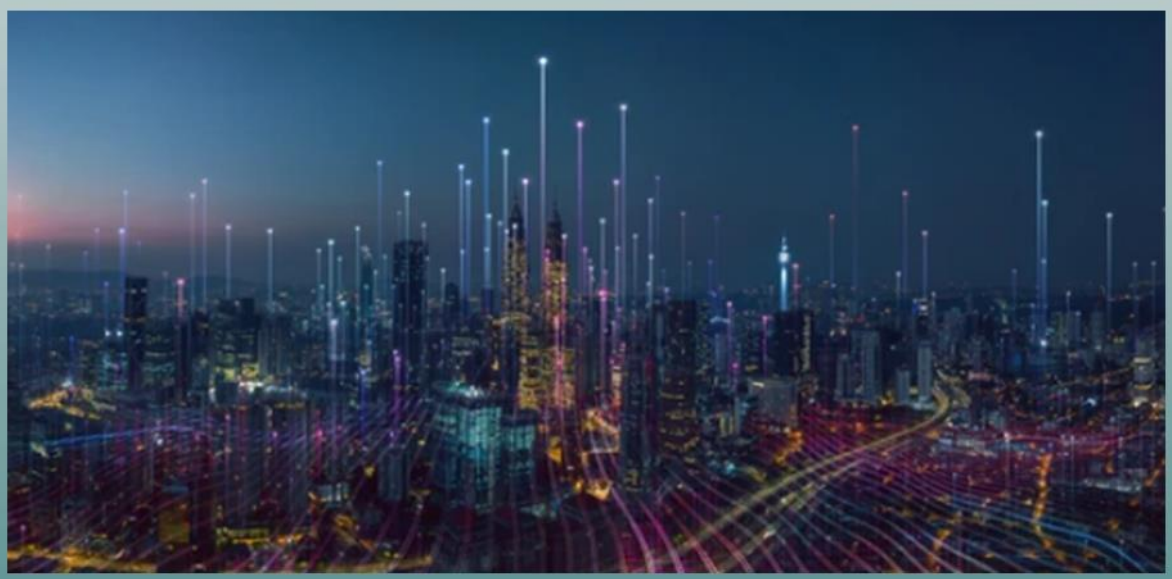
Volume 6, Issue 3, 2024

Print ISSN: 2663-1938
Online ISSN: 2663-1946

JOURNAL OF COMPUTER SCIENCE AND ELECTRICAL ENGINEERING



Journal of Computer Science and Electrical Engineering

Volume 6, Issue 3, 2024

Published by Upubscience Publisher

Journal of Computer Science and Electrical Engineering
Print ISSN: 2663-1938 Online ISSN: 2663-1946
Email: info@upubscience.com
Website: http://www.upubscience.com/

Table of Content

Journal of Computer Science and Electrical Engineering
ISSN: 2663-1946
DOI: 10.61784/jcsee3009

RESEARCH ON FLEXIBLE AND VERIFIABLE SECURITY PROTOCOLS IN CLOUD COMPUTING ENVIRONMENTS

XiaoYin Yi[1,2*], Qian Huang[3]
[1]*Chongqing Key Laboratory of Public Big Data Security Technology, Qijiang 401420, Chongqing, China.*
[2]*Chongqing College of Mobile Communication, Qijiang 401420, Chongqing, China.*
[3]*Chongqing Polytechnic University of Electronic Technology, Shapingba401331, Chongqing,China.*
Corresponding Author:XiaoYin Yi, Email:17323415956@189.cn

Abstract: With the rapid development of cloud computing technology, an increasing number of enterprises and individuals are migrating their data and applications to cloud environments. Although cloud computing offers convenient services and flexible resource management, security issues remain a primary concern for users. This paper explores flexible and verifiable security protocols in cloud environments with the aim of enhancing data security and user trust. By analyzing and improving existing security protocols, a novel security protocol framework is proposed and experimentally validated. The results indicate that this framework demonstrates significant advantages in both flexibility and security.
Keywords: Cloud Storage; Data Integrity; Hierarchical Merkle Tree; Multi-Grained

1. INTRODUCTION

The popularization of cloud computing has brought about a revolutionary change in the way data is stored and processed. However, the accompanying security issues have also become increasingly prominent. Users face risks such as data leakage and unauthorized access when storing data in the cloud. Therefore, researching flexible and verifiable security protocols is crucial for ensuring the security of cloud environments.
Structurally, a cloud storage system [1] can be divided into four layers: storage layer, basic management layer, application interface layer, and access layer. Generally speaking, the architecture of a cloud storage system may vary depending on the application environment and research objectives. Figure 1 illustrates the general architecture of existing cloud storage systems.

