE modules, using Wi-Fi modules has three advantages: first, the transmission bandwidth is larger, which can meet the real-time transmission of large-capacity data such as high-definition images; second, the power consumption is lower, saving about 20% compared to LTE modules, which helps to reduce the energy consumption of the entire vessel; third, the connection stability is stronger, reducing data packet loss in complex water environment. When the navigation distance of the unmanned vessel exceeds the coverage radius of Wi-Fi, the system automatically switches to the mobile data module to maintain remote communication, ensuring stable data transmission for the unmanned vessel at longer distances or in complex water areas[13].

3.3 Dving System

This project adopts a drive system that uses two electric motors as the propulsion units to precisely control the ship's navigation status. The specific connection method is as follows: Connect the brushless power regulator with a rated current of 30A to the wires of the ROV brushless motor. The pin 12 on the main board is connected to the yellow signal of the power regulator to control the working state of the brushless motor. Finally, the positive pole of the power supply is connected to the red power line of the power regulator, and the negative pole of the power supply and the GND of the main board, as well as the black power line of the power regulator, are connected to provide normal power supply for the brushless motor (Figure 4 Connection diagram of the drive system).

After the control terminal sends the operation instructions, the Raspberry Pi motherboard acquires the relevant information and outputs the PWM signal. By adjusting the PWM pulse width, it controls the motor driver, thereby controlling the rotational speed of the brushless motor. This control system uses the rotational speed difference between the two motors to form the steering torque to complete the course adjustment, ensuring the stability of the ship's movement in complex waters.

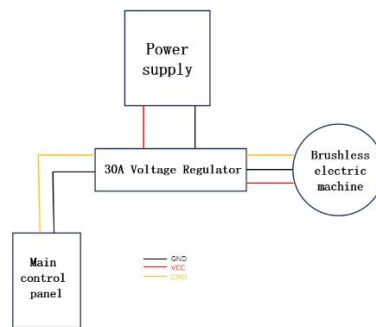


Figure 4 Connection Diagram of the Drive System

3.4 Camera Module

This project uses the camera module from Yabo Intelligent Company. The camera adopts a USB2.0 direct plug-in interface. Connect the USB2.0 interface of the camera to the USB interface of the Raspberry Pi motherboard. Thus, the camera connection is completed.

3.5 Ultrasonic Module

This project uses the HC-SR04 ultrasonic module. The VCC of the ultrasonic module is connected to the 5V pin of the Raspberry Pi motherboard, and the Trig and Echo pins of the ultrasonic module are connected to the GPIO ports of the Raspberry Pi. Finally, the GND of the ultrasonic module is connected to the GND of the Raspberry Pi motherboard. This completes the circuit connection of the ultrasonic module. The HC-SR04 ultrasonic module uses IO triggering for distance measurement. It sends a high-level signal of at least 10us. The module will automatically send 8 40kHz ultrasonic wave pulses and simultaneously start the echo detection mechanism to determine the existence of the reflected signal. When the Raspberry Pi motherboard receives the high-level echo signal from the ultrasonic module, the duration of this high-level pulse corresponds to the round-trip propagation time of the ultrasonic wave from emission to reflection and back to reception. Based on this, the actual distance between the unmanned ship and the target obstacle can be calculated. (Figure 5 The working principle of the ultrasonic module).

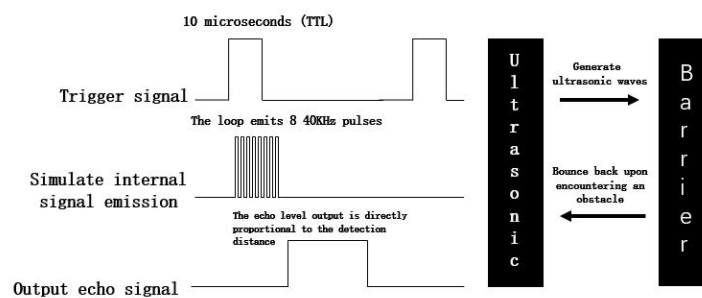


Figure 5 The Working Principle of the Ultrasonic Module

3.6 Sensor Module

To achieve the detection function of the unmanned ship, this project integrates various types of sensors, including the TS-30 turbidity sensor, pH sensor, 18B20 temperature sensor, and TDS sensor, etc. In practical applications, the test sensors can be selected according to actual needs. To achieve the detection function of the unmanned ship, this project integrates various types of sensors, including the TS-30 turbidity sensor, pH sensor, 18B20 temperature sensor, and TDS sensor, etc. The specific process is as follows: Connect the signal pins of the test sensor to the ANALOG IN interface on the Raspberry Pi motherboard, connect the positive power supply to the VCC interface on the main control board, and connect the negative power supply to the GND interface. (Figure 6 Principle structure diagram of the TDS sensor).

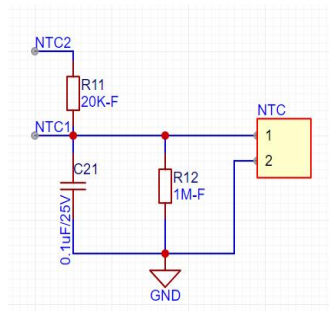


Figure 6 Principle Structure Diagram of the TDS Sensor

According to the actual testing requirements, the sensors can be connected to the Raspberry Pi motherboard. When the sensors collect data and transmit it to the Raspberry Pi motherboard, the motherboard will process the data and send it to the cloud server through the server. Then, users can directly obtain the relevant data information on their mobile phones.

3.7 Power Drive Module

The wind power drive module mainly consists of wind turbines and solar photovoltaic panels. On unmanned ships, waterproof solar photovoltaic panels and small wind turbines are installed. Based on efficient photovoltaic materials and advanced aerodynamics principles, they can start and generate electricity under different light intensities and lower wind speeds. They provide continuous renewable energy input for the power system of unmanned ships. Their surfaces have been specially treated. The solar panels and wind turbines have excellent resistance to seawater corrosion and UV aging, which is sufficient to adapt to harsh sea conditions. The pictures of the wind turbine and the solar panel are shown in Figures 7 and 8.



Figure 7 Small Wind Turbine Generator



Figure 8 Solar Photovoltaic Panel

3.8 Hull Design

To reduce the running resistance of the unmanned ship, the hull design of this project adopts a streamlined shape. In terms of waterproofing, a double-layer waterproof design is used. A waterproof layer is placed inside the hull, and a sealing layer for the cabin is added between the deck and the boat to prevent water from entering the hull. To enhance the visibility of the unmanned ship during operation, the hull is coated in yellow, which makes it easy to quickly determine the location where the unmanned ship is working. At the same time, it is driven by two high-speed motors and two propellers, resulting in lower energy consumption during operation (Figure 9 shows the physical image of the unmanned ship's hull).

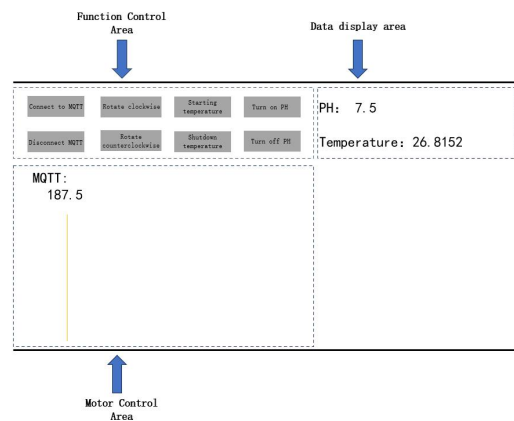


Figure 9 Physical Image of the Unmanned Ship's Hull

4 UNMANNED VESSEL CONTROL SOFTWARE

4.1 Terminal Program Design

This project uses the WxBit programming software to design the mobile control terminal App and the control interface. This software can program the control interface according to the actual needs of users, and the program is simple to write. The position detection function of the unmanned boat can obtain the specific position of the unmanned boat and monitor the target environment by using the Baidu Map API. The control interface is shown in Figure 10.

**Figure 10** Terminal Application of Unmanned Vessel

4.2 Terminal Control System Design

This project combines software and hardware to design the control system. It uses the user terminal to control the unmanned boat and thereby achieve monitoring. The first step is to connect the Raspberry Pi motherboard to the computer. The second step is that the program can write the software Visual Studio Code onto the motherboard and use the mobile App to control the navigation direction of the unmanned boat. The interface of the APP can obtain detection data through various sensors. The third step is to obtain the corresponding IP through the connection of the Raspberry Pi motherboard, and the user logs into the local network IP on the web page to set the camera. The camera image is displayed through the web page on the computer end. The position of the unmanned boat is obtained by using the GPS positioning module on Baidu Map[14].

5 OPERATION AND DEBUGGING OF UNMANNED VESELS

5.1 Test Preparation

To ensure the safety, reliability and accuracy of the unmanned vessel in practical applications, tests are conducted before its deployment in the actual environment, prior to the laboratory setting: the sealing performance of the hull is tested to verify whether it has the ability to withstand underwater pressure and prevent water vapor penetration; at the same time, the operation performance of the motor is comprehensively inspected to ensure that it can stably output power in the water environment and drive the unmanned vessel to navigate efficiently; for the WiFi module, the stability and smoothness of signal transmission are analyzed to ensure that the unmanned vessel can achieve real-time and accurate data interaction in complex water environments; in addition, various sensor modules also need to undergo strict tests, including attitude sensors, environmental sensors, etc., to test whether they can accurately perceive the surrounding environmental information when working in the water environment, providing a reliable basis for the intelligent decision-making of the unmanned vessel. The power supply voltage of each module is checked to ensure its stability, whether there are errors in the data transmission process, and the accuracy of the sensor measurement results, so as to comprehensively ensure the safety of the unmanned vessel in practical applications.

5.2 Actual Operation test

This project developed an unmanned ship that operates in an autonomous mode and has detailedly recorded its navigation trajectory and yawing conditions. During the actual water area tests, the unmanned ship ran normally. The hull has good sealing performance, ensuring that all modules in the cabin can operate normally. At the same time, the Wi-Fi signal transmission was normal. Even in the sea area, its data could maintain stable transmission, and the terminal equipment could receive various data in real time, ensuring the timeliness and accuracy of information interaction.

Moreover, the unmanned ship has a large-capacity battery and a wind-solar hybrid power generation system, providing energy for its long-term operation. The test results show that the endurance and stability of this unmanned ship are both good.

5.3 Test Result

The unmanned vessel successfully followed the preset autonomous route, and the actual route was basically coincident with the preset route. The maximum yaw angle was 2.5° , and the average yaw angle was 0.8° . The navigation time was 12 minutes, which was within the expected time range. The unmanned vessel demonstrated excellent stability and accuracy in the autonomous navigation mode, and the yaw situation was within the acceptable range, verifying its self-sufficient navigation capability.

During the autonomous navigation test of the unmanned ship, the ship demonstrated excellent navigation capabilities. During the test, it was able to accurately follow the preset route and conduct fixed-point sampling without deviating from the course or colliding with obstacles. Its autonomous obstacle avoidance system worked properly, responding promptly when encountering potential obstacles and adjusting the navigation direction to ensure navigation safety. At each sampling point, the speed was adjusted to ensure the accuracy of the samples. During the test, the unmanned ship could maintain a stable navigation state and could promptly sense environmental changes and adjust its operating state. In addition, the navigation efficiency of the unmanned ship also met the expected goals. During the test, it could complete the tasks at a relatively fast speed while maintaining low energy consumption. The actual measurement and navigation of the unmanned vessel are shown in Figure 11-12.



Figure 11 Measured Data of the Unmanned Ship



Figure 12 Navigation Map of the Unmanned Ship

6 CONCLUSION

This project develops an intelligent unmanned ship controlled by Raspberry Pi. It has diverse functions, including autonomous navigation, data collection, and the ability to perform complex tasks such as underwater topography mapping and water quality analysis. In terms of power supply, a power system combining wind turbines and solar charging panels was developed. This design aims to utilize the complementarity of wind energy and solar energy to create an efficient, environmentally friendly, and sustainable energy solution. By using a mixed power drive method of wind and solar energy, the unmanned ship can achieve long voyages and ensure stable power supply, enabling it to adapt more flexibly to various environmental conditions[15].

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

The research was supported by the Characteristic Innovation Project of Zhanjiang Preschool Education College in 2022 (Natural Science Research) (ZY2022TSCX02); the Guangdong College Students' Science and Technology Innovation Cultivation Special Fund in 2025 (Project Number: pdjh2025bc407), Project Name: Amphibious Robot Developed Based on Raspberry PI.

REFERENCES

- [1] Xin J, Gong H, Zhai B et al. Application analysis of single and multi-beam unmanned vessel underwater measurement. *Shaanxi Water Resources*, 2025(3): 126–128, 132.
- [2] Ji X. Application of unmanned monitoring vessel in river pollution source tracking. *Heilongjiang Environmental Bulletin*, 2024, 37(11): 142–144.

- [3] Lai J, Lin Z, Chen X et al. Trajectory tracking control strategy for unmanned vessel for seawater quality detection. *Journal of Guangzhou Maritime University*, 2024, 32(2): 26–32, 49.
- [4] Li J. Design of positioning and navigation system for garbage collection unmanned vessel based on depth camera. *Integrated Circuits Application*, 2024, 41(11): 60–62.
- [5] Wang Y, Ji F, Lu S et al. Research status of ocean intelligent unmanned equipment detection based on multi-physical field. *Journal of Underwater Unmanned Systems*, 2024, 32(6): 1131–1140.
- [6] Zhang W, Li Z, Zheng Y. Current development status and research trends of unmanned vessels abroad and domestic. *Ship Science and Technology*, 2024, 46(15): 79–83.
- [7] Tong J, Yu Y. Visual detection of unmanned vessel based on improved Mask R-CNN. *Ship and Ocean Engineering*, 2025, 54(1): 6–12.
- [8] Zhang Q. Current situation and impact analysis of unmanned vessels on maritime supervision. *China Maritime*, 2023(3): 42–45.
- [9] Xiong S. Single-image visual detection and tracking research for unmanned vessel based on infrared image. *Ship Science and Technology*, 2024, 46(7): 159–162.
- [10] Wu H, Wu W, Liao J et al. Design of fixed-point cruising system for unmanned vessel based on ROS-Web interaction. *Ship Electronic Engineering*, 2023, 43(11): 64–69.
- [11] Zhai J, Pan H. Research on task reliability test method of unmanned vessel. *Ship Standardization and Quality*, 2023(4): 53–57, 30.
- [12] Wu H, Wu Z, Meng Q et al. Autonomous cruising unmanned vessel system in water environment monitoring scenarios. *Electronic Making*, 2023, 31(2): 14–17.
- [13] Liu Y, Yang X, Chu H et al. Design and implementation of unmanned vessel control system. *Manufacturing Automation*, 2022, 44(10): 127–131.
- [14] Jin J, Zhang J, Liu D et al. Research on visual tracking of high-speed unmanned vessel targets based on improved SiamRPN. *Advances in Marine Science*, 2024, 42(3): 545–554.
- [15] Wang H, Li C, Wang K. Application of unmanned vessel measurement technology in feasibility analysis of river sand mining. *Zhihai*, 2023(11): 26–27.

THE INTERACTIVE EFFECTS OF TITLE SENTIMENT AND LENGTH ON VIDEO DISSEMINATION: OPTIMAL LENGTH AND THE DIMINISHING RETURNS OF EMOTION

YaWei He^{1*}, ZiYi Guan¹, Wen Liu^{2,3}

¹*School of Media and Communication, Shanghai Jiao Tong University, Shanghai 200240, China.*

²*AI Technology Innovation Group, School of Economics and Management, Communication University of China, Beijing 100024, China.*

³*Department of Engineering, University of Cambridge, Cambridge CB3 0FS, United Kingdom.*

Corresponding Author: YaWei He, Email: heyawei1992@sjtu.edu.cn

Abstract: In the highly competitive ecosystem of online video platforms, metadata elements such as titles are critical determinants of user engagement. This study uses computational communication methods to explore the interactive effects of title sentiment intensity and title length on video view counts for automotive videos on the Bilibili platform. Through hierarchical regression analysis on a large-scale dataset of 984 videos, and supported by non-parametric Generalized Additive Models (GAM) and bootstrap analysis for robustness, this research reveals two key findings: First, there is a significant inverted U-shaped relationship between title length and video views, meaning views increase with length up to a peak at approximately 26-30 characters, and then decline. Second, title length plays a negative moderating role in the relationship between sentiment intensity and views. This finding suggests that emotional language is more effective in attracting user attention in short titles, but its appeal diminishes in longer, more information-dense titles, and may even have a negative impact if perceived as "clickbait." These findings offer nuanced, data-driven strategies for content creators and digital marketers for optimizing video metadata to maximize audience engagement on competitive digital platforms.

Keywords: User engagement; Sentiment analysis; Video analytics; Title optimization; Moderation analysis; Computational communication

1 INTRODUCTION

The proliferation of online video platforms has fundamentally reshaped the landscape of media consumption and information dissemination. Platforms like YouTube, TikTok, and Bilibili have evolved into primary arenas where billions of users seek entertainment, education, and product information [1]. For high-involvement consumer goods such as automobiles, these platforms serve as a critical channel where potential buyers engage in extensive information-seeking, watching reviews, comparisons, and test-drive experiences to inform their purchasing decisions [2]. In this densely saturated media environment, the competition for user attention is ferocious. An estimated 500 hours of video are upload to YouTube every minute, creating a situation where the supply of content vastly exceeds the audience's limited attentional capacity [3].

This content overload places immense importance on the role of metadata—particularly video titles—in the initial, crucial stage of content discovery. A video's title is often the first point of contact with a potential viewer, functioning as a powerful signal that influences the decision to click and watch [4]. It must be compelling enough to stand out in a crowded feed of recommendations and search results, effectively communicating the video's content and value proposition in a fleeting moment. Consequently, understanding the specific characteristics of titles that drive viewership is a paramount concern for content creators, digital marketers, and platform algorithm designers alike.

This study focuses on Bilibili, a uniquely influential platform in the Chinese digital ecosystem. Unlike its global counterparts, Bilibili has cultivated a distinctive community culture centered around its predominantly Generation Z user base, user-generated content from creators (known as "UPs"), and its signature "bullet comments" (danmu)—a real-time, overlaying commentary system that fosters a strong sense of co-viewing and community participation [5]. This highly engaged and interactive environment makes Bilibili an ideal laboratory for studying the dynamics of content engagement [6].

Previous research has identified two primary title characteristics as key drivers of engagement: emotionality and length. Studies in communication and marketing have consistently shown that content evoking strong emotions, whether positive or negative, is more likely to be shared and consumed [7]. Simultaneously, the length of a headline or title has been shown to influence readability and cognitive load, with a general consensus that an optimal length exists, avoiding the pitfalls of being either un-informatively brief or overwhelmingly verbose [8].

However, a significant gap persists in the literature. Most studies have examined these factors in isolation, neglecting their potential interactive effects. It is plausible that the effectiveness of an emotional cue is not uniform but is contingent on the overall structure of the title, such as its length. For instance, is a strong emotional appeal equally effective in a concise, 15–

character title as it is in a descriptive, 40-character one? Furthermore, many studies assume a linear relationship, whereas the true effect could be curvilinear, featuring points of diminishing returns.

This research addresses this gap by investigating the individual, non-linear, and interactive effects of title sentiment intensity and title length on the view count of automotive videos on Bilibili. The automotive sector has chosen for its status as a high-involvement category where viewers are motivated information seekers, making title accuracy and appeal particularly salient [9]. By employing a computational approach on a large-scale dataset, this study seeks to answer the following research questions:

- (1) What is the nature of the relationship between video title length and view count? Is it linear or curvilinear?
- (2) What is the relationship between the emotional intensity of a video title and its view count?
- (3) Does title length moderate the relationship between sentiment intensity and view count?

By answering these questions, this paper aims to provide a more nuanced and holistic understanding of video title optimization. The findings will contribute to communication theory by elucidating the complex interplay of informational and emotional cues in digital media, while also offering practical, data-driven guidelines for stakeholders in the digital content ecosystem.

2 LITERATURE REVIEW AND HYPOTHESES DEVELOPMENT

2.1 Video Titles as Information Scent

In the vast information landscape of the internet, users behave like foragers, seeking valuable content while trying to minimize effort. Information Foraging Theory (IFT) posits that users decide which content to pursue based on "information scent"—cues in their environment that signal the potential value and cost of accessing a piece of information [10]. In the context of video platforms, a video's title, along with its thumbnail, constitutes the primary information scent. A strong scent convinces the user that clicking the video will lead to relevant, useful, or entertaining content that satisfies their current information need [11].

The quality of this scent has determined by the characteristics of the title. Two fundamental components contribute to this quality: the informational content and the affective (emotional) content. The length of the title has directly related to the amount of information it can convey. A very short title may provide insufficient scent, leaving the user uncertain about the video's content. Conversely, a very long title might create an overwhelming scent, increasing cognitive load and making it difficult for the user to parse its relevance quickly [12]. This suggests that the relationship between the informational value of a title and its length is not linear. There likely exists a "sweet spot" where the title is descriptive enough to provide a strong scent without inducing cognitive friction [13]. This leads to the hypothesis of a non-linear, inverted U-shaped relationship between title length and user engagement.

H1a: There is a positive linear relationship between video title length and view count.

H1b: There is a negative quadratic relationship between video title length and view count, suggesting an inverted U-shaped curve where view count increases with length up to an optimal point and then decreases.

2.2 The Role of Emotionality in Driving Engagement

The second critical component of information scent is its affective quality. Decades of research in psychology and communication have established that emotion is a powerful driver of attention, memory, and behavior [14]. Content that elicits an emotional response is more likely to be noticed, processed, and shared. This phenomenon is explained by theories such as the Elaboration Likelihood Model (ELM), which suggests that cues evoking emotion can act as powerful peripheral signals that influence attitudes and behaviors, especially when viewers are not deeply engaged in systematic processing (e.g., when scrolling through a feed) [15].

Berger & Milkman (2012) found that The New York Times articles that evoked high-arousal emotions (like awe, anger, or anxiety) were more likely to become viral [16]. Similarly, studies of social media posts have found that both positive and negative sentiment can increase engagement, often mediated by the level of emotional arousal [17].

The intensity of the emotion, rather than just its positive or negative valence, appears to be a key mechanism. A title with high emotional intensity—whether strongly positive (e.g., "The Most INCREDIBLE Car I've Ever Driven!") or strongly negative (e.g., "Why This Car is a COMPLETE Disaster!")—creates a stronger affective scent, promising a more engaging or dramatic viewing experience compared to a neutral, purely factual title (e.g., "A Review of the 2024 Model X") [18]. The distribution of sentiment scores in the dataset, which peaks at moderate levels but also shows a wide spread, suggests that creators are experimenting with varying levels of emotionality. The bivariate analysis indicates a complex, U-shaped relationship, where both very low (neutral) and very high-intensity titles perform better than moderately emotional ones. This suggested that being emotionally unambiguous, either by being purely factual or highly emotional, is a superior strategy [19].

H2: Higher emotional intensity in a video title is associated with a higher view count.

2.3 The Interaction of Length and Sentiment: A Moderation Hypothesis

The central argument of this paper is that the effects of title length and sentiment intensity are not independent. We propose that title length moderates the effect of sentiment on viewership. This hypothesis is grounded in the principles of

cognitive load and signal clarity. The cognitive resources a user is willing to expend on evaluating a single video title are extremely limited. The processing of emotional cues must therefore be efficient [20].

When a title is short, it has low cognitive load. In this context, an emotional word or phrase acts as a very clear and potent signal. It is the dominant component of the information scent and can effectively capture attention. For example, in a short title like "Mind-Blowing Speed!", the emotional component is unambiguous and powerful.

However, as a title gets longer, its informational density increases. It contains more descriptive keywords, technical specifications, or contextual phrases. This increased length demands more cognitive effort from the user to read and comprehend. In this high-load context, the role of the emotional cue changes. A strong emotional term embedded in a long, descriptive title may be perceived differently. It could be seen as incongruous with the factual tone, potentially reducing the title's credibility and making it appear as "clickbait" [21]. Alternatively, the emotional signal might simply get lost amidst the other informational cues, its potency diluted. The user, already expending effort to process the title's length, may be less receptive to a strong emotional appeal. This suggests that the positive effect of sentiment intensity should weaken as title length increases, and may even become negative for very long titles [22].

This proposed mechanism is visually suggested in the binned heat-map in the present paper, which displays the actual mean view counts for different combinations of title length and sentiment score bins. For instance, within the longest title length bin (31-36), the view count does not consistently increase with sentiment, showing a much more erratic pattern compared to shorter length bins. This provides preliminary evidence for an interaction [23].

H3: Video title length negatively moderates the relationship between sentiment intensity and view count. Specifically, the positive effect of sentiment intensity on view count will be stronger for shorter titles and weaker (or even negative) for longer titles.

3 METHODOLOGY

3.1 Data Collection and Sampling

The dataset for this study was compiled from Bilibili (bilibili.com), a prominent video-sharing platform in China known for its extensive user-generated content and highly engaged community. Data collection was performed using a custom Python-based web scraping script.

The collection period spanned from July 2019 to December 31, 2024, to capture a wide range of videos over time.

The sampling process focused on videos within the automotive content category, which includes a variety of sub-topics such as vehicle reviews, test drives, model comparisons, and brand events. To ensure the sample's diversity and representativeness, videos were collected from creators of varying scales and audience levels, rather than focusing solely on top-tier influencers. This approach enhances the generalizability of the findings across the broader automotive content ecosystem on the platform.

An initial raw dataset was subjected to a data cleaning process. Videos were excluded if they met any of the following criteria: poor image quality, absence of key information (e.g., identifiable persons or vehicles), or asynchronous audio and video. After this filtering process, a final sample of $N=984$ videos was retained for analysis. For each video in the final dataset, key metadata were extracted, including the video title, view count, and publication date.

3.2 Measurement of Variables

3.2.1 Dependent variable: view count

The primary dependent variable, representing the reach and popularity of a video, is its total view count. As is common with online media engagement data, the raw distribution of view counts was highly right-skewed, as illustrated in Figure 1A. The mean view count (813,272) was significantly larger than the median (563,728), indicating the presence of high-performance outliers that violate the assumptions of ordinary least squares (OLS) regression. To normalize the distribution and stabilize the variance, a natural logarithmic transformation was applied ($\text{Log_View_Count} = \ln(\text{View_Count})$). The resulting transformed variable, shown in Figure 1B, closely approximates a normal distribution, making it suitable for parametric statistical analysis.

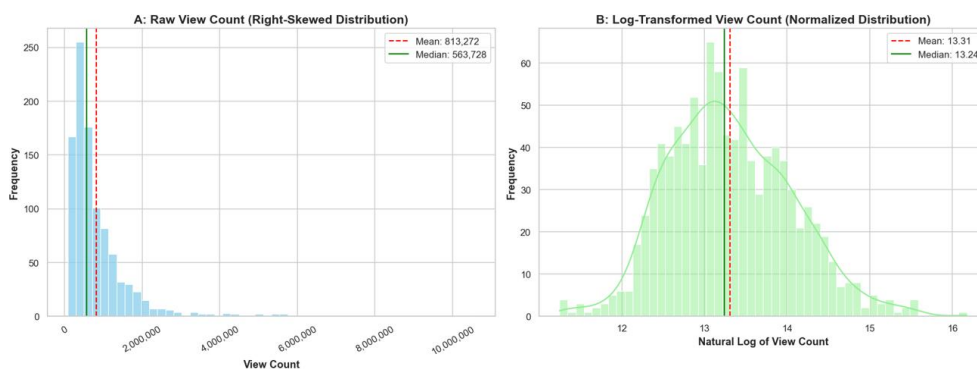


Figure 1 Distribution of View Count

3.2.2 Independent variables: title length and sentiment intensity

Title Length: This variable was measured as the total number of characters in the video's title. The distribution of title lengths, shown in Figure 2, ranged from 3 to 49 characters, with a mean of 17.94 and a median of 17.00.

Sentiment Intensity Score: The emotional intensity of each title was quantified using a sophisticated two-step process. First, a pre-trained BERT (Bidirectional Encoder Representations from Transformers) multilingual sentiment classification model was employed. This model trained on a five-star rating system, is adept at capturing both the polarity (positive/negative) and magnitude of sentiment in text. For each title, the model calculated a raw sentiment score (Sraw) by computing a weighted probability average, which was then standardized to a [0, 1] interval. On this scale, 0 represents the most negative sentiment, 1 represents the most positive, and values around 0.5 are neutral.

Second, recognizing that user engagement is often driven by the magnitude of emotion rather than its direction, the raw score was converted into an intensity score (Sintensity). This was achieved using the formula: $\text{Sintensity} = 2 * |\text{Sraw} - 0.5|$. This transformation maps titles with sentiment scores near the neutral midpoint (0.5) to intensity scores near 0, and titles with scores near the extremes (0 or 1) to intensity scores near 1. The final sentiment intensity score, therefore, ranges from 0 (completely neutral) to 1 (maximum emotional intensity). The distribution of this final score is presented in Figure 3.

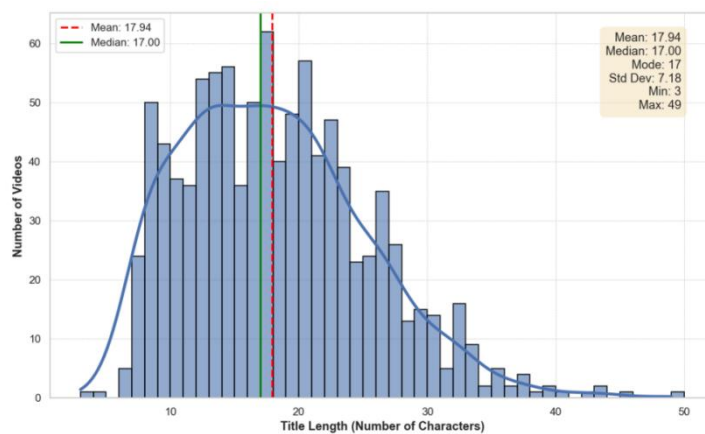


Figure 2 Distribution of Title Length

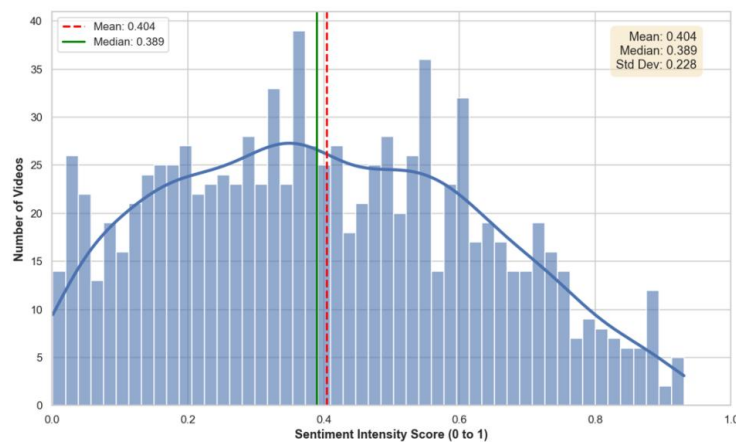


Figure 3 Distribution of Sentiment Intensity Score

3.2.3 Control variable

To account for potential temporal effects, such as platform algorithm changes or shifts in user behavior over the long data collection period, a control variable for publication date was included. Videos were categorized into two groups: those published before July 2023 ("older") and those published on or after July 2023 ("newer").

3.3 Analytical Strategy

The data analysis was conducted in three main stages. First, descriptive statistics and visualizations were used to characterize the distributions of all key variables and to conduct a preliminary exploration of the bivariate relationships. This included plotting the distribution of view counts before and after log transformation (Figure 1), the distributions of title length (Figure 2) and sentiment intensity (Figure 3), and creating binned plots to visualize the mean view counts across different strata of sentiment intensity (see Figure 4) and title length (see Figure 5).

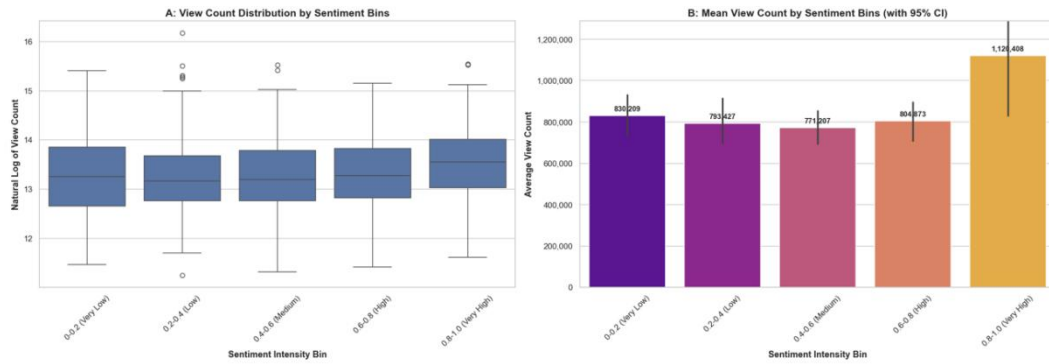


Figure 4 View Count Analysis by Binned Sentiment Intensity

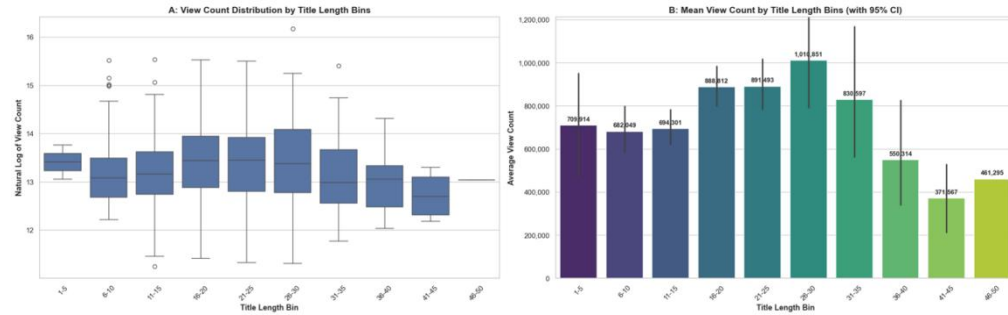


Figure 5 View Count Analysis by Binned Title Length

Second, to formally test the hypotheses, a hierarchical OLS regression analysis was performed on the log-transformed view count. This approach allows for a systematic assessment of the explanatory power added by each set of variables. Four nested models were specified:

- Model A: Included only the sentiment variable ($\text{Log_View_Count} \sim \text{Sentiment} + \text{Sentiment_sq}$).
- Model B: Included only the length variables ($\text{Log_View_Count} \sim \text{Length} + \text{Length_sq}$).
- Model C (Main Effects Model): Included all main effect terms ($\text{Log_View_Count} \sim \text{Sentiment} + \text{Sentiment_sq} + \text{Length} + \text{Length_sq}$).
- Model D (Interaction Model): Included all main effects plus the interaction term to test the moderation hypothesis ($\text{Log_View_Count} \sim \text{Sentiment} + \text{Sentiment_sq} + \text{Length} + \text{Length_sq} + \text{Sentiment} * \text{Length}$).

The quadratic terms (Sentiment_sq and Length_sq) were included to model the hypothesized non-linear relationships. Model fit was compared across the hierarchy using the Adjusted R^2 statistic.

Third, to ensure the robustness and validity of the findings from the primary regression analysis, two additional analyses were conducted. A non-parametric Generalized Additive Model (GAM) was fitted to visualize the interaction surface without the constraints of a pre-specified functional form. Furthermore, a bootstrap analysis (500 resamples) was performed on the interaction coefficient from Model D to assess its stability and to compute a robust 95% confidence interval.

4 RESULTS

4.1 Descriptive Statistics

The final dataset comprised 984 videos. The primary dependent variable, view count, exhibited a wide range, with a mean of 813,272 and a standard deviation of 798,253. As shown in Figure 1, the log-transformed view count was normally distributed with a mean of 13.31 and a standard deviation of 0.747. The independent variables also showed considerable variation. The title sentiment intensity score had a mean of 0.404 ($SD = 0.228$). The title length ranged from 3 to 49 characters, with a mean of 17.94 ($SD = 7.18$). The distributions of sentiment and length are visualized in Figure 2 and Figure 3, respectively.

4.2 Bivariate Analysis

Preliminary analysis of the relationships between the independent variables and view count was conducted by grouping the data into bins. Figure 4 shows the average view count across five sentiment intensity bins. A U-shaped or J-shaped pattern is observable. The lowest average view count was found in the medium sentiment bin (0.4–0.6), while the highest average was in the very high sentiment bin (0.8–1.0), which had a mean view count of 1,120,408. This suggests that both emotionally neutral and highly charged titles outperform moderately emotional ones.

Figure 5 displays the average view count across ten title length bins, revealing a clear inverted U-shaped relationship. Viewership increases with title length, peaks in the 26–30 character bin with a mean of 1,010,851 views, and then declines steadily for longer titles. The lowest performance was observed in the 41–45 character bin. This provides strong initial support for H1b, positing a curvilinear effect of title length.

4.3 Hierarchical Regression Analysis

To formally test the hypotheses, a hierarchical OLS regression was performed. The progression of model fit is visualized in Figure 6A. Model A (sentiment only) explained very little variance ($\text{Adj. } R^2=0.0030$). Model B (length only) provided a substantial improvement ($\text{Adj. } R^2=0.0164$). The main effects model (Model C) offered a slight improvement over Model B ($\text{Adj. } R^2=0.0181$). The final interaction model (Model D) yielded the best fit, with an Adjusted R^2 of 0.0220. The statistically significant increase in explanatory power from Model C to Model D underscores the importance of the interaction term.