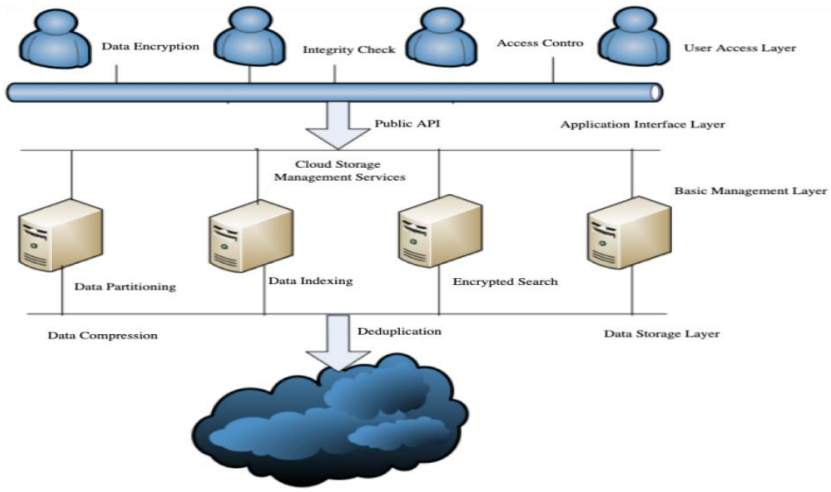


Figure 1 General Architecture of Cloud Storage Systems

The issue with cloud storage services is finding an efficient way for users to verify the authenticity and integrity of data on remote cloud servers. When the service provider is untrustworthy, whether due to dishonesty or system failures, this security threat can be mitigated by integrating data fault tolerance with data integrity and availability verification. When integrity verification schemes involve a third-party auditor, this third party, as an entity within the cloud storage architecture, may also be untrustworthy. The third party could collude with the cloud service provider to help conceal data that has been lost or altered by the provider [2]. Literature confirms users' concerns by indicating that after multiple integrity verifications, as long as a sufficient number of linear equations involving identical blocks are collected, the content of these blocks can be easily obtained by solving the equations. When the third-party auditor is untrustworthy, further integration of data encryption or privacy protection schemes is necessary to address this issue.
The greatest challenge in cloud storage performance is the ability to move data between users and remote cloud storage providers [3], primarily through TCP, which is a connection-oriented protocol based on packet acknowledgment and controls data flow from peer endpoints. Multi-tenancy is a key feature of cloud storage architecture. The characteristics

of cloud storage include: manageability, access methods, performance, multi-tenancy, scalability, availability, control, efficiency, and cost. For details, see Table 1 below.

Table 1 Characteristics of Cloud Storage

Characteristic	Description
Manageability	The ability to manage the system with minimal resources
Access Methods	The protocols used for accessing public cloud storage
Performance	Performance measured based on bandwidth and latency
Multi-Tenancy	Support for multiple users (or tenants)
Scalability	The ability to handle higher demands or scale appropriately
Storage Efficiency	Measurement of how efficiently raw storage is utilized
Cost	Measurement of storage costs (typically per GB)
Data Availability	Measurement of system uptime
Control	The ability to control the system—especially configuring for cost, performance, or other features

Currently, in cloud storage, file chunking is handled with fixed chunk sizes corresponding to BLS signatures, where 20 bytes are used as the basic data unit for authentication. This results in a large Merkle hash tree structure with significant storage costs and slow search operations. Moreover, this approach is only suitable for verifying small data chunks, leading to excessive communication volume and inefficiency during verification [4]. Additionally, the granularity of verification and update operations is limited to 20-byte data blocks, which does not integrate verification environments and actual dynamic operation needs, thus limiting user control. This paper addresses these shortcomings through detailed research and analysis, proposing a reliable and efficient verifiable scheme that further improves and refines the existing verification methods.

The main contributions of this paper are as follows:

(1) We have researched and compared existing cloud storage data integrity verification protocols. Based on the shortcomings of these schemes, we propose a dynamic data security solution that can be flexibly verified in an untrusted environment.