The detailed results for the final model (Model D) are visualized in the coefficient plot in Figure 6B. The key findings are as follows:

- (1) Title Length: The linear term for 'Length' was positive and highly significant ($\beta=0.0660$, $p<0.001$), while the quadratic term 'Length_sq' was negative and highly significant ($\beta=-0.0012$, $p<0.001$). This pair of coefficients provides strong statistical confirmation for the inverted U-shaped relationship hypothesized in H1b.
- (2) Sentiment Intensity: In the final model, the main effects for 'Sentiment' ($\beta=0.2499$, $p=0.524$) and 'Sentiment_sq' ($\beta=0.5976$, $p=0.169$) were not statistically significant. This suggests that, after accounting for title length and the interaction effect, sentiment does not have a simple, direct impact on view count, lending indirect support to the moderation hypothesis.
- (3) Interaction Effect: Crucially, the interaction term 'Sentiment_Length' was negative and statistically significant ($\beta=-0.0325$, $p=0.027$). This result supports H3, indicating that title length negatively moderates the effect of sentiment intensity on log-transformed view count.

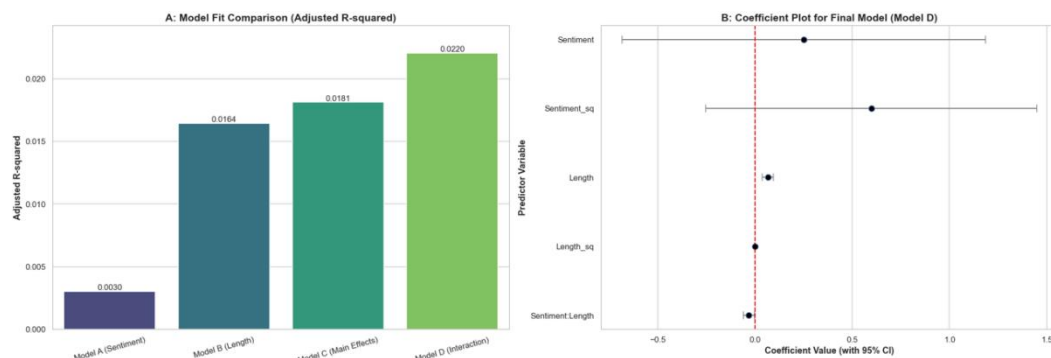


Figure 6 View Count Analysis by Binned Title Length

4.4 Visualization of the Interaction Effect

The nature of the significant interaction effect is detailed in the visualizations presented in Figure 7. The Conditional Effects Plot (Panel B) is key to interpretation. It displays the predicted relationship between sentiment score and view count at three different levels of title length (short, medium, and long). For short titles (e.g., ~10.8 characters), the slope is positive and steep, indicating that higher sentiment strongly predicts higher viewership. For medium-length titles (e.g., ~18.0 characters), the positive slope is weaker. For long titles (e.g., ~26.8 characters), the slope becomes slightly negative. This plot vividly illustrates the moderating role of title length. The Interaction Heatmap (Panel C) visualizes the model's predictions across the entire continuous range of both variables, confirming the interaction.

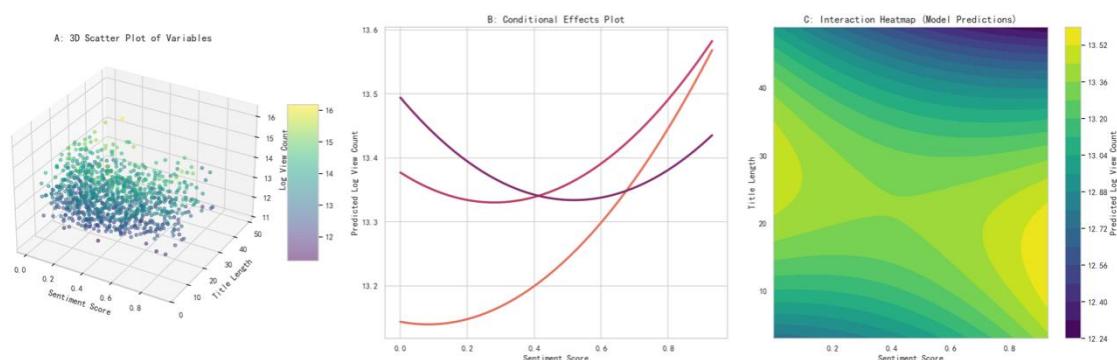


Figure 7 Interaction Visualizations

5 ROBUSTNESS CHECKS

To ensure the reliability of the regression results, two additional analyses conducted. Further diagnostic plots for the sentiment variable and model residuals are provide in Figure 8, showing acceptable patterns for the regression assumptions. The 2D density analyses in Figure 9(Hexbin Plot) and Figure 10(Binned Heatmap) provide a non-parametric view of the data, revealing the actual concentrations of high-viewership videos and supporting the patterns identified by the regression model.

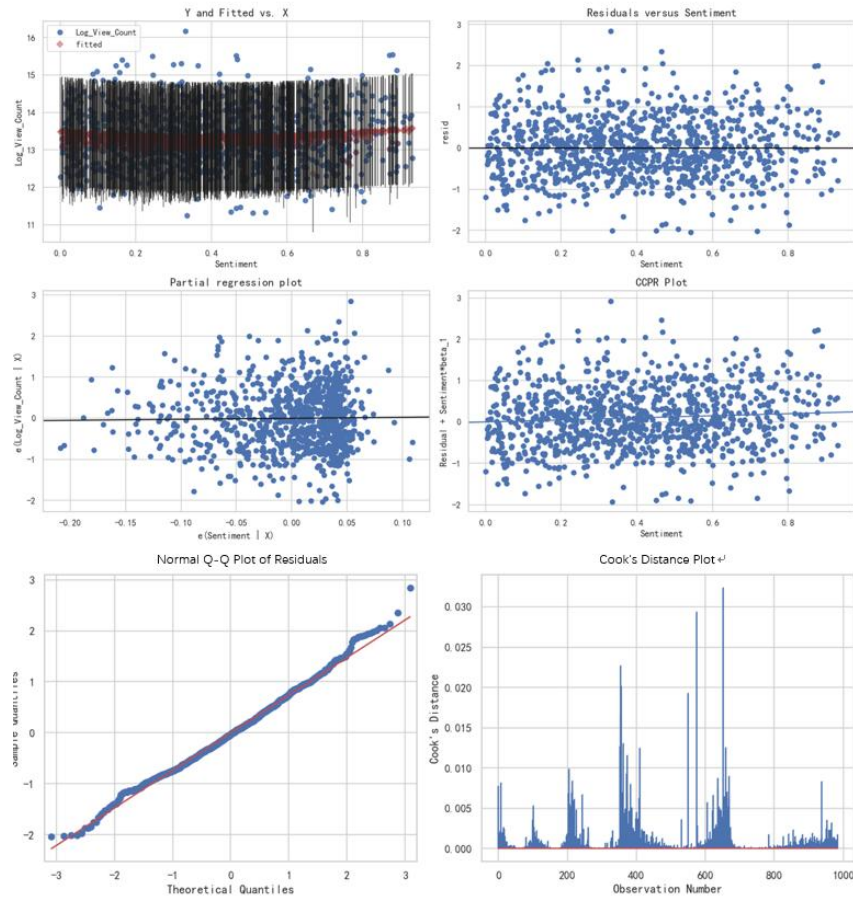


Figure 8 Diagnostic Plots for Sentiment Variable 2D Density Analysis

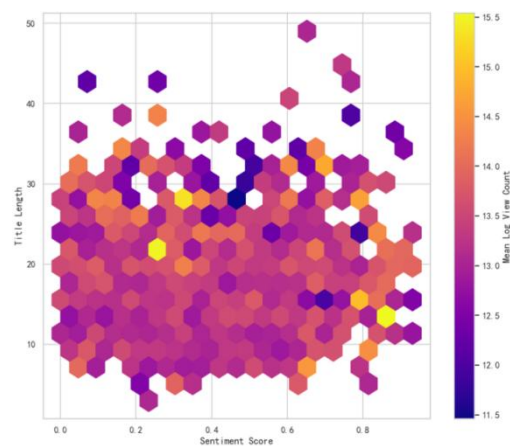


Figure 9 2D Density Analysis (Hexbin Plot)

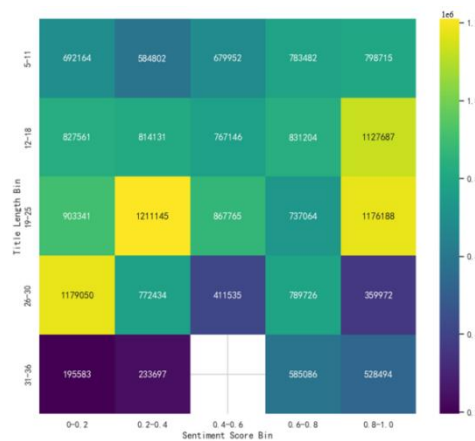


Figure 10 Binned Coupling Analysis (2D Binned Heatmap)

5.1 Non-Parametric Validation with Generalized Additive Model (GAM)

The OLS regression model assumes a specific functional form. To relax this assumption, a Generalized Additive Model (GAM) was fitted. Figure 11 presents the interaction surface plot from the GAM analysis. The GAM plot reveals a complex interaction surface that, while more flexible, corroborates the primary finding from the regression model. It shows that the effect of sentiment on viewership is not uniform across all title lengths. The general pattern observed in the GAM surface is consistent with the interaction found in Model D, lending non-parametric support to our conclusions.

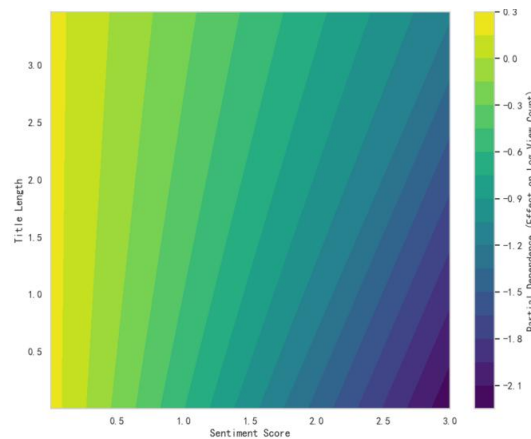


Figure 11 GAM Interaction Surface

5.2 Bootstrap Analysis of the Interaction Coefficient

To assess the stability of the estimated interaction coefficient, a bootstrap analysis was performed (500 resamples). Figure 12 displays the bootstrap distribution of the Sentiment*Length, interaction coefficient. The distribution is approximately normal. Most importantly, the 95% confidence interval derived from the bootstrap distribution is $[-0.064, -0.003]$. As this interval does not contain zero, we can confidently conclude that the negative interaction effect is statistically significant and robust. This analysis strongly reinforces the central finding of the study.

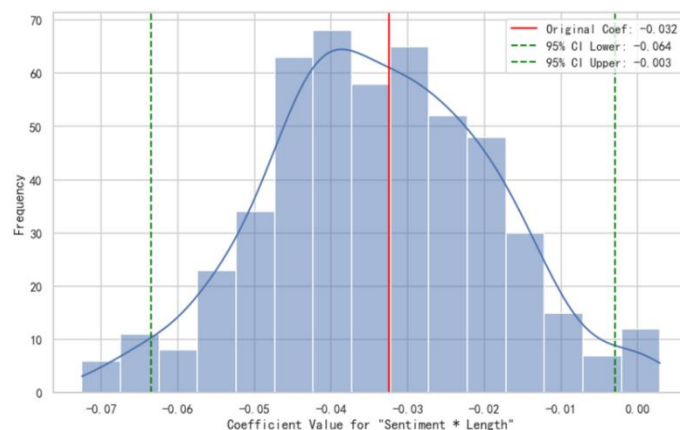


Figure 12 Bootstrap Distribution of the Interaction Coefficient

6 DISCUSSION

6.1 Summary of Key Findings

This study set out to dissect the complex, interactive relationship between video title sentiment and length in predicting viewership on the Bilibili platform. The results from our hierarchical regression, supported by robust visualization and validation techniques, offer a nuanced perspective that advances both communication theory and practical content strategy. The investigation yielded two primary findings. First, we confirmed a significant inverted U – shaped relationship between title length and view count. Viewership tends to increase as titles become more descriptive, but only up to an optimal point of approximately 26–30 characters, after which performance declines. This suggests a trade-off between providing sufficient information scent and imposing excessive cognitive load.

The second and more novel finding is the significant moderating role of title length on the effect of sentiment intensity. Our analysis revealed a negative interaction, meaning the positive impact of a strong emotional hook on attracting viewers is most potent in shorter titles. As titles grow longer and more informationally dense, the effectiveness of this emotional appeal diminishes and can even become detrimental. This interaction is the core contribution of our paper, highlighting that the "rules" for using emotional language in titles are not universal but are conditional on other structural characteristics of the title.

6.2 Theoretical Implications

These findings have important implications for communication theories applied to digital media. They enrich Information Foraging Theory (IFT) by demonstrating that "information scent" is not a simple sum of its parts. Rather, different cues within the scent—in this case, informational cues (length) and affective cues (sentiment)—interact with each other. A strong affective cue can enhance a weak informational signal (a short title) but may conflict with or be diluted by a strong informational signal (a long title). This suggests that the cognitive calculus users perform when evaluating content is interactive and context-dependent.

The results also resonate with the Elaboration Likelihood Model (ELM). The interaction can be interpreted through the lens of central versus peripheral processing. A short, emotional title may function primarily as a peripheral cue, quickly capturing attention and prompting a click with minimal cognitive effort. A long, descriptive title invites more central, systematic processing, where the user evaluates its informational relevance. In this mode, a strong emotional plea might be perceived as a persuasive tactic rather than a genuine signal of content quality, leading to skepticism or dismissal. Our study suggests that the effectiveness of a peripheral cue (emotion) is attenuated when the message format encourages more central processing (length).

6.3 Practical Implications

The conclusions of this research offer direct, actionable insights for a range of stakeholders in the online video ecosystem:

(1) For Content Creators and Digital Marketers: The findings provide a data-driven guide for title optimization. The simple advice to "be more emotional" is insufficient. Instead, a more nuanced strategy is required:

1. Optimal Length: Aim for titles in the 26–30 character range, as this appears to be the "sweet spot" for maximizing average views in the automotive category.
2. Conditional Emotionality: If a short, punchy title is desired, incorporating strong emotional language is an effective strategy to make it stand out. If a longer, more descriptive title is necessary to convey complex information, it may be wiser to adopt a more neutral, factual tone to maintain credibility and clarity.
3. Avoid Extremes: Very long titles (>40 characters) consistently underperform and should generally be avoided.

(2) For Platform Algorithm Designers: Recommendation and search algorithms could be refined to account for these interaction effects. Instead of simply up-ranking content with high-sentiment titles, algorithms could learn to weigh sentiment differently based on title length and other features. Recognizing that the "ideal" title is contextual could lead to more accurate predictions of user engagement and better content recommendations.

6.4 Limitations and Future Research

While this study provides valuable insights, several limitations must be acknowledged. First, the data is correlational, and therefore we cannot make definitive causal claims. While we have identified strong associations, experimental studies (e.g., A/B testing different title versions for the same video) would be necessary to establish causality.

Second, our findings are specific to one content genre (automotive) on a single platform (Bilibili). The optimal title length and the nature of the sentiment-length interaction may differ across other content categories (e.g., entertainment, education) and other platforms with different user demographics and interface designs. Future research should seek to replicate these findings in different contexts to assess their generalizability.

Third, this study focused exclusively on the title. A video's success is determined by a multitude of factors, including the thumbnail image, video tags, creator reputation, and the actual quality of the content. Future studies should aim to incorporate these variables into more comprehensive predictive models. For instance, analyzing the congruence between the emotion conveyed in the title and the emotion conveyed in the thumbnail could yield further insights.

Finally, while we controlled for the publication period, the exact time node for collecting view count data was not specified in the original dataset description. Future work should standardize this metric (e.g., view count at 7 days or 30 days post-publication) to ensure perfect comparability across all videos.

7 CONCLUSION

In the relentless battle for audience attention on digital platforms, a video's title serves as its essential opening gambit. This study demonstrates that crafting the perfect title is an exercise in managing a delicate balance between informational clarity and emotional appeal. Our analysis of 968 automotive videos from Bilibili reveals a distinct inverted U-shaped relationship between title length and viewership, pinpointing an optimal range of 26-30 characters that yielded the highest average of over 1 million views. Crucially, we moved beyond simple main effects to uncover a significant negative interaction ($\beta = -0.0325$) between title length and sentiment intensity. This finding indicates that the effectiveness of an emotional appeal is not universal. High emotional intensity, while capable of generating the highest average viewership (over 1.1 million views in the highest sentiment category), is most beneficial in shorter titles. As titles become longer and more descriptive, the positive impact of sentiment diminishes and can even become detrimental, a pattern confirmed by our conditional effects analysis. This research, validated through non-parametric models and bootstrap analysis, provides a robust, empirical foundation for understanding the complex dynamics of user engagement with video metadata. It moves the conversation from what works (e.g., emotion, length) to when and how these elements work together, offering a more sophisticated framework for content optimization in the digital age.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

The authors are grateful for the finance support from the Ministry of Education Industry-Academia Collaborative Education Program (No.231104575131914), General Program of National Social Science Fund of China (No. 22BXW040) and Shanghai Jiao Tong University Participation in Research Program (No. T200PRP42015).

REFERENCES

- [1] Mishra A. The Rise of Short-Form Video: A Digital Revolution. *International Journal of Research Publication and Reviews*, 2025, 6(6): 1-5.
- [2] Chaudhari P. To Study the Influence of Digital Platforms on Consumer Behavior in the Indian Automobile Industry. *International Journal of Research Publication and Reviews*, 2025, 6(5): 1-7.
- [3] Charle Agency. 32 YouTube Statistics 2025: Key Insights & Trends You Need to Know. 2025.
- [4] Erdogmuş N, Esen E. The Influence of Video Format on Engagement and Performance in Online Learning. *Brain Sciences*, 2021, 11(2): 128.
- [5] Guo W, Chen X. The Impact of Bilibili on Gen Z's Cohort Identity and Aesthetic Resonance. In: *Proceedings of the 2023 International Conference on Human-Computer Interaction*, 2023: 143-150.
- [6] Chen Z, He Q, Mao Z, et al. Governing Generation Z in China: Bilibili, bidirectional mediation, and technological politics. *The Information Society*, 2023, 39(2): 82-94.
- [7] Alamäki A, Dirin A, Huotari ML, et al. Expanding knowledge on emotional dynamics and viewer engagement in travel vlogs: An automated content analysis approach. *Journal of Hospitality and Tourism Management*, 2024, 60: 220-230.
- [8] Prinz-Weiß A, Grünewald K, Nottelmann L, et al. Caption it! The impact of headings on learning from texts. *Applied Cognitive Psychology*, 2023, 37(5): 1041-1052.
- [9] Al-Wreikat A, Bin Kadir Shahar H, Abu Seman N, et al. The impact of social media influencers on customers' perceived values, attitudes and purchase intentions in the fashion apparel industry: moderating effect of brand reputation. *Journal of Business and Socio-economic Development*, 2025.
- [10] Chen X, Zhang X. Exploring Prototypicality Adherence Effects on User Engagement in Video Recommender Systems. *Electronic Commerce Research and Applications*, 2024, 65: 101400.
- [11] Liu Y, Zhang L, Chen X. Developers' information seeking in Question & Answer websites through a gender lens. *Information and Software Technology*, 2024, 171: 107429.
- [12] Kim J, Yang X. The effect of subtitle format on cognitive load and comprehension. *Technology, Pedagogy and Education*, 2024: 1-18.13.
- [13] Chen Y. Impacts of titles on user engagement in short video platforms. Master's thesis, Singapore Management University, 2023.
- [14] Tyng CM, Amin HU, Saad MNM, et al. The Influences of Emotion on Learning and Memory. *Frontiers in Psychology*, 2017, 8: 1454.
- [15] Kitchen PJ, Kerr G, Schultz DE, et al. A meta-analysis of the elaboration likelihood model in the electronic word of mouth literature. *International Journal of Consumer Studies*, 2022, 46(5): 1880-1900.
- [16] Berger J, Milkman KL. What Makes Online Content Viral? *Journal of Marketing Research*, 2012, 49(2): 192-205.

- [17] Hasell A, Weeks BE. Engagement in Emotional News on Social Media: Intensity and Type of Emotions. *Journalism & Mass Communication Quarterly*, 2020, 99(1): 95-117.
- [18] Kim J, Yang X. Less is more: Engagement with the content of social media influencers. *Journal of Business Research*, 2024, 172: 114830.
- [19] Hasell A, Weeks BE. Engagement in Emotional News on Social Media: Intensity and Type of Emotions. *Journalism & Mass Communication Quarterly*, 2020, 99(1): 95-117.
- [20] Skulmowski A, Rey GD. Understanding Cognitive Load in Digital and Online Learning: A New Perspective on Extraneous Cognitive Load. *Educational Psychology Review*, 2021, 34(1): 171-196.
- [21] Pal A, Chua AY, Goh DH-L, et al. Click me... ! The influence of clickbait on user engagement in social media and the moderating role of digital literacy. *PLoS ONE*, 2022, 17(6): e0266743.
- [22] Kuiken J, Schuth A, Spitters M, et al. The impact of advertisement disclosure on consumer engagement in short video advertising. *Behaviour & Information Technology*, 2025: 1-15.
- [23] Lee J, Hong IB. Predicting video views through a sentiment analysis of titles and thumbnails. *Journal of Business Research*, 2024, 174: 114971.

MQPF: A MULTI-DIMENSIONAL QUALITY-AWARE PATH FUSION FRAMEWORK FOR QUESTION ANSWERING

XinYi Wang¹, Bo Liu^{2*}

¹National University of Defense Technology, Changsha 410073, Hunan, China.

²Academy of Military Sciences, Beijing 100091, China.

Corresponding Author: Bo Liu, Email: m19105020705@163.com

Abstract: In multi-hop question answering (MHQA) tasks, existing methods typically integrate multiple reasoning paths from knowledge graphs (KGs) and chains of thought (CoTs). Early KG-enhanced methods primarily focus on obtaining relevant knowledge but fail to consider the multi-dimensional quality of reasoning paths. Subsequent works filter paths but treat all retained paths as equally important without further differentiation. Although some recent works attempt to rank paths by quality, they only provide a relative order without quantifying the actual quality differences between paths. To address these limitations, we propose a Multi-dimensional Quality-aware Path Fusion (MQPF) framework. MQPF introduces a multi-dimensional evaluation mechanism that quantifies path quality from semantic, structural, and outcome-based dimensions. Based on the overall scores, MQPF first filters out low-quality paths to reduce noise and then assigns adaptive weights to the remaining paths according to their scores. This method effectively removes unreliable information and enhances the utilization of trustworthy information during reasoning. Experiments show that MQPF performs comparably to baselines on multiple datasets. Moreover, as a model-agnostic module, it can be used as a plug-and-play module to enhance the performance of existing multi-path reasoning methods.

Keywords: Question answering; Large language model; Knowledge graph

1 INTRODUCTION

Effective multi-hop question answering (MHQA) requires combining different reasoning paths to produce accurate answers [1-3]. This involves two key components: knowledge graph (KG) subgraphs that provide structured knowledge and chains of thought (CoTs) that support step-by-step logical reasoning. By integrating both types of reasoning paths, MHQA systems can achieve broader coverage of relevant information while maintaining robust reasoning capabilities [4,5].

Early KG-enhanced large language model (LLM) reasoning methods construct reasoning paths by retrieving or traversing KGs to improve reasoning accuracy and interpretability [3,6-8]. However, these methods only focus on obtaining relevant knowledge, overlooking the importance of evaluating the quality of reasoning paths across multiple dimensions. As a result, low-quality paths that appear highly relevant may introduce noise and biases into the subsequent reasoning steps. Subsequent works try to filter paths by majority voting or top-n sampling [9-12]. However, when generating final answers, they treat all retained paths as equally important, overlooking the quality differences between them. This equal treatment makes it hard for the LLM to effectively leverage the most critical information, especially when paths contain different answers. Some later works attempt to rank retained paths by quality [13-16], but they only establish a relative ordering. These methods cannot quantify the actual quality differences between paths, lacking a mechanism to reflect how much better one path is than another. As a result, two key challenges remain: (1) how to measure multiple quality aspects of reasoning paths, and (2) how to combine these evaluations to best guide path integration and final answer generation.

To address these challenges, we propose a Multi-dimensional Quality-aware Path Fusion (MQPF) framework. Our method begins with a multi-dimensional evaluation framework with three dimensions: Semantic Quality (S LLM): A powerful LLM is prompted as a reasoning quality evaluator to score paths based on logical coherence and factual correctness; Structural Quality (SS truct): Evaluates the structural effectiveness of paths using subgraph density (for KG paths), reasoning path length (for CoT paths), and question-path relevance; Result Quality (SRM): A reward model fine-tuned on human preferences evaluates the reliability of answers from each path. Each path receives a composite score based on these three dimensions. Based on this score, MQPF first filters out low-quality paths by applying a threshold. Next, it assigns adaptive weights to valid paths, making higher-quality paths have greater influence for answer generation. Finally, it makes the LLM prioritize high-weight paths by promoting. This integrated method provides multi-dimensional quality evaluation for hybrid reasoning paths and employs a quality-adaptive fusion strategy that translates the scores into influence weights during answer generation. In summary, our contributions are:

- We propose a comprehensive multi-dimensional quality evaluation framework for evaluating KG subgraphs and CoT reasoning paths in MHQA.
- We introduce a quality-driven fusion mechanism that assigns weights based on path quality to guide path fusion and answer generation.
- Method Experiments show that our method outperforms all baselines and can serve as a plug-and-play module to enhance existing multi-path reasoning methods.

2 RELATED WORKS

Early methods that enhance LLM reasoning with knowledge graphs (KGs) typically retrieve or traverse KGs to obtain reasoning paths or subgraphs. RoG proposes a planning-search-reasoning framework that retrieves paths from KGs to guide LLM reasoning[3]. GNN-RAG retrieves candidate answers from KG subgraphs and extracts the shortest paths between question entities and answer candidates as reasoning paths[6]. SubgraphRAG retrieves relevant triples as subgraphs to generate accurate and interpretable answers[7]. GraphReader uses an LLM to explore the KG and dynamically updates a notebook to record relevant information[8]. However, these methods only focus on retrieving relevant knowledge like reasoning paths and KG subgraphs without considering their quality across multiple dimensions. As a result, low-quality paths that appear highly relevant may lead to wrong final answers.

Subsequent works try to filter paths by majority voting or top-n sampling. Self-consistency CoT first generates multiple reasoning paths and answers and then selects the most consistent one via majority voting[9]. RoK uses CoT to expand the query and find more related entities[10]. It then builds a subgraph of the KG by matching paths between these entities and finally filters the top-n reasoning paths from this subgraph. Forest-of-Thought maintains multiple reasoning trees for parallel reasoning and applies consensus voting to determine the final answer[11]. ARG-KBQA performs multi-hop beam search over the KGs and converts the top-n reasoning paths into logical forms to enhance LLM reasoning[12]. However, these methods treat all retained paths as equally important during answer generation without further differentiation. This uniform treatment makes it difficult for the LLM to effectively leverage the most reliable information.

Some later works attempt to rank the retained paths by quality. AdaPCR uses a dense retriever to fetch candidate passages and performs context-aware reranking by concatenating each passage with the question to form a new query[13]. PromptRank uses an LLM to compute the generation probability of a question given a path as a relevance score, which is then used to rank the paths[14]. RAGtifier employs a Pinecone retriever and a BGE reranker to improve retrieval quality[15]. AttnRank finds that LLMs pay more attention to the documents at the beginning and the end[16], so it places the most relevant documents in these high-attention positions. While these methods order paths by relevance, they do not quantify the actual quality differences between paths, making it difficult to assess how much better one path is compared to another.

3 METHOD

To address the issue that multi-path information is not well utilized in multi-hop QA tasks, we propose a multi-dimensional quality assessment framework and a path filtering and weighted fusion mechanism. Specifically, we first generate multiple reasoning paths based on the question (including structured KG paths and textual CoTs), and then evaluate these paths from semantic, structural, and outcome-based dimensions. Based on the overall score of the path, we filter out low-quality paths and assign weights to the remaining valid paths to guide the LLM in generating the final answer. The overall framework is shown in Figure 1.

3.1 Multi-Dimensional Quality Assessment Framework

To effectively capture the quality differences among various reasoning paths, we propose a multi-dimensional quality assessment framework that quantifies the quality of reasoning paths from three perspectives: semantic, structural, and outcome-based dimensions. The goal of this framework is to assess the overall quality of each reasoning path and assign it an overall score. These scores are then used in the next step to weight and merge the reasoning paths.

3.1.1 Semantic quality

Semantic quality assessment measures whether a reasoning path is logically consistent and factually accurate. We use GPT-4, a large language model with strong reasoning and prompting capabilities, to evaluate each reasoning path through structured prompts. The prompt template is shown in Figure 2.

To improve scoring consistency, we evaluate each reasoning path three separate times and take the average as the semantic quality score (S_{LLM}). The result is considered valid only when the standard deviation of the three scores is below 0.5.

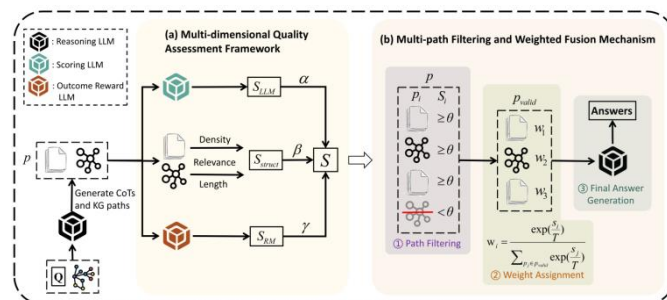


Figure 1 The Overview of MQPF. (a) Multi-dimensional Quality Assessment Framework quantifies the quality of reasoning paths from semantic, structural, and outcome-based dimensions. (b) Multi-path Filtering and Weighted Fusion Mechanism contains 3 parts: ① Path Filtering, ② Weight Assignment, and ③ Final Answer Generation.

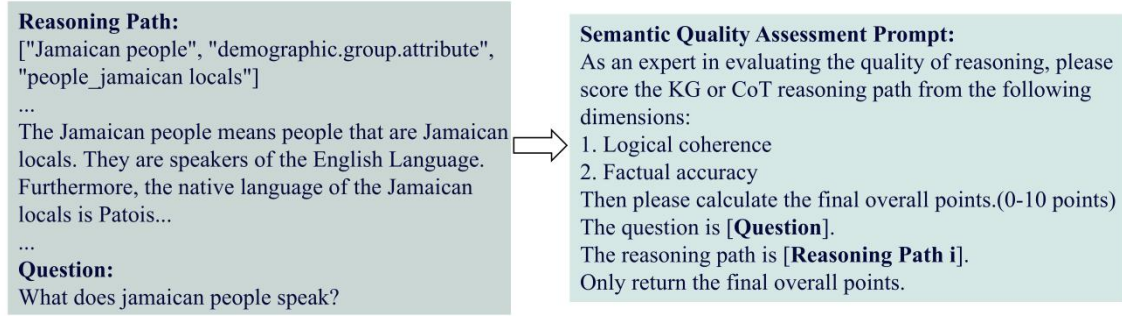


Figure 2 Prompt Template for Semantic Quality Assessment

3.1.2 Structural quality

Structural quality assessment focuses on evaluating the effectiveness of a reasoning path's structure. Paths that are more compact and contain higher information density receive higher scores. Considering the structural differences among path types, we designed separate evaluation strategies for structured paths (KG subgraphs) and textual paths (CoTs).

Structured Paths The structural quality of this type of path consists of two components: (1) Graph Density, which measures the connectivity and informational richness between nodes to avoid isolated or sparse paths; (2) Semantic Relevance between the path and the question, where higher relevance leads to a higher score. For the former, we use the number of relations and entities in the structured path to calculate it; for the latter, we convert the structured path into a sequence of triples and encode it into vectors. We then measure its relevance to the question embedding by calculating the cosine similarity between them. For a structured path $p_i(r, e)$ and question Q , the structural quality score S_{struct} is:

$$S_{struct} = \frac{\text{num}(r)}{\text{num}(e)} \cdot \text{correlation}(Q, p_i) \quad (1)$$

where r are the relationship edges and e are the entity nodes in the path. The correlation function calculates the relevance between question Q and reasoning path p_i with cosine similarity.

Textual Paths The scoring criteria for this type of path include length and relevance to the question. Although techniques like CoT improve the interpretability and correctness of the reasoning process, they often lead LLMs to produce unnecessarily long and redundant steps [17,18]. Therefore, we link the structural score of textual paths to their length. Paths that are more concise and information-dense receive higher scores. For a textual reasoning path p_i , the structural quality score is:

$$S_{struct} = e^{-\eta \cdot \text{len}(p_i)} \cdot \text{correlation}(Q, p_i) \quad (2)$$

where η is a hyperparameter that controls how quickly the score decreases as the path gets longer. The correlation function measures the relevance between question Q and reasoning path p_i with cosine similarity.

3.1.3 Outcome-based quality

Outcome-based quality evaluates whether the final answer from a reasoning path aligns with human preferences and factual standards. We use an outcome reward model (ORM) named SkyworkRM-Llama3.1-8B, which has been fine-tuned on human preference data, to assess each reasoning path. Since this reward model can only evaluate textual inputs, we first convert structured paths into a textual sequence of connected triplets. The textual path is then fed into the ORM to obtain an outcome-based quality score S_{RM} .

3.1.4 Overall score

The overall score S_i for each reasoning path p_i is computed by a weighted combination of scores from semantic (S_{LLM}), structural (S_{struct}), and outcome-based dimensions (S_{RM}).

$$S_i = \alpha \cdot S_{LLM} + \beta \cdot S_{struct} + \gamma \cdot S_{RM} \quad (3)$$

where α , β , and γ are the weight coefficients assigned to the semantic, structural, and outcome-based dimensions. They are used to keep each type of score within the same range, so the scores from each dimension can be considered equally and comprehensively when calculating the overall score. This overall score reflects the comprehensive quality of the path across multiple aspects and serves as the basis for weighted integration in the subsequent path fusion stage.

3.2 Multi-path Filtering and Weighted Fusion Mechanism

After obtaining overall scores for multiple reasoning paths, we propose a score-based multi-path filtering and weighted fusion mechanism to enhance the accuracy of the final answer. This mechanism consists of three parts: path filtering, weight assignment, and final answer generation. It is designed to select high-quality reasoning paths, dynamically assign fusion weights based on their scores, and ultimately guide LLMs to generate the final answer through prompts.

3.2.1 Path filtering

To reduce the influence of redundant and erroneous information in the subsequent fusion stage, we apply a threshold-based filtering strategy to remove low-quality reasoning paths. Specifically, a threshold θ is set, and all paths satisfying $S \geq \theta$ are retained to form a valid path set p_{valid} .

$$p_{valid} = \{p_i | S_i \geq \theta\} \quad (4)$$

The set of valid paths p_{valid} will then be assigned weights based on their scores in the next step.

3.2.2 Weight assignment

To better distinguish the quality differences among paths in the valid set p_{valid} , we assign each path a weight based on its score S_i . Paths with higher scores receive correspondingly greater weights, and larger score gaps result in more distinct differences in weight allocation. Specifically, we use a Softmax function with a temperature parameter T to compute the weights. For a path p_i in p_{valid} , its weight w_i is:

$$w_i = \frac{e^{\frac{S_i}{T}}}{\sum_{p_j \in p_{valid}} e^{\frac{S_j}{T}}} \quad (5)$$

where the temperature parameter T is used to control the concentration of the weight distribution, making the weighting more sensitive to variations in the scores. The weights w_i will be used in the subsequent fusion of reasoning paths and final answer generation.

3.2.3 Final answer generation

To effectively incorporate the weighted multi-path information into the LLMs for generating the final answer, we design a prompt template. This template explicitly indicates each reasoning path along with its assigned weight, making the LLMs prioritize information from higher-weighted paths. The detailed prompt template is shown in Figure 3. By explicitly embedding the paths and their weights into the input context, the template makes the LLMs recognize quality differences among reasoning paths to achieve weight-based integration and reasoning. This method enhances the accuracy and reliability of the final answer.

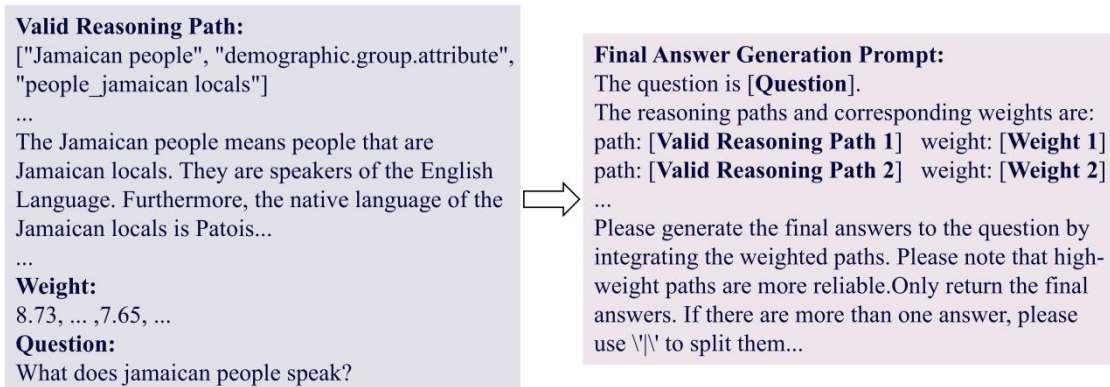


Figure 3 Prompt Template for Final Answer Generation

4 EXPERIMENTS

4.1 Experiment Setup

4.1.1 Datasets

Following previous works [3,19,20], we conduct experiments on two datasets, WebQuestionSP (WebQSP) and Complex WebQuestions (CWQ) [21, 22]. The statistics of the datasets are given in Table 1. We follow previous works (Luo et al., 2024a) to use the same train and test splits for fair comparison. The questions in WebQSP are 1-hop or 2-hop, and the questions in CWQ are 2-4 hops. The two datasets test the model's ability to understand and answer questions with multiple facts and reasoning steps. The KG for both datasets is Freebase [23].