(2) The scheme is applied in an environment where all three parties (users, third-party auditors, and cloud storage servers) are untrustworthy. It includes an operation for the cloud storage server to verify root nodes, enabling the server to more promptly assess the trustworthiness of users and accept user data. Additionally, we have constructed a new hierarchical Merkle hash tree.

(3) The scheme supports dynamic data operations, public verification, and privacy protection. It takes into account and integrates the varying needs of different users, enhancing overall efficiency. Verification operations support chunks and sub-chunks, and dynamic operations can also be supported at these levels. This approach better meets actual user needs and increases the flexibility of verification by elevating the granularity from a singular to an optionally selectable level.

2. RELATED WORK

2.1 File Partitioning

First, the file F is divided into I chunks, represented as $F=(m_1,m_2,...,m_I)$. Then, each chunk m_i is further divided into J sub-chunks, represented as $F=(m_{1,1},m_{1,2},...,m_{I,J})$. Finally, each sub-chunk $m_{i,j}$ is further divided into K basic blocks, at which point the file is represented as $F=(m_{1,1,1},m_{1,1,2},m_{1,1,K},m_{1,2,1},...,m_{I,J,K})$. From these three definitions of the file F, the file can be viewed as a 3-dimensional file structure, as shown in Figure 2. In simple terms, this means decomposing the file from coarse granularity to various sizes of finer granularity [5]. In this paper, 160-bit chunks are used as basic blocks in our scheme, with data sub-chunk sizes of 8KB and data chunk sizes of 512KB. The sub-chunks serve as the J-layer leaf nodes, and data chunks serve as the I-layer leaf nodes to construct a hierarchical Merkle hash tree.

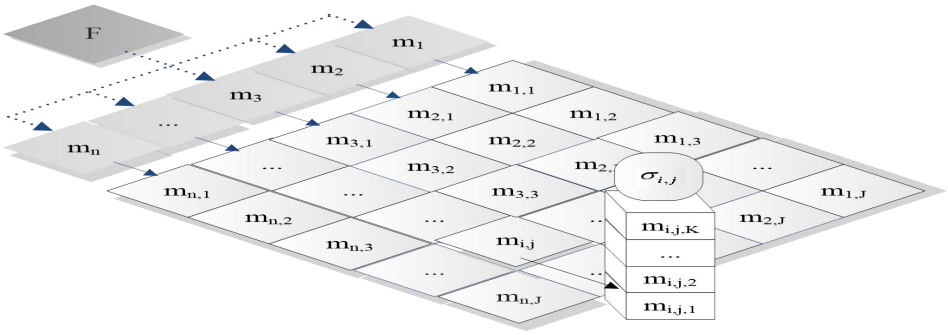


Figure 2 File Chunking Diagram

2.2 Hierarchical Merkle Hash Tree

A hierarchical Merkle hash tree is a logical structuring of the Merkle Hash Tree, aimed at providing a more intuitive representation of the integrity verification process based on two levels of granularity, as well as supporting dynamic operations on blocks. The structure of the hierarchical Merkle hash tree is shown in Figure 3, with the entire tree divided into two levels: Level *I* and Level *J*. Below the dashed line is Level *J*, where a leaf node at Level *J* represents the hash $h(H(m_{i,j}))$ of a sub-chunk label $H(m_{i,j})$.Above the dashed line is Level *I*, where a leaf node at Level *I* represents the hash $h(H^{'}(m_i))$ of a data chunk label $H^{'}(m_i)$. The hash value of the root node of the *J*-level subtree is the same as the hash value of the leaf nodes in the *I*-level tree. Additionally, the relationship between the data chunk label $H^{'}(m_i)$ and the block label $H(m_{i,j})$ is defined as $H^{'}(m_i)=\prod_{1\le i\le n,1\le j\le r} H(m_{i,j})$. Here, *H*(.)is a one-way mapping function $\{0.1\}^{*}\rightarrow G_1$, and *h*(.) is a one-way secure hash function.