Table 1 Statistics of Datasets

Dataset	#Train	#Test	Max #hop
WebQSP	2826	1628	2
CWQ	27639	3531	4

4.1.2 Baselines

We compare MQPF with 13 baselines grouped into 3 categories: (1) LLM only, (2) Multi-path LLM, (3) KG LLM.

(1) LLM-only methods use only LLMs for reasoning without other enhancement methods.

Qwen2-7B is one of a series of LLMs developed by the Alibaba Cloud Tongyi Qianwen team, with a parameter size of 7 billion[24].

Llama-2-7B is one of the Llama 2 series of LLMs developed by Meta AI, with a parameter size of 7 billion[25].

Llama-3.1-8B is one of the Llama 3 series of LLMs developed by Meta AI, with a parameter size of 8 billion[26].

(2) Multi-path+LLM methods prompt LLMs to generate multiple KG paths and CoTs and then generate the final answers. The reasoning LLMs for this kind of baselines are Llama-2-7B and Llama-3.1-8B.

(3) KG+LLM methods use KGs to enhance LLM reasoning.

G-Retriever retrieves the relevant nodes and edges, then constructs the relevant subgraph using the bonus Steiner tree method[27].

GRAG retrieves text subgraphs and performs soft pruning to identify relevant subgraph structures effectively, and proposes a new cue strategy[28].

SubgraphRAG generates accurate and explainable answers by efficiently retrieving relevant subgraphs from KGs and leveraging LLMs for reasoning[7].

RoG proposes a planning-search-reasoning framework, which retrieves reasoning paths from KGs to guide LLMs in reasoning[3].

GNN-RAG integrates graph neural networks (GNNs) as retrieval mechanisms to extract structured knowledge paths from KGs, which are then verbalized and fed into LLMs for answer generation[6].

4.1.3 Evaluation metrics

Following previous works [3,20,28], we use Hit@1 and the F1 score as evaluation metrics. Hit@1 checks if the ground truth exists in the generated answers. The F1 score is a harmonic average of accuracy and recall, providing a metric that balances false positives and false negatives.

4.1.4 Implementations

We choose Llama-3.1-8B-Instruct and Llama-2-7B-Chat as the reasoning LLMs. The number of CoT paths and KG paths is both set to 4 to get multiple reasoning paths. The hyperparameter η for S_{struct} is set to 1×10^{-4} , and the threshold θ for filtering is set to 2. Both of the parameters are selected based on the experimental results. We set the weight coefficients α , β and γ to ensure that the scores from each dimension are considered equally when calculating the overall score and to restrict the overall score to a range of (0 – 10). Considering that S_{LLM} is in the range of (0 – 10) and S_{struct} is in the range of (0 – 1), we set $\alpha = \frac{1}{3}$ and $\beta = \frac{10}{3}$. Because S_{RM} varies with path quality, we dynamically adjust γ based on its maximum and minimum values to map the score to a range of $(0 - \frac{10}{3})$.

4.2 Main Results

We compare our method, MQPF, to other baselines on the datasets. As Table 2 shows, MQPF performs best on CWQ and is comparable to other baselines on WebQSP. On CWQ, the F1 score and Hit@1 are 2.6% and 6.1% above the best baseline. On WebQSP, although MQPF itself performs slightly worse than RoG and SubgraphRAG, it can improve the performance of the two baselines when combined with them. The MQPF+baselines increase the F1 score of the corresponding baseline by up to 1.4% and Hit@1 by up to 1.0% on WebQSP. It shows that MQPF can effectively enhance the multi-path reasoning of the LLMs as a plug-and-play module.

It is found that the performance of Multi-path+LLM baselines is better than that of LLM only baselines. The F1 score is improved by up to 74.8% and Hit@1 by up to 92.9%. It indicates that multiple reasoning paths can enhance the coverage and accuracy of the final answers. Also, the overall performance of KG-enhanced methods (Multi-path+LLM and KG+LLM baselines) is better than that of LLM only baselines, indicating that KGs are important in MHQA tasks.

We also observe that larger LLMs do not always perform better than smaller LLMs. For the performance of Llama-2-7B and Llama-3.1-8B, Llama-2-7B has a higher F1 and Hit@1 on the WebQSP dataset, by 4.9% and 1.6%. In the CWQ dataset, its Hit@1 is still higher. This suggests that increasing the parameters does not inherently enhance the graph reasoning ability of LLMs.

Table 2 Model Performance on Two Datasets Comparing Three Categories of Methods

Category	Method	WebQSP		CWQ	
		F1 Score	Hit@1	F1 Score	Hit@1
LLM only	Qwen2-7B [24]	0.3550	0.5080	0.2160	0.2530
	Llama-2-7B [25]	0.3650	0.5640	0.2140	0.2840
	Llama-3.1-8B [26]	0.3480	0.5550	0.2240	0.2810
Multi-path+LLM	Llama-2-7B [25]	0.4625	0.7168	0.3740	0.5411
	Llama-3.1-8B [26]	0.4601	0.7137	0.3810	0.5421
KG+LLM	G-Retriever [27]	0.4674	0.6808	0.3396	0.4721
	GRAG [28]	0.5022	0.7236	0.3649	0.5018
	SubgraphRAG [7]	0.7057	<u>0.8661</u>	0.4716	0.5698
	RoG [3]	0.7080	0.8570	0.5620	0.6260
	GNN-RAG [6]	0.7130	0.8060	0.5940	0.6170
	MQPF+Llama-3.1-8B	0.6431	0.8337	0.6094	0.6640
Our method	MQPF+Llama-2-7B	0.6427	0.8340	<u>0.5990</u>	<u>0.6639</u>
	MQPF+RoG	<u>0.7145</u>	0.8636	0.5762	0.6370
	MQPF+SubgraphRAG	0.7154	0.8718	0.5027	0.5892

Note: The best results are **bolded**, and the second best results are underlined.

4.3 Ablation Study

We conduct a series of evaluations of MQPF to see which component plays a key role in the overall score, including removing S_{LLM} , removing S_{struct} , and removing S_{RM} . We can see that the performances of variables all decrease, as shown in Table 3. It indicates that every component is indispensable. Among them, removing S_{struct} drops model performance the most. On the CWQ dataset, the F1 score decreased by 4.1% and Hit@1 decreased by 3.1%. The situation is similar on WebQSP. This suggests that S_{struct} plays a more central role in the overall score. It evaluates the structural validity and information density of reasoning paths to select more coherent and concise ones, which effectively improves the accuracy of the final answers.

Table 3 Performances of Three Model Variables

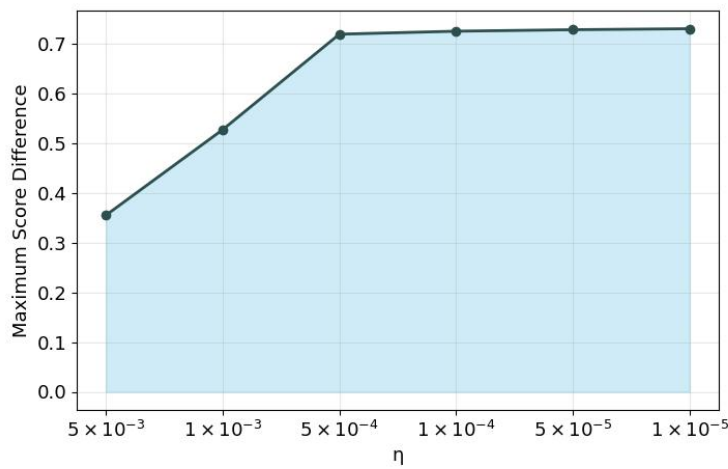
Method	WebQSP		CWQ	
	F1 Score	Hit@1	F1 Score	Hit@1
MQPF	0.6431	0.8337	0.6094	0.6640
w/o S_{LLM}	0.6271	0.8188	0.5955	0.6593
w/o S_{struct}	0.6196	0.8139	0.5846	0.6432
w/o S_{RM}	0.6293	0.8237	0.5915	0.6511

Note: The best results are **bolded**.

4.4 Analytical Experiments

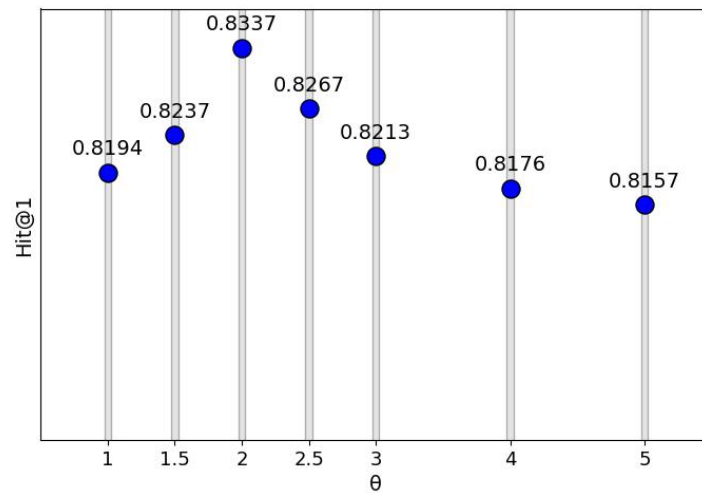
4.4.1 Structural hyperparameter analysis

For the hyperparameter for S_{struct} for textual paths η , we conduct multiple experiments with different values of it. To make the distinction between different reasoning paths greater, we set η while keeping the score differences of the reasoning paths relatively large. As shown in Figure 4, the maximum score difference continues to improve as η decreases. However, as η decreases, S_{struct} will increasingly ignore the path length and rely more on the correlation. After comprehensive consideration, we set η to 1×10^{-4} .

**Figure 4** Performances on Different Values of η

4.4.2 Threshold analysis

For the threshold θ for path filtering, we conduct multiple experiments with different values of it. As shown in Figure 5, when $\theta = 2$, Hit@1 is the highest. Therefore, the threshold θ is set to 2.

**Figure 5** Performances on Different Values of θ

5 CONCLUSION

In this paper, we propose a Multi-dimensional Quality-aware Path Fusion framework (MQPF) for path quality assessment and multi-path fusion in MHQA tasks. It introduces a multi-dimensional evaluation mechanism that quantifies path quality from semantic, structural, and outcome-based dimensions. Based on the overall scores, it filters

out low-quality paths to reduce noise and then assigns adaptive weights. Experiments show that MQPF performs comparably to baselines and can be used as a plug-and-play module to enhance the performance of other multi-path reasoning methods.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCE

- [1] Bi Z, Hajjaligol D, Sun Z, et al. Stoc-tot: Stochastic tree-of-thought with constrained decoding for complex reasoning in multi-hop question answering. 2024. DOI: <https://doi.org/10.48550/arXiv.2407.03687>.
- [2] Lee S, Shin J, Ahn Y, et al. Zero-shot multi-hop question answering via monte-carlo tree search with large language models. 2024. DOI: <https://doi.org/10.48550/arXiv.2409.19382>.
- [3] Luo L, Li Y F, Haffari G, et al. Reasoning on graphs: Faithful and interpretable large language model reasoning. International Conference on Learning Representations. 2024.
- [4] Park J, Patel A, Khan O Z, et al. Graph-guided reasoning for multi-hop question answering in large language models. 2023. DOI: <https://doi.org/10.48550/arXiv.2311.09762>.
- [5] Chen L Y, Tong P R, Jin Z M, et al. Plan-on-graph: Self-correcting adaptive planning of large language model on knowledge graphs. Advances in Neural Information Processing Systems (NeurIPS) 37, 2024. DOI: <https://doi.org/10.48550/arXiv.2410.23875>.
- [6] Mayromatis C, Karypis G. Gnn-rag: Graph neural retrieval for large language model reasoning. 2024. DOI: <https://doi.org/10.48550/arXiv.2405.20139>.
- [7] Li M, Miao S, Li P. Simple is effective: The roles of graphs and large language models in knowledge-graph-based retrieval-augmented generation. 2025. DOI: <https://doi.org/10.48550/arXiv.2410.20724>.
- [8] Li S, He Y, Guo H, et al. Graphreader: Building graph-based agent to enhance long-context abilities of large language models. 2024. DOI: <https://doi.org/10.48550/arXiv.2406.14550>.
- [9] Wang X, Wei J, Schuurmans D, et al. Self-consistency improves chain of thought reasoning in language models. 2023. DOI: <https://doi.org/10.48550/arXiv.2203.11171>.
- [10] Wang Y, Jiang B, Luo Y, et al. Reasoning on efficient knowledge paths: knowledge graph guides large language model for domain question answering. 2024. DOI: <https://doi.org/10.48550/arXiv.2404.10384>.
- [11] Bi Z, Han K, Liu C, et al. 2025. Forest-of-thought: Scaling test-time compute for enhancing llm reasoning. 2025. DOI: <https://doi.org/10.48550/arXiv.2412.09078>.
- [12] Tian Y, Song D, Wu Z, et al. Augmenting reasoning capabilities of llms with graph structures in knowledge base question answering. Findings of the Association for Computational Linguistics: EMNLP 2024, Association for Computational Linguistics, Miami, Florida, USA, 2024, 11967-11977.
- [13] Ko T W, Jiang J Y, Cheng P J. Beyond independent passages: Adaptive passage combination retrieval for retrieval augmented open-domain question answering. 2025. DOI: <https://doi.org/10.48550/arXiv.2507.04069>.
- [14] Khalifa M, Logeswaran L, Lee M, et al. Few-shot reranking for multi-hop qa via language model prompting. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Toronto, Canada, 2023, 15882-15897. DOI: [10.18653/v1/2023.acl-long.885](https://doi.org/10.18653/v1/2023.acl-long.885).
- [15] Cofala T, Astappiev O, Xion W, et al. Ragtifier: Evaluating rag generation approaches of state-of-the-art rag systems for the si-gir liverag competition. 2025. DOI: <https://doi.org/10.48550/arXiv.2506.14412>.
- [16] Yi Z, Zeng D, Ling Z, et al. Attention basin: Why contextual position matters in large language models. 2025. DOI: <https://doi.org/10.48550/arXiv.2508.05128>.
- [17] Chiang C H, Lee H Y. Over-reasoning and redundant calculation of large language models. Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 2: Short Papers). Association for Computational Linguistics, St. Julian's, Malta, 2024, 161-169.
- [18] Nayab S, Rossolini G, Simoni M, et al. Concise thoughts: Impact of output length on llm reasoning and cost. 2024. DOI: <https://doi.org/10.48550/arXiv.2407.19825>.
- [19] Wang K, Duan F, Wang S, et al. Knowledge-driven cot: Exploring faithful reasoning in llms for knowledge-intensive question answering. 2023. DOI: <https://doi.org/10.48550/arXiv.2308.13259>.
- [20] Luo L, Zhao Z, Haffari G, et al. Graph-constrained reasoning: Faithful reasoning on knowledge graphs with large language models. 2024. DOI: <https://doi.org/10.48550/arXiv.2410.13080>.
- [21] Yih W t, Richardson M, Meek C, et al. The value of semantic parse labeling for knowledge base question answering. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), Association for Computational Linguistics, Berlin, Germany, 2016, 201-206.
- [22] Talmor A, Berant J. The web as a knowledge-base for answering complex questions. Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies Volume 1 (Long Papers), Association for Computational Linguistics, New Orleans, Louisiana, 2018, 641-651.

- [23] Bollacker K, Evans C, Paritosh P, et al. Freebase: a collaboratively created graph database for structuring human knowledge. Proceedings of the 2008 ACM SIGMOD international conference on Management of data (SIGMOD '08). Association for Computing Machinery, New York, NY, USA, 2008, 1247-1250. DOI: <https://doi.org/10.1145/1376616.1376746>.
- [24] Yang A, Yang B, Hui B, et al. Qwen2 technical report. 2024. DOI: <https://doi.org/10.48550/arXiv.2407.10671>.
- [25] Touvron H, Martin L, Stone K, et al. Llama 2: Open foundation and fine-tuned chat models. 2023. DOI: <https://doi.org/10.48550/arXiv.2307.09288>.
- [26] Grattafiori A, Dubey A, Jauhri A, et al. The llama 3 herd of models. 2024. DOI: <https://doi.org/10.48550/arXiv.2407.21783>.
- [27] He X, Tian Y, Sun Y, et al. G-retriever: Retrieval-augmented generation for textual graph understanding and question answering. Advances in Neural Information Processing Systems, 2025(37): 132876-132907.
- [28] Hu Y, Lei Z, Zhang Z, et al. Grag: Graph retrieval-augmented generation. 2025. DOI: <https://doi.org/10.48550/arXiv.2405.16506>.

SEMANTIC SEGMENTATION OF SHIP HULLS FOR UNDERWATER CLEANING ROBOTS

Lu Li[#], Wei Fang[#], YiGe Shang, Zhi Zhang, MingFu Li, DaoYi Chen^{*}

Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518000, Guangdong, China.

[#]Lu Li and Wei Fang contributed equally to this work and they are both first authors.

Corresponding Author: DaoYi Chen, Email: chen.daoyi@sz.tsinghua.edu.cn

Abstract: Underwater barnacle cleaning robots require a semantic segmentation algorithm deployable onboard to effectively segment barnacles, obstacles, and background on ship hulls. This study independently constructed a dedicated ship hull semantic segmentation dataset and implemented an improved BiSeNet network. The approach integrates features from both a spatial path and a context path. Building upon the cross-entropy loss function, class weights were strategically assigned during training according to the distribution of different target categories within the dataset. This method achieved significantly high Intersection over Union (IOU) and F1 scores.

Keywords: Spatial features; Contextual features; Semantic segmentation; Ship hull dataset; Weighted loss function

1 INTRODUCTION

During prolonged oceanic transits, submerged hull surfaces inevitably develop complex biofouling communities comprising diverse marine organisms. Crustaceans within the order Cirripedia, particularly barnacles, constitute the most economically consequential fouling organisms due to their exceptional adhesive properties and rapid reproductive cycles. Barnacles form permanent calcareous attachments to hull substrates shortly after larval settlement through proteinaceous cement secreted from specialized glands. Hydrodynamic investigations confirm that extensive barnacle colonization beyond certain threshold levels substantially increases hydrodynamic drag by fundamentally disrupting boundary layer flow regimes. This hydrodynamic resistance directly compromises propulsion efficiency, necessitating considerable engine power augmentation to sustain operational speeds.

Beyond drag-related energy penalties, barnacles induce severe structural deterioration via biocorrosion pathways[1]. Their basal structures release concentrated organic acids that electrochemically compromise protective coatings. This process generates microscale electrochemical corrosion cells where barnacle-covered zones function as anodes relative to adjacent cathodic bare steel, markedly accelerating metal oxidation rates. Industry assessments verify that such biodeterioration significantly curtails vessel service lifetimes, elevates dry-dock maintenance frequency, and generates substantial latent asset depreciation not routinely accounted for in operational expenditures.

Conventional biofouling management predominantly employs hazardous diver operations or scheduled dry-docking—both approaches exhibit critical constraints. Human divers confront safety hazards and limited operational endurance, resulting in restricted cleaning coverage with inconsistent outcomes. Dry-dock procedures, while comprehensive, immobilize vessels for extended periods at considerable expense, triggering logistical disruptions throughout supply network.

Recent advancements in remotely operated and autonomous underwater vehicles have redirected industry attention toward robotic cleaning platforms[2], offering dramatic reductions in human risk exposure and substantial operational efficiency gains. However, achieving genuinely intelligent robotic cleaning requires overcoming perceptual limitations in turbid marine environments. Conventional optical imaging systems suffer severe performance degradation underwater due to light attenuation and particulate scattering, causing traditional edge detection methodologies to misclassify sedimentary particulates as biological fouling boundaries at elevated error rates[3].

This research addresses these sensory constraints through specialized semantic segmentation frameworks, enabling precise real-time differentiation of biofouling, structural obstacles, and background elements under validated turbidity conditions. This capability establishes an automated inspection-cleaning-validation cycle that drastically reduces hull maintenance durations while significantly suppressing biofouling recurrence through optimized cleaning protocols. Collectively, these innovations transform robotic systems from basic mechanized tools into intelligent marine maintenance platforms, delivering fundamental improvements in global maritime operational efficiency and environmental stewardship.

2 METHODOLOGY

2.1 Multi-Level Supervised Dual-Path Network Framework

BiSeNet is a lightweight neural network specifically designed for real-time semantic segmentation[4]. Its dual-path architecture balances spatial detail and contextual semantic information, reconciling speed and accuracy. The core design comprises a Spatial Path (SP) and a Context Path (CP). A Feature Fusion Module (FFM) merges features

extracted from these paths. To accelerate convergence, a multi-level supervision strategy was employed, incorporating four auxiliary segmentation heads into the intermediate layers of the Context Path. The auxiliary loss from each head is calculated, contributing to the total loss function, as defined in Formula (1):

$$L_{\text{total}} = L_{\text{main}} + \sum_{i=1}^4 L_{\text{aux}}^i \quad (1)$$

2.2 Weighted Cross-Entropy Loss Function

To address class imbalance (e.g., barnacle pixels being substantially fewer than background pixels), a class weighting strategy was introduced into the standard cross-entropy loss function. This elevates the learning weight for minority classes. The standard cross-entropy loss is expressed mathematically as Formula (3)[5]. The class-weighted variant is given in Formula (4), where C represents the total number of classes, y_i is the one-hot encoded ground truth label for each pixel, p_i is the predicted probability of the pixel belonging to each class, and w_i is the class weight, inversely proportional to the class's pixel frequency a_i .

$$w_i = \frac{1}{a_i} \quad (2)$$

$$L_{\text{CE}} = -\sum_{i=1}^C y_i \log(p_i) \quad (3)$$

$$L_{\text{CE}} = -\sum_{i=1}^C w_i y_i \log(p_i) \quad (4)$$

3 EXPERIMENTAL RESULTS

3.1 Dataset Description

A custom dataset simulating the underwater ship hull environment was constructed. Images were captured using an RGB camera from a simulated robotic perspective. Three target categories were annotated: Background, Barnacle Region, and Hull Obstacle. The original dataset comprised 408 images, annotated using LabelMe. Through augmentation techniques including rotation, cropping, and brightness adjustment, the dataset was expanded to 1,872 images. Pixel distribution analysis revealed: Background (Category 0) covered 52.87%, Barnacle Region (Category 1) covered 39.32%, and Hull Obstacle (Category 2) covered 7.80%. The dataset was split into a training set (1,498 images) and a test set (374 images) using a 4:1 ratio. Representative dataset samples are shown in Figs. 1 to 3.



Figure 1 Image Containing Seawater Background and Barnacle Region



Figure 2 Image Containing Cleaned Hull Surface and Barnacle Region



Figure 3 Image Containing Hull Obstacle, Barnacle Region, and Seawater Background

3.2 Analysis of Results

A significant imbalance exists in the pixel proportion of the three target classes within the dataset, as summarized in Table 1.

Table 1 Dataset Class Distribution and Description

Pixel Ratio	Description
52.87%	Clean hull areas & seawater background
39.32%	Barnacle adhesion regions
7.80%	Fins, propellers, etc.

Post-implementation of the class weighting strategy, segmentation accuracy for barnacles and obstacles markedly improved. Performance was evaluated using the F1 score and IoU metrics[6,7]. The F1 score represents the weighted harmonic mean of precision and recall Formula(5), while IoU denotes the average Intersection over Union Formula(6), A represents the predicted area, and B represents the ground truth area.

$$F1_Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (5)$$

$$IoU = \frac{A \cap B}{A \cup B} \quad (6)$$

Results before weighting application are detailed in Tables 2 and 3.

Single Scale(SS) denotes an inference approach where predictions are generated using the original image resolution or a single fixed scale, without any multi-scale augmentation.

Multi-Scale Fusion (MSF) involves aggregating predictions from multiple scaled versions of the input image, followed by fusion mechanisms such as averaging or max-voting to consolidate results across scales.

Multi-Scale Fusion with Cropping (MSFC) combines the principles of MSF with image cropping, applying multi-scale fusion to cropped sub-regions to further enhance segmentation accuracy, especially for fine-grained details in high-resolution scenarios.

Table 2 F1 Scores (Baseline - No Class Weighting)

F1 Score	ratio	ss	msf	msfc
cat 0	0.528740	0.852262	0.866208	0.871694
cat 1	0.393239	0.811935	0.845999	0.847779
cat 2	0.078021	0.4704	0.442935	0.531036
macro_F1	~	0.711532	0.718381	0.75017
micro_F1	~	0.818883	0.837874	0.845198

Table 3 IoU Scores (Baseline - No Class Weighting)

IoU	ratio	ss	msf	msfc
cat 0	0.528740	0.742559	0.763993	0.77257
cat 1	0.393239	0.68341	0.733101	0.735779
cat 2	0.078021	0.422108	0.543133	0.532094
mlous	~	0.577834	0.593854	0.623284
fw_mlous	~	0.685358	0.714432	0.72603

Results after applying class weighting are presented in Tables 4 and 5.

Table 4 F1 Scores (With Class Weighting)

F1 Score	ratio	ss	msf	msfc
cat 0	0.528739	0.872693	0.891672	0.891887
cat 1	0.393239	0.846798	0.870601	0.867972

cat 2	0.078022	0.68807	0.74849	0.75926
macro_F1	~	0.80252	0.836921	0.839706
micro_F1	~	0.85155	0.874783	0.874697

Table 5 IoU Scores (With Class Weighting)

IoU	ratio	ss	msf	msfc
cat 0	0.528739	0.77414	0.804521	0.804871
cat 1	0.393239	0.734302	0.770854	0.766741
cat 2	0.078021	0.524472	0.59807	0.611942
mIous	~	0.677638	0.724481	0.727851
fw_mIous	~	0.738994	0.775174	0.774824

Following the implementation of the class weighting strategy, significant gains were observed during single-scale evaluation: Barnacle Region (Category 1): F1 score increased from 0.811 to 0.846; IoU increased from 0.683 to 0.734. Obstacle Region (Category 2): F1 score increased from 0.470 to 0.688; IoU increased from 0.30 to 0.524. On the Jetson Orin NX embedded platform, the model inference speed reaches 24 FPS, meeting the real-time control frequency requirements.

3.3 Inference Visualization

Qualitative results generated by the trained model (with weighting) on different test image types are depicted below (see Figure 4-9), confirming effective segmentation of all three classes:



Figure 4 An Image Containing a Cleaned Hull Surface and Barnacles

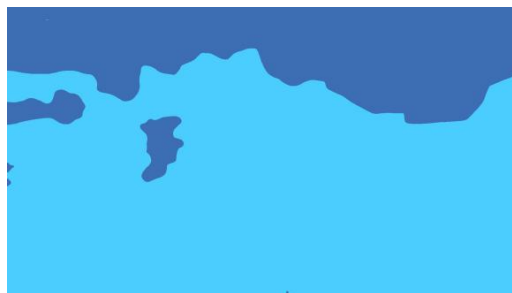


Figure 5 Segmentation Map of an Image Containing a Cleaned Hull Surface and Barnacles



Figure 6 An Image Containing a Hull Obstacle, Barnacle Region, and Seawater Background

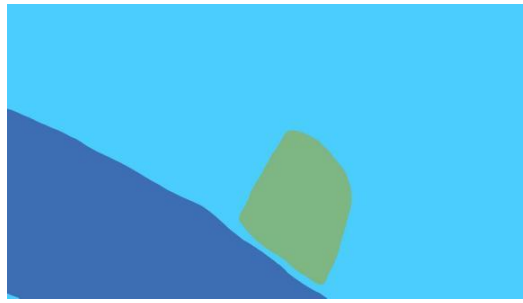


Figure 7 Segmentation Map of an Image Containing a Hull Obstacle, Barnacle Region, and Seawater Background



Figure 8 An Image Containing Seawater Background and Barnacle Region



Figure 9 Segmentation Map of an Image Containing Seawater Background and Barnacle Region

4 CONCLUSION

This research confronts the fundamental technological barrier of precise hull-adhering contaminant segmentation. This precision is essential for deploying truly intelligent underwater cleaning robotics across global maritime operations. It proposes an architecturally refined lightweight semantic segmentation paradigm, anchored in an optimized BiSeNet framework. This framework implements a sophisticated bilateral feature fusion mechanism, which coordinates high-resolution spatial pathways with deep contextual streams. This coordination overcomes inherent precision-context trade-offs in turbid subaquatic environments. The paradigm also strategically deploys dynamically class-weighted loss functions. These functions probabilistically recalibrate learning priorities to counteract extreme categorical imbalances. These imbalances exist among densely aggregated barnacle colonies, sparse structural obstructions, and heterogeneous background interfaces.

Concurrently, it integrates hierarchical multi-level supervision modules. These modules inject auxiliary gradients across complementary feature hierarchies to stabilize convergence pathways and amplify discriminative feature representation capabilities. Collectively, these innovations enable unprecedented pixel-accurate real-time identification of biofouling distributions and hazardous protrusions. This identification occurs across complex curved hull surfaces under real-world visibility constraints. The system sustains real-time computational efficiency, which is essential for continuous hull scanning during robotic transit. Thereby, it establishes an industrial-grade perception backbone for autonomous underwater maintenance systems. These systems are capable of executing millimeter-precision cleaning trajectories around thrusters, sensors, and anodes with zero collision tolerance thresholds. This capability fundamentally redefines vessel husbandry paradigms by transitioning from reactive labor-intensive scrubbing toward strategically optimized biofouling management systems. These new systems directly mitigate billions in global shipping fuel penalties and coating degradation costs annually. They also provide environmentally sustainable alternatives to toxic anti-fouling chemical treatments.

Future evolutionary research vectors explicitly target multi-physics convolutional architectures intrinsically resilient to severe scattering noise prevalent in harbor silting conditions through wavelength-adaptive optical modeling coupled with hybrid optical-acoustic sensor fusion frameworks that synergize RGB imaging with bathymetric laser scanning and ultrasonic thickness mapping to circumvent spectral attenuation limitations. They also advance predictive maintenance cognition through temporal consistency networks that correlate sequential segmentation maps into 4D hull corrosion

progression models facilitating preventative structural interventions and lifecycle durability forecasts. Together, these developments collaboratively position next-generation marine robotic platforms as comprehensive vessel health guardians reconciling operational safety compliance, economic sustainability objectives, and oceanic ecosystem preservation imperatives within unified autonomous service frameworks[8-10].

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

- [1] Trif László, Abdul Shaban, Judit Telegdi. Electrochemical and surface analytical techniques applied to microbiologically influenced corrosion investigation. *Corrosion Reviews*, 2018, 36(4): 349-363.
- [2] Lander Karl. Proactive Cleaning, Maintenance and Inspection of Vessel Hulls via Autonomous Robots. Paper presented at the AMPP Annual Conference + Expo, Denver, Colorado, USA, 2023.
- [3] Huang Q, Huang J. Comprehensive review of edge and contour detection: From traditional methods to recent advances. *Neural Computing and Applications*, 2025, 37(4): 2175-2209.
- [4] S Kim, K Jo. BiSeNet with Depthwise Attention Spatial Path for Semantic Segmentation. 2022 International Workshop on Intelligent Systems (IWIS), Ulsan, Korea, Republic of. 2022, 1-4. DOI: 10.1109/IWIS56333.2022.9920717.
- [5] N Das, S A Begum. An Empirical Study of Loss Functions for Aspect Category Detection in Imbalanced Data Scenario. 2025 10th International Conference on Signal Processing and Communication (ICSC), Noida, India. 2025, 247-252. DOI: 10.1109/ICSC64553.2025.10968230.
- [6] G R S N Kumar, R S Sankuri, S P K Karri. Multi Scale aided Deep Learning model for High F1-score classification of fundus images based Diabetic Retinopathy and Glaucoma. 2023 International Conference on Computer, Electronics & Electrical Engineering & their Applications (IC2E3), Srinagar Garhwal, India. 2023, 1-6. DOI: 10.1109/IC2E357697.2023.10262434.
- [7] Y Chen, X Li, Z He, et al. IoU_MDA: An Occluded Object Detection Algorithm Based on Fuzzy Sample Anchor Box IoU Matching Degree Deviation Aware. *IEEE Access*, 2024(12): 47630-47645. DOI: 10.1109/ACCESS.2024.3375109.
- [8] Hennigh O, Narasimhan S, Nabian M A, et al. NVIDIA SimNet™: An AI-accelerated multi-physics simulation framework. *International conference on computational science*. Cham: Springer International Publishing. 2021, 447-461.
- [9] Ferreira I O, Andrade L C, Teixeira V G, et al. State of art of bathymetric surveys. *Boletim de Ciências Geodésicas*, 2022, 28(01): e2022002.
- [10] Wang X, Lin M, Li J, et al. Ultrasonic guided wave imaging with deep learning: Applications in corrosion mapping. *Mechanical Systems and Signal Processing*, 2022(169): 108761.

AHP-TOPSIS-BASED ASSESSMENT FRAMEWORK FOR CYBERSECURITY POLICY EFFECTIVENESS: EMPIRICAL EVIDENCE FROM GCI INDICATORS

JiaWei Wen*, Sai Yang, ChenGxin Zhang

Zijin School of Geology and Mining Fuzhou University, Fuzhou University, Fuzhou 350108, Fujian, China.

Corresponding Author: JiaWei Wen, Email: 15083450070@163.com

Abstract: With the acceleration of global digitalization, cybercrime has exhibited transnational and covert characteristics, posing severe challenges to national cybersecurity. To scientifically evaluate the cybersecurity level of various countries and identify effective policy pathways, this study proposes a comprehensive evaluation model integrating the Analytic Hierarchy Process and Technique for Order Preference by Similarity to Ideal Solution. Based on the five pillars of the International Telecommunication Union Global Cybersecurity Index, an evaluation system comprising 15 sub-indicators is constructed. Indicator weights are determined through expert scoring, and the TOPSIS algorithm is employed to conduct quantitative ranking and grading of the cybersecurity levels of 27 countries. The results show that there is a significant positive correlation between cybersecurity level and comprehensive national strength. Technologically advanced countries such as Finland and the United States are rated as "Excellent", while less developed countries such as Zimbabwe fall into the "Weak" category. Furthermore, Pearson correlation analysis on U.S. cybercrime data and demographic characteristics reveals that internet usage, employment to population ratio of men aged 15-24 and GDP exhibit significant statistical associations with cybercrime incidence. This study provides a data-driven decision-making framework for policymakers to optimize cybersecurity strategies and verifies the core role of synergistic governance of law and technology.

Keywords: AHP-TOPSIS; National cybersecurity level assessment; GCI; Cybercrime prevention and control

1 INTRODUCTION

In recent years, the rapid development of information technology has profoundly promoted global connectivity. While significantly enhancing social productivity, it has also amplified cybersecurity risks. Cyber threats exhibit strong transnationality and concealment. To address systemic risks, countries around the world are actively formulating cybersecurity strategies. Therefore, constructing a scientific and quantifiable national cybersecurity level assessment system is of great significance for identifying weak links, optimizing resource allocation, and improving global cyber resilience. This study intends to integrate the Analytic Hierarchy Process (AHP) and Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) to establish a comprehensive evaluation model. In existing studies, Ashour M and Mahdiyar A pointed out that the application of the AHP method in multi-criteria decision-making has been continuously expanding, with significant advantages especially in risk assessment and resource allocation, however, current research has paid insufficient attention to the complex and multi-dimensional field of cybersecurity, and its applicability in dynamic cyber threat environments remains to be improved [1]. To this end, Alhakami combined fuzzy AHP with TOPSIS to construct a comprehensive assessment framework, which was used to evaluate the effectiveness of intrusion detection methods under Gen V multi-vector attacks [2]. Nevertheless, Wang et al. argued that although AHP and TOPSIS perform well in determining indicator weights, existing studies have not fully explored the impact of multi-source information such as expert trust information and corporate preference information on assessment accuracy [3]. In addition, Yu proposed a cybersecurity situation assessment method based on a fusion model, which improved the quality and efficiency of feature extraction but failed to fully cover the specific cybersecurity needs of the Industrial Internet of Things (IIoT) [4]. This paper proposes a new model combining AHP and TOPSIS methods, which aims to provide comprehensive support for the scientific assessment of national cybersecurity levels.

The main contributions of this paper include: 1) Developed a comprehensive evaluation model based on AHP-TOPSIS, realizing the quantitative assessment of national cybersecurity levels. 2) Innovatively deconstructs the five pillars of the Global Cybersecurity Index (GCI) into 15 sub-indicators and resolves the coupling among the indicators through the AHP. 3) Revealed multidimensional correlations between cybercrime and demographic characteristics through Pearson correlation analysis, providing empirical evidence for targeted policy formulation.

The structure of this paper is as follows: The first part is the introduction, which introduces the research background, current research status, and contribution points; The second part is related theories, which expounds on AHP, TOPSIS and Pearson; The third part is model establishment and analysis, which describes in detail the model ideas and establishment process; The fourth part is model conclusions, which summarizes the research results; The fifth part is the conclusion, which expounds on the research significance and future prospects.

2 RELATED THEORIES

The AHP is a systematic decision-making method. It decomposes complex problems into multiple levels and factors, then conducts pairwise comparisons of these factors through expert judgments to construct a judgment matrix, thereby determining the weights of each factor and ultimately providing decision support. The core of the AHP algorithm lies in integrating qualitative analysis with quantitative analysis: through a hierarchical structure, it breaks down complex decision-making problems into multiple comparable subproblems, and determines the relative importance of each factor based on expert experience [5]. This method is particularly applicable to decision-making problems that are difficult to directly subject to quantitative analysis, such as evaluating the relative importance of different threat factors in cybersecurity risk assessment [6].