In the data integrity verification process, the prover needs to provide the verifier with auxiliary authentication information (abbreviated as AAI). AAI consists of the sibling nodes along the path from the leaf node to the root node in the hierarchical MHT. For different levels of granularity verification, examples include sub-chunks $x_{8.2}$, $x_{8.8}$ and chunks x_2, x_4. The verifier possesses the actual h_r. When the cloud storage server receives verification requests for sub-chunks $x_{8.2}$, $x_{8.8}$ and chunks x_2, x_4, the cloud storage server provides the verifier with reliable AAI, which are as follows:

$\Omega_{8,2}$ =<$h(x_{8,1})$,$h(x_{8,d})$,$h(x_{8,b})$,$h(x_7)$,$h(x_e)$,$h(x_a)$>,

$\Omega_{8,8}$ =<$h(x_{8,7})$,$h(x_{8,e})$,$h(x_{8,a})$,$h(x_7)$,$h(x_e)$,$h(x_a)$>

Ω_2 =<$h(x_1)$,h_d,h_b>, Ω_4 =<$h(x_3)$,h_c, h_b>

The verifier computes:

$h(x_{8,c})=h(h(x_{8,1})||h(x_{8,2}))$,$h(x_{8,a})=h(h(x_{8,c})||h(x_{8,d}))$,$h(x_8)=h(h(x_{8,a})||h(x_{8,b}))$,

$h(x_{8,f})=h(h(x_{8,7})||h(x_{8,8}))$,$h(x_{8,b})=h(h(x_{8,e})||h(x_{8,f}))$,$h(x_8)=h(h(x_{8,a})||h(x_{8,b}))$,

$h_f=h(h_7||h_8)$,$h_b=h(h_e||h_f)$,$h_c=h(h_1||h_2)$,$h_d=h(h_3||h_4)$,$h_a=h(h_c||h_d)$,$h_r=h(h_a||h_b)$,

Then, the verifier checks whether the computed h_r matches the actual one.

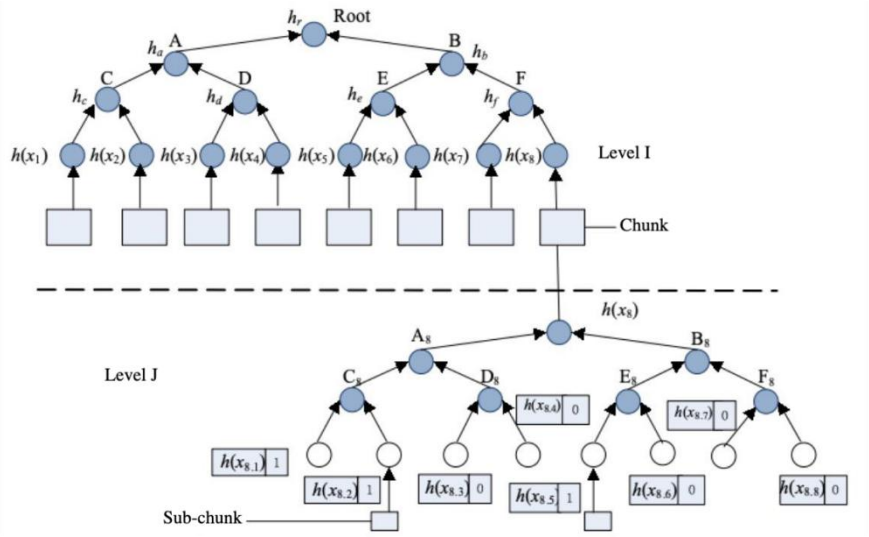


Figure 3 Hierarchical Merkle Hash Tree structure diagram

3. PAPER AND TEXT FORMAT A FLEXIBLE VERIFICATION SECURITY PROTOCOL FRAMEWORK

3.1 Support for Two Levels of Granularity in Integrity Verification Protocols

In this paper, the verification protocol supporting two levels of granularity involves three main stages and employs five algorithm functions. The five algorithm functions are listed in Table 2. The three stages are: the setup stage, the server verification stage, and the user or third-party verification stage. The five algorithm functions are: KeyGen, SigGen, VerifyUser, GenProof, and VerifyProof.

Table 2 The Five Algorithm Functions

Algorithm	Input	Output
KeyGen	A signature pair (*spk*, *ssk*) and security parameter *k*	Public key *pk* and private key *sk*
SigGen	Private key *sk* and outsourced data file *F*	Sub-blocks of the file, the set of signatures for the blocks Φ, Φ', and the metadata for the hierarchical Merkle hash tree root signature $sig_{sk}(H(R))$
VerifyUser	User file and signature	Accuracy of file data
GenProof	File *F*, set of signatures Φ, Φ', and challenge value *chal*	Aggregated results and data proof for the challenged sub-blocks or blocks
VerifyProof	Aggregated results and data proof	Verification results for the challenged sub-blocks or blocks

3.2 Support for Dynamic Data Operations at Two Levels of Granularity

In an untrusted environment, data integrity verification schemes can efficiently handle fully dynamic data operations based on a hierarchical Merkle hash tree as the storage verification structure. In a remote cloud storage environment, without retrieving the entire data file, dynamic update operations utilize four algorithm functions: VerifyUser, ExecUpdate, SearchUpdate, and VerifyUpdate, as shown in Table 3. Remote support for users includes operations such as insertion, deletion, and modification of data files.