The TOPSIS algorithm is a commonly used multi-criteria decision analysis method. Its core idea is to construct the ideal solution and the negative ideal solution, then calculate the distances between each alternative and the ideal solution as well as the negative ideal solution, and determine the ranking of each alternative in terms of superiority or inferiority based on these distances. The basic steps of the TOPSIS algorithm include: constructing the decision matrix, normalizing the decision matrix, determining the ideal solution and the negative ideal solution, calculating the distances between each alternative and the ideal solution as well as the negative ideal solution, calculating the relative closeness of each alternative, and finally ranking the alternatives based on the relative closeness [5]. The TOPSIS algorithm boasts advantages such as clear conceptualization, simplicity in calculation, and effectiveness in handling multi-criteria decision-making problems. In the field of network security, this algorithm can be applied to evaluate the effectiveness of different security solutions. For instance, it enables the selection of the optimal security solution by considering multiple indicators of security schemes, including cost, performance, and reliability [7].

The Pearson correlation coefficient (PCC) is a commonly used statistical measure of the linear relationship between two variables. Its core concept is to quantify the strength of the positive or negative change in one variable when the other variable changes. It determines the correlation by calculating the ratio of the covariance between two variables to the product of their standard deviations. The result ranges from -1 to +1, where +1 indicates perfect positive linear correlation, -1 indicates perfect negative linear correlation, and 0 indicates no linear correlation [8-10]. The Pearson correlation coefficient has a wide range of applications, playing a significant role not only in basic statistical analysis but also being integrated into various complex decision-making, prediction, and machine learning models [11-14].

3 EXPERIMENTS

National cybersecurity level is related to multiple factors. Given the existence of assessment tools such as the International Telecommunication Union (ITU)'s GCI, we adopted the five indicators and 15 sub-indicators from this research to construct a national cybersecurity level evaluation index system [15]. Based on this information, we ultimately employed a combined method of AHP and TOPSIS to evaluate a country's cybersecurity level. When analyzing complex decision-making problems, the AHP decomposes the problem into different levels, including the goal layer, criterion layer, and alternative layer. Subsequently, a judgment matrix is constructed by comparing the relative importance of elements within the same level, and the weights of each element are calculated. Finally, ranking and decision-making are performed based on the weights and the scores of the alternative layer.

Constructing the judgment matrix: By reviewing relevant literature, this study obtained expert ratings for the five indicators in the main criterion layer. The rating result for each indicator is the average of the expert ratings, from which the final ratings are derived [16], and then the judgment matrix is determined as shown in Table 1. Among them, F1: Legal, F2: Technical, F3: Organizational, F4: Capacity development, F5: Cooperation.

Table 1 AHP Judgment Matrix of Main Criteria Layer Based on Expert Ratings

	F1	F2	F3	F4	F5
F1	1	9	2	8	6
F2	1/9	1	1/6	1/2	1/3
F3	1/2	6	1	3	2
F4	1/8	2	1/3	1	1/2
F5	1/6	3	1/2	2	1

Calculation of eigenvectors: For the judgment matrix R , compute its largest eigenvalue. Eigenvector v is a nonzero vector satisfying the following equation:

$$Rv = \lambda v \quad (1)$$

The eigenvalues are obtained by solving the characteristic equations:

$$\det(R - \lambda I) = 0 \quad (2)$$

For each eigenvalue λ , substituting into the matrix equation:

$$(R - \lambda I)v = 0 \quad (3)$$

By solving this system of linear equations, we obtain the corresponding eigenvectors v . Normalize the feature vectors so that the weights sum to 1:

$$w = \frac{v}{\sum_{i=1}^n v_i} \quad (4)$$

Calculate Consistency Ratio: Calculation of consistency indicators CI the formula is as follows

$$CI = \frac{\lambda_{max} - n}{n - 1} \quad (5)$$

Next, the consistency ratio is calculated, with its formula as follows:

$$CR = \frac{CI}{RI} \quad (6)$$

If $CR < 0.1$ then the consistency of the judgment matrix is better. If $CR > 0.1$ then the judgment matrix needs to be readjusted. The calculated CR value is 0.998, which proves that the consistency of the judgment matrix is acceptable.

For the word criterion layers S1, S2, S3S15, we use the same method to figure out their weights. The weights of the criterion and sub-criterion layers are used to obtain the combined weights of the factors. For a given criterion, its final weight is:

$$w(C_k) = \sum_{i=1}^{m_k} w(C_k, S_i) \times w(S_i) \quad (7)$$

The TOPSIS algorithm is then used to score and rank each country [17]. For each country, normalize the ratings in each sub-criterion by dividing the rating by the maximum value of that sub-criterion across all countries:

$$X_{ij} = \frac{x_{ij}}{\max(x_{ij})} \quad (8)$$

Apply the aggregated weights of the sub-criteria to the normalized decision matrix:

$$V_{ij} = X_{ij} \times w_j \quad (9)$$

Determine the ideal and anti-ideal solutions:

$$A^+ = (\max(V_{ij})) \quad (10)$$

$$A^- = (\min(V_{ij})) \quad (11)$$

Compute the Euclidean distance of each country from the ideal and anti-ideal solutions:

$$D_i^+ = \sqrt{\sum_{j=1}^m (V_{ij} - A_j^+)^2} \quad (12)$$

$$D_i^- = \sqrt{\sum_{j=1}^m (V_{ij} - A_j^-)^2} \quad (13)$$

Calculate the relative closeness of each country to the ideal solution:

$$C_i = \frac{D_i^-}{D_i^+ + D_i^-} \quad (14)$$

Next, this study analysed the correlation between cybercrime and demographic data in the United States between 2000 and 2022. Since demographic characteristics have many kinds, so we firstly choose many kinds of demographic characteristics data, and based on the number of cybercrimes under different years, we conduct Spearman correlation analysis on them. Since the crime situation in the United States is characterized by a large amount of data and volatility relative to the crime situation in other countries, and the number of crimes in other countries is often in a state of zero relatively over a period of time, which is not conducive to data analysis, we have chosen the U.S. data for correlation analysis here.

This study uses Pearson's product-moment correlation coefficient to quantify the linear correlation between demographic indicators and cybercrime rates. For a bivariate dataset containing $n=23$ observations, the correlation coefficient r is calculated using the following formula:

$$r_{XY} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (15)$$

To test the statistical significance of r , construct the t-statistic:

$$t = r \sqrt{\frac{n-2}{1-r^2}} \sim t(n-2) \quad (16)$$

4 RESULTS

This study employs the AHP-TOPSIS model to evaluate the cybersecurity levels of 27 countries, with its key findings manifested in two dimensions:

Characteristics of indicator weight distribution: The weights of the five pillars determined by the model via the AHP algorithm show significant differences (Figure 1). The legal dimension dominates absolutely with a weight of 52.32%, among which personal data protection (PP, 33.14%) and effective data protection regulations (DPF, 13.63%) are key sub-indicators. The technical dimension has the lowest weight (4.33%), with the adoption of cybersecurity standard frameworks (FCS, 0.42%) among its sub-indicators showing a weak contribution. The organizational (24.17%), cooperative (11.9%), and capacity-building (7.28%) dimensions together form a secondary support layer. This weight distribution indicates that cybersecurity levels are mainly driven by legal completeness, with the marginal utility of technical means being limited.

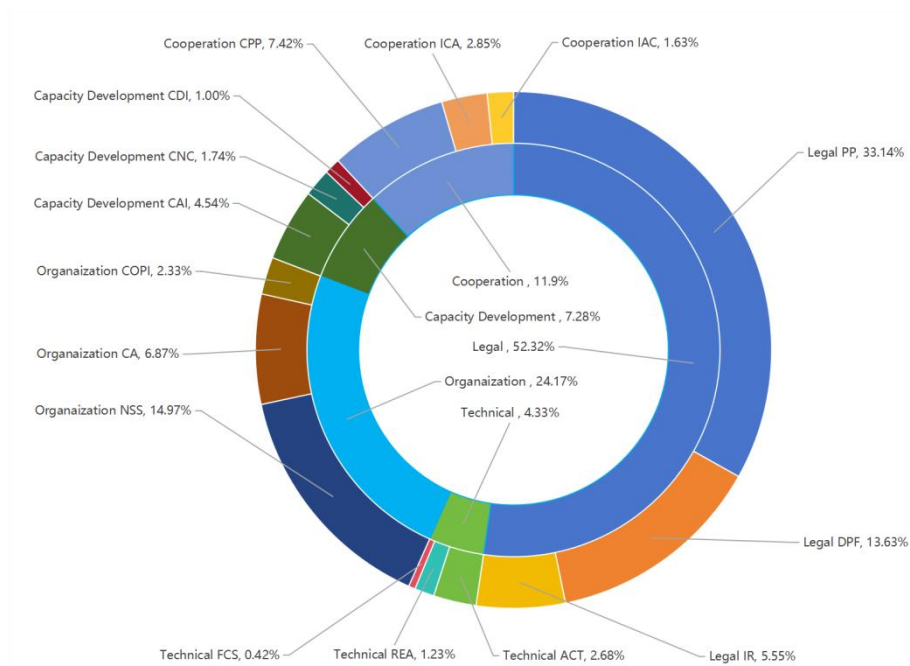


Figure 1 Indicator Weights

Polarization of national scoring grades: The ranking of C_i values based on the TOPSIS algorithm (Figure 2) reveals a polarization in the cybersecurity levels of the 27 countries: High-score group ($C_i \geq 0.8$): 13 countries including Finland (0.91), the United States (0.89), and Singapore (0.87), with their legal dimension scores all exceeding 0.9 (out of 1) and organizational dimension scores >0.85 ; Low-score group ($C_i \leq 0.2$): 10 countries including Zimbabwe (0.15), Niger (0.17), and Equatorial Guinea (0.18), with their legal dimension scores all below 0.2 and technical dimension scores less than 0.1.

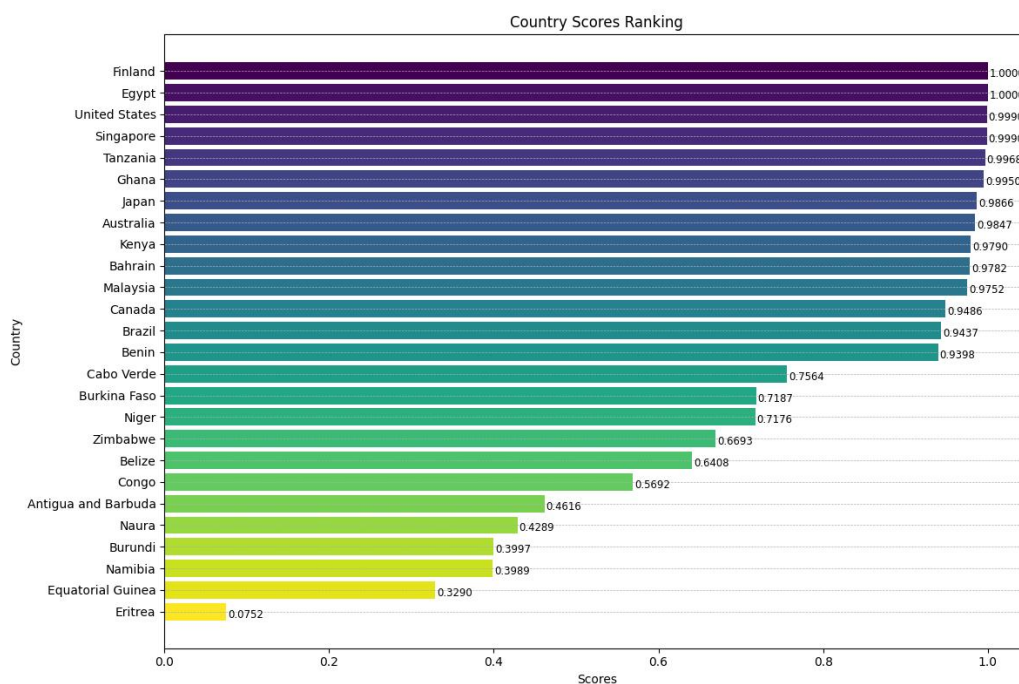


Figure 2 Score Ranking

The heat map of Pearson's correlation analysis is shown in Figure 3, and some demographic characteristics t-test results are shown in Table 2. The result revealed that a statistically robust negative correlation between youth employment rates (aged 15–24) and cybercrime incidence, where higher workforce engagement correlates with reduced criminal activity. Concurrently, internet usage exhibits a strong positive association, confirming that expanded digital access inherently broadens attack surfaces. Gross Domestic Product demonstrates a paradoxical role: while economic advancement enables technological resilience, it simultaneously elevates cyber-risk exposure through increased digitalization. Critically, these demographic and economic dynamics are subordinate to legal infrastructure—which

holds 52.32% weight in our AHP-TOPSIS framework—as evidenced by nations like Finland maintaining low cybercrime despite high internet/GDP metrics through stringent data protection laws (PP: 33.14% sub-weight). Other variables, including inflation and educational attainment, lack significant explanatory power. Thus, effective cybersecurity policy must prioritize youth employment stabilization and calibrated digital inclusion, all anchored by legislative rigor.

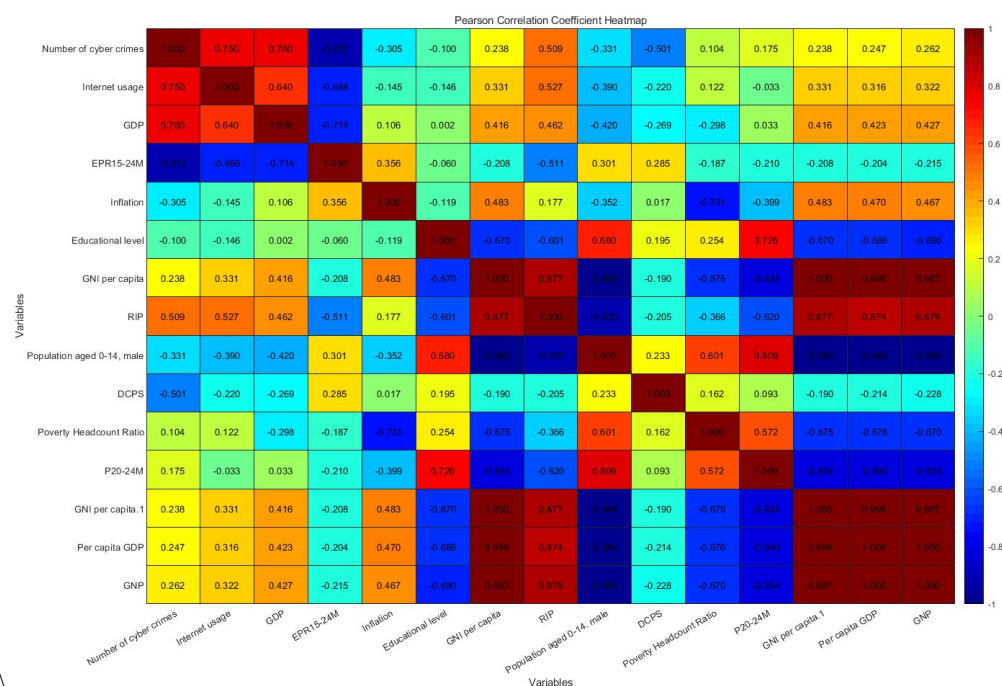


Figure 3 Pearson Correlation Analysis Heat Map

Table 2 Some Demographic Characteristics t-Test Results

Variable	r	t	Significance level
EPR15-24M	-0.912	-10.19	1%
Internet usage	0.750	5.20	1%
GDP	0.760	5.36	1%
Inflation	-0.305	-1.47	Not significant
Educational level	-0.100	-0.46	Not significant

5 CONCLUSIONS

This study innovatively constructs an AHP-TOPSIS comprehensive evaluation model to achieve a quantitative assessment of national cybersecurity levels, effectively addressing the shortcomings of existing research in adapting to multi-dimensional complex cybersecurity environments. The model is based on the five pillars of the International Telecommunication Union's Global Cybersecurity Index, innovatively decomposing them into 15 sub-indicators. Expert scores were used to determine the weights, and the TOPSIS algorithm was applied to rank and classify the cybersecurity levels of 27 countries. The results indicate that cybersecurity levels are significantly positively correlated with a country's overall strength, with the legal dimension accounting for over half of the weighting, making it the core factor influencing cybersecurity, while the marginal utility of technical measures is relatively limited. Further Pearson correlation analysis reveals that factors such as youth employment rates, internet usage rates, and GDP are significantly associated with cybercrime rates in the United States. This study not only provides data-driven decision support for policymakers but also validates the critical role of legal and technical collaborative governance in cybersecurity. Future research could consider incorporating more country samples and dynamic indicators to further optimize the timeliness and applicability of the evaluation system.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

- [1] Ashour M, Mahdiyar A. A comprehensive State-of-the-art survey on the recent modified and hybrid analytic hierarchy process Approaches. *Applied Soft Computing*, 2024, 150: 111014.

- [2] Alhakami W. Evaluating modern intrusion detection methods in the face of Gen V multi-vector attacks with fuzzy AHP-TOPSIS. *PloS one*, 2024, 19(5): e0302559-e0302559.
- [3] Wang K, Zhang M, Hong Y, et al. Airborne Network Information Security Risk Assessment Method Based on Improved STPA-TOPSIS. *Aerospace*, 2025, 12(5): 442-442.
- [4] Yunhao Y. A network security situation assessment method based on fusion model. *Discover Applied Sciences*, 2024, 6(3).
- [5] Jianjun Z. A new scoring function in multi-criteria decision-making based on vague set. 2024.
- [6] Alevizos L, Ta VT. Threat-Informed Cyber Resilience Index: A Probabilistic Quantitative Approach to Measure Defence Effectiveness Against Cyber Attacks. 2024.
- [7] Mayra M, Chunming W, Walter F. Adversarial examples: A survey of attacks and defenses in deep learning-enabled cybersecurity systems. *Expert Systems With Applications*, 2024, 238(PE).
- [8] Ramduny J, Kelly C. Connectome-based fingerprinting: Reproducibility, precision, and behavioral Prediction. *Neuropsychopharmacology*, 2024, 50(1): 114–123.
- [9] Iglesias-Martínez ME, Castro-Palacio JC, Scholkmann F, et al. Correlations between Background Radiation inside a Multilayer Interleaving Structure, Geomagnetic Activity, and Cosmic Radiation: A Fourth Order Cumulant-based Correlation Analysis. *arXiv*, 2020.
- [10] Wang J, Zheng N. Measures of Correlation for Multiple Variables. *arXiv*, 2014(2014).
- [11] Ho LH, Lin YL, Chen TY. A Pearson-like correlation-based TOPSIS method with interval-valued Pythagorean fuzzy uncertainty and its application to multiple criteria decision analysis of stroke rehabilitation Treatments. *Neural Computing and Applications*, 2019, 32(12): 8265–8295.
- [12] Zeng S, Luo D, Zhang C, et al. A Correlation-Based TOPSIS Method for Multiple Attribute Decision Making with Single-Valued Neutrosophic Information. *International Journal of Information Technology & Decision Making*, 2020, 19(01): 343–358.
- [13] Jin Y, Wu H, Sun D, et al. A Multi-Attribute Pearson's Picture Fuzzy Correlation-Based Decision-Making Method. *Mathematics*, 2019, 7(10): 999.
- [14] Zhou Q, Ma Y, Xing Z, et al. Pearson correlation Coefficient-guided large-scale fuzzy cognitive maps learning Algorithm. *Fuzzy Sets and Systems*, 2025, 519: 109523.
- [15] International Telecommunication Union. *Global Cybersecurity Index 2024: 5th Edition*. Geneva: ITU Publications, 2024: 2.
- [16] Bruggemann R, Koppatz P, Scholl M, et al. Global Cybersecurity Index (GCI) and the Role of its 5 Pillars. *Social Indicators Research*, 2021(prepublish): 1-19.
- [17] Subrata C. TOPSIS and Modified TOPSIS: A comparative analysis. *Decision Analytics Journal*, 2022, 2.

DESIGN OF A PHOTOVOLTAIC TRACKING SYSTEM CONTROL SYSTEM

MingMao Gong*, XingHui Liu

School of Electronic Information Engineering, Sichuan Technology and Business University, Chengdu 611745, Sichuan, China.

Corresponding Author: MingMao Gong, Email: fengxu0217@126.com

Abstract: With the continuous advancement of the "dual carbon" goals, improving the efficiency of solar energy utilization and reducing the cost of power generation have become key directions for technological innovation in the photovoltaic industry. This paper, starting from practical application needs, designs a photovoltaic tracking support system based on the CH31. The system is built on the Alibaba Cloud IoT platform, combined with 4G wireless communication technology, and uses the MQTT communication protocol to complete data upload and command download. The system integrates tilt sensors, light sensors, wind speed and direction sensors, and rain and snow sensors, and uses GPS to obtain longitude and latitude and RTC time. Using astronomical algorithms, the system achieves forward and reverse tracking of the photovoltaic system. To facilitate user management, the system includes a manual mode, which allows for the adjustment of the photovoltaic panel tilt angle through buttons or the cloud, at which point the tracking function is disabled. Combined with meteorological sensors, when extreme weather is detected, the photovoltaic panel is set to a fixed tilt angle mode to achieve self-protection.

Keywords: CH32; Photovoltaic tracking system; MQTT; Astronomical algorithm

1 INTRODUCTION

As the global energy structure accelerates toward clean and low-carbon transformation, solar energy, with its abundant reserves, renewable, and pollution-free advantages, has become a core force in replacing traditional fossil fuels. In the rapid development of the photovoltaic industry, maximizing the conversion efficiency of solar energy and reducing the cost per kilowatt-hour (kWh) have always been core goals of technological innovation.[1] The photovoltaic tracking system is one of the methods that can effectively improve the utilization of solar energy. The system uses GPS data combined with astronomical algorithms to calculate the solar altitude angle and azimuth angle, aligning the photovoltaic panel perpendicularly to the incident solar radiation, thereby achieving forward tracking of the photovoltaic system[2]. To prevent shading caused by excessive tilt angles of the photovoltaic panel during sunrise and sunset and to facilitate user management, the system introduces an inverse tracking mode and a manual mode based on forward tracking. In the manual mode, users can adjust the tilt angle of the photovoltaic panel through the built IoT platform or local buttons, and the tracking function is disabled in this mode. Additionally, the system uses multiple sensors to detect external weather conditions, and in extreme environments, the system activates a self-protection mode, setting the photovoltaic panel to a fixed tilt angle[3]. The attitude sensor continuously detects the tilt angle of the photovoltaic panel in real-time (September 15, 2025), achieving closed-loop control of the photovoltaic panel angle, thereby maximizing the capture of solar energy.

2 DESIGN SCHEME FOR PHOTOVOLTAIC TRACKING SYSTEM

The overall architecture of the system includes power management, sensor data acquisition, drive control module, and IoT platform. The system inputs a 12V power supply and provides stable 5V and 3.3V power to the system through two voltage reduction circuits, ensuring that all modules work properly. In the design of the photovoltaic tracking system, GPS is selected to obtain the location information of the device, and the latitude, longitude, and RTC time data are analyzed. Astronomical algorithms are used to automatically calculate the solar altitude angle and azimuth angle. Based on the calculated angle and tracking mode, the motor is controlled to rotate, thereby driving the photovoltaic panel. In addition, an attitude sensor is used to detect whether the actual angle of the photovoltaic panel matches the theoretical angle and achieve angle feedback. At the same time, rain and snow sensors, lighting sensors, and wind speed and direction sensors are introduced to sense the local weather and lighting conditions in real time. When detecting abnormal external environment or light intensity values below the threshold, the photovoltaic panel switches to a fixed mode. The system is set up for remote control of the Internet of Things and local control of buttons, achieving manual mode of the system. In this mode, the tracking function is turned off. Simultaneously upload data to the cloud platform through the 4G module to achieve remote monitoring.

2.1 Tracking Mode

On a day with good weather conditions, the large tilt angle of the photovoltaic panels calculated at sunrise and sunset

can cause the rear photovoltaic panels to be obscured by the shadows of the front row. Therefore, the system adopts reverse tracking mode at sunrise and sunset, and forward tracking mode at other times. The system obtains the latitude and longitude of its location and RTC time through GPS, and uses astronomical algorithms to calculate the solar altitude angle α , declination angle δ , time angle ω , and azimuth angle λ [4].

Solar altitude angle α : The angle between the direct sunlight and the plane of the observation point. The calculation formula is:

$$\alpha = \arcsin(\sin \delta * \sin \varphi + \cos \delta * \cos \varphi * \cos \omega) \quad (1)$$

Among them, δ is the solar declination angle, φ is the latitude of the measured earth, and ω is the time angle.

Solar declination angle δ : The angle between the incident light of the sun and the equatorial plane. The calculation formula can be expressed as:

$$\delta = 0.3723 + 23.2567 \sin \theta + 0.1149 \sin 2\theta - 0.1712 \sin 3\theta - 0.785 \cos \theta + 0.3656 \cos 2\theta + 0.201 \cos 3\theta \quad (2)$$

θ represents the solar angle, and its calculation formula is:

$$\theta = \frac{2\pi (N - N_0)}{365.2422} \quad (3)$$

N represents the product day, which is the number of days from January 1st of each year to the calculation day, and distinguishes between normal and leap years. For example, the product day of December 31st in a normal year is 365, while in a leap year it is 366; N_0 is the correction constant for the product of days, with an approximate value of 79.67. The specific expression is:

$$N_0 = 79.6764 + 0.2422 * (\text{Year} - 1985) - \text{INT}\left(\frac{\text{Year} - 1985}{4}\right) \quad (4)$$

Time angle ω : The angle between the incident sunlight and the local radial plane of the Earth. The calculation formula is: $\omega = (T_s - 12) * 15^\circ$.

The calculation formula for T_s is: $T_s = T_0 \pm \frac{120 - J}{15} + \frac{E}{60}$, T_0 is Beijing time, J is the longitude of the measurement location, taken as + for longitude west and - for longitude east; E is the time difference. The time difference E can be expressed as:

$$E = 0.0028 - 1.9857 \sin \theta + 9.059 \sin 2\theta - 7.0924 \cos \theta - 0.6882 \cos 2\theta \quad (5)$$

After calculating the above data, if the system is in forward tracking mode, the schematic diagram of the solar altitude angle α , photovoltaic panel inclination angle β , and photovoltaic panel theoretical tracking angle γ are shown in Figure 1.

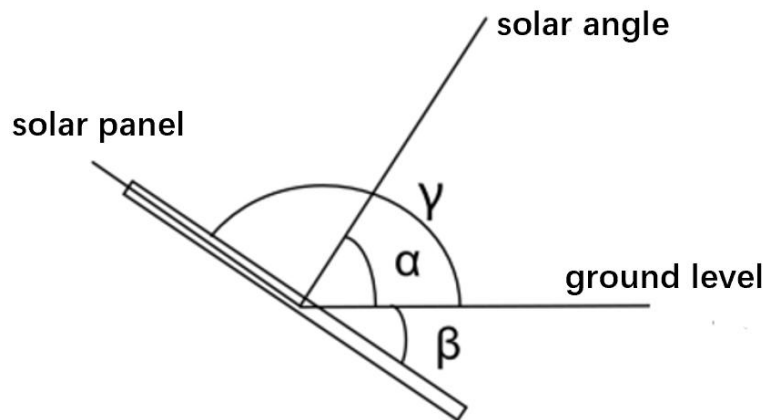


Figure 1 Schematic Diagram of Angle in Forward Tracking Mode

When the system is in reverse tracking mode, first calculate the minimum distance $D = d \sin \lambda + B \sin \alpha$ between adjacent photovoltaic supports in the east-west direction, where d is the projection length of direct sunlight between adjacent photovoltaic supports; B is the width of the photovoltaic panel. And its projection length d can be expressed as:

$$d = \frac{B \sin \sigma}{\tan \tau}, \text{ It's shown in Figure 2.}$$

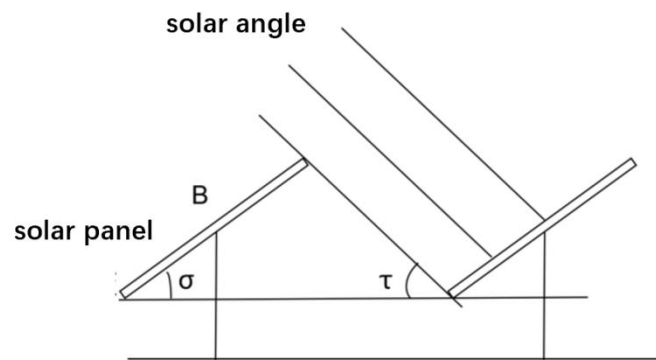


Figure 2 Schematic Diagram of Reverse Tracking Mode Angle

2.2 Fixed Mode

When the wind speed and direction sensor of the system detects that its wind speed value exceeds the threshold, in order to prevent damage to the photovoltaic bracket, the photovoltaic panel switches to a fixed mode, that is, the photovoltaic panel rotates to the windward side at an angle of 5° . At this time, the theoretical tracking angle is 175° or 185° , and the photovoltaic panel stops tracking until the wind speed is detected to be normal; When the rain and snow sensor of the system detects heavy rainfall or snow, in order to prevent snow accumulation on the photovoltaic panel from causing damage to the photovoltaic bracket or rainwater erosion of the photovoltaic panel, the photovoltaic panel switches to fixed mode, that is, the photovoltaic panel rotates to 135° and remains unchanged until the environment returns to normal; When the light sensor detects that the current ambient light intensity value is below the threshold, the system switches to a fixed mode, and the photovoltaic panel maintains 185° to reduce the power consumption generated when the motor drives the photovoltaic panel to rotate.

2.3 Manual Mode

For the convenience of user management, installation, and debugging, the system has added a manual mode. Users can control the rotation of the photovoltaic panel through local system integration buttons, or issue instructions through the Internet of Things platform to temporarily adjust the tilt angle of the photovoltaic panel. In this mode, the photovoltaic tracking function is turned off to prevent functional conflicts; Dust can have a significant impact on the performance of photovoltaic panels [5]. When the amount of dust reaches 12.64g/m^2 , its power generation will decrease by about 20%; Due to the large size of photovoltaic panels in the application scenario of photovoltaic tracking brackets, the system mode needs to be switched to manual mode when cleaning the photovoltaic panels, in order to adjust the tilt angle of the photovoltaic panels and facilitate cleaning and maintenance by cleaning personnel[6].

3 HARDWARE CIRCUIT DESIGN

The photovoltaic tracking bracket control system is designed and developed based on the CH32 microcontroller. The hardware circuit consists of several parts, including the control unit, power unit, data acquisition unit, communication unit, and drive unit. Among them, the power supply unit is used to first convert the 12V power input by the system into 5V through LM7805T to supply power to the 4G module, and then complete the 5V to 3.3V power conversion through AM1117-3.3 to increase the stable voltage for the main control chip and peripheral sensors; The data acquisition unit completes the data acquisition function through lighting sensors, rain and snow sensors, wind speed and direction sensors, attitude sensors, and GPS. The collected data will be sent to the main control chip for filtering and processing, thereby reducing data jitter and providing a reliable data source for its logical execution; Using 4G modules to complete data upload and distribution, achieve communication functions, and upload data collected by sensors to the Alibaba Cloud platform[7]. Users can remotely monitor the current device status and issue instructions through the Alibaba Cloud platform in case of abnormalities, ensuring the normal operation of the system; The control unit analyzes the latitude, longitude, and RTC time based on the data obtained from GPS, calculates the corresponding theoretical value of the photovoltaic panel tilt angle using astronomical algorithms, and compares it with the photovoltaic panel tilt angle value collected by the attitude sensor to control the driving unit and unify the theoretical value with the actual value[8]. In addition, buttons are integrated in the control unit to achieve local control functions. The specific hardware circuit diagram is shown in Figure 3.

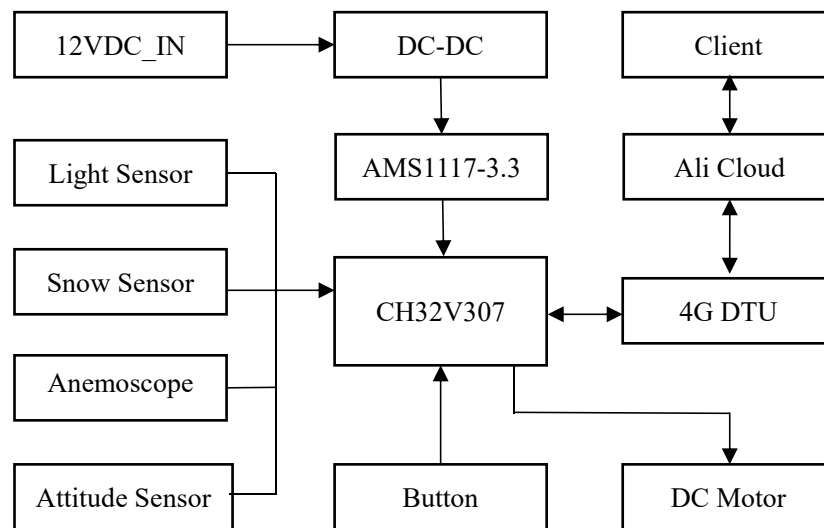


Figure 3 Hardware Circuit Diagram

4 EXPERIMENTAL DATA TESTING AND ANALYSIS

During the testing process, the system selects outdoor as the testing location and conducts fixed-point timed sampling at 9 equidistant time points throughout the day, namely 5:00, 7:00, 9:00, 11:00, 13:00, 15:00, 17:00, 19:00, and 21:00. Based on the above formula, the theoretical and actual tracking angles of the photovoltaic tracking support control system are calculated for one day. After each collection, print out the theoretical angle value through the serial port and record it. The specific test results are shown in Table 1.

Table 1 Angle Test Results

Test time	theoretical angle (°)	actual angle (°)	absolute error (°)	relative error (%)
5: 00	180.0	179.1	theoretical angle	0.50
7: 00	144.7	143.9	0.8	0.55
9: 00	168.5	167.7	0.8	0.47
11: 00	192.1	191.5	0.6	0.31
13: 00	217.4	218.1	0.7	0.32
15: 00	241.2	240.5	0.7	0.29
17: 00	190.9	190.1	0.8	0.42
19: 00	184.5	183.9	0.6	0.32
21: 00	180.0	180.8	0.8	0.44

According to the test results, the angle of the photovoltaic panel changes linearly with time from 9:00 to 15:00. The angle of rotation of the photovoltaic panel every 2 hours is approximately 24 °. During this time period, the system adopts forward tracking mode; The angles between 5:00-9:00 and 15:00-21:00 show nonlinear changes over time, and the system adopts reverse tracking mode. Due to the addition of an attitude sensor in the system to analyze the angle of the photovoltaic panel in real time and provide feedback to CH32, CH32 compares the calculated theoretical value with the actual value returned by the attitude sensor to control the low-speed motor for adjustment until they are consistent. Therefore, the error between the theoretical value and the actual value of the photovoltaic panel angle is relatively small, with an absolute error of less than 1 ° and a relative error of less than 0.6%. Meet the design requirements for system tracking accuracy.

5 CONCLUSION

This design is aimed at the photovoltaic tracking bracket control system, covering the entire process of hardware design, software debugging, and system testing. Advanced sensors have been selected for data acquisition in hardware to ensure the accuracy and effectiveness of the data. The data processing module verifies, filters, and analyzes the collected environmental data, improving the accuracy and reliability of the system. Determine the current operating mode of the system based on the processed sensor data. The system also integrates a 4G wireless communication module, which can transmit real-time monitoring data to the IoT platform. Users can also issue instructions based on the system's various data, completing remote monitoring and control functions. The system testing results show that the tracking error of the system is within $\pm 1^\circ$, which meets the system design requirements and can effectively improve the capture efficiency of sunlight.

CONFLICT OF INTEREST

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

- [1] Akpınar E, Katircioglu G G, Das M. Effects of Solar Tracking on Different Types of Solar Panels; Experimental Study for Thermal and Photovoltaic Types. *International Journal of Hydrogen Energy*, 2025, 144, 611-620.
- [2] Chiteka K, Arora R, Sridhara S N, et al. Influence of Irradiance Incidence Angle and Installation Configuration on the Deposition of Dust and Dust-Shading of a Photovoltaic Array. *Energy*, 2021, 216, 119289. DOI: <https://doi.org/10.1016/j.energy.2020.119289>.
- [3] Wang L, Li R, Dou H, et al. Flat Single Axis Photovoltaic Tracking Bracket - Basic Load Characteristics and Optimization Design. *Acta Energiæ Solaris Sinica*, 2024, 45(10): 345-353.
- [4] Fedenko V, Dzundza B, Pavlyuk M, et al. Design of a Complex Dual-Axis Solar Tracker with an Integrated Solar PV Monitoring System. *Eastern-European Journal of Enterprise Technologies*, 2025, 3(8): 6-13.
- [5] Barrios Sánchez J M, Tlapanco Ríos E I. Dual-Axis Solar Tracking System for Enhanced Photovoltaic Efficiency in Tropical Climates. *Sustainability*, 2025, 17(3): 1117-1117.
- [6] Aguila-Leon J, Vargas-Salgado C, Chiñas-Palacios C, et al. Solar Photovoltaic Maximum Power Point Tracking Controller Optimization Using Grey Wolf Optimizer: A Performance Comparison Between Bio-Inspired and Traditional Algorithms. *Expert Systems With Applications*, 2023, 211, 118700. DOI: <https://doi.org/10.1016/j.eswa.2022.118700>.
- [7] Samosir A S, Gusmedi H, Purwiyanti S, et al. Modeling and Simulation of Fuzzy Logic Based Maximum Power Point Tracking (MPPT) for PV Application. *International Journal of Electrical and Computer Engineering (IJECE)*, 2018, 8(3): 1315-1323. DOI: 10.11591/ijece.v8i3.pp1315-1323.
- [8] Du L, Ma X, Liu Z, et al. Photovoltaic Dual Axis Tracking System Based on Microcontroller. *Journal of Yunnan Normal University (Natural Science Edition)*, 2017(5): 8-12.