Table 3 The Foure Algorithm Functions

Algorithm	Input	Output
VerifyUser	User file and signature	Accuracy of file data
ExecUpdate	Dynamic operation execution request	Updated file, set of signatures after F', operation Φ_{xin}, $\Phi_{\text{xin}'}$, and update proof P_{update}
SearchUpdate	Dynamic operation search request	Results of leaf nodes corresponding to the searched sub-blocks and blocks
VerifyUpdate	Update operation verification request	Verification results

3.2.1 Sub-block or block insertion

This paper will introduce block insertion; insertion of sub-blocks is similar. Suppose a client needs to insert a data block m_i^* after an existing data block (here $1 \le i \le n$). The specific protocol operation process is as follows:

(1) The user first decomposes the data block m^* into J data sub-blocks, represented as $(m_{i,1}^*, m_{i,2}^*, ..., m_{i,J}^*)$. Finally, each data sub-block $m_{i,j}^*$ is further decomposed into K basic blocks, referred to as $(m_{i,1,1}^*, m_{i,1,2}^*, m_{i,1,k}^*, m_{i,2,1}^*, ..., m_{i,J,K}^*)$. The user then signs the data block m_i^* to be inserted and the data sub-blocks that compose it as $\sigma_i^* = (H'(m_i^*).\prod_{k=1}^{k} u_k^{\sum m_{i,j,k}^*})^a$ and $\sigma_{i,j}^* = (H(m_{i,j}^*).u^{m_{i,j}^*})^a$. Next, the user constructs a Merkle Hash Tree (MHT) with the data block m_i^* as the root node and sends the data insertion request information "$update = (I, i, m_i^*, \sigma_i^*, \sigma_{i,j}^*)$" to the server.

(2) When the cloud server receives the insertion request, it first verifies the data block to be inserted using the VerifyUser algorithm and then performs the insertion operation. The process is as follows: First, store the data block m_i^* and save all its sub-block signatures $\sigma_{i,j}^*$. Using the SearchUpdate algorithm, locate node $h(H'(m_i))$ in the hierarchical Merkle Hash Tree (MHT) and retain its auxiliary authentication information Ω_i. Next, insert the MHT with $h(H'(m_i^*))$ as the root node, and add a new node as the parent of both node $h(H'(m_i))$ and node $h(H'(m_i^*))$. The value of this parent node is the hash of the concatenation of its child nodes $h(H(m_i))$ and $h(H'(m_i^*))$. Then, update the information of all sibling nodes along the path from this parent node to the root node of the tree, i.e, update the hash values of all nodes along this path, and compute the new root R'. Add the signature σ_i^* to Φ'. Finally, the cloud server sends the result of the data insertion operation $P_{update} = (\Omega_i, H'(m_i), Sig_{sk}(H(R), R'))$ back to the user.

(3) When the user receives the operation result, they first verify the validity of m_i using the auxiliary authentication information $\{\Omega_i, H'(m_i)\}$ and the verification algorithm. If the verification is invalid, the result is FALSE. If the verification is valid, calculate the old root node R based on the relevant information and use equation to verify the authenticity of the tree root R. If the equation does not hold, output FALSE. If the equation holds, compute the information of the new root R_{new} using $\{\Omega_i, H'(m_i), H'(m^*)\}$ and compare it with R'. If the values are different, output FALSE. Otherwise, output TRUE. The user then signs the new root node R' on the hierarchical MHT and sends the signature of the new root to the cloud server. Upon receiving the new root's signature, the cloud server saves the new root signature and deletes the old root signature file. Finally, the user performs a complete integrity verification of the blocks or sub-blocks. If the result is TRUE, it indicates that the server has completed the data insertion operation. At this point, the user deletes the locally stored information $\{Sig_{sk}(H(R')), P_{update}, m_i^*, \sigma_i^*, \sigma_{i,j}^*\}$.

3.2.2 Deletion of Sub-blocks or Blocks

Next, we will describe the deletion of a block, which is similar for the deletion of sub-blocks. Suppose the user needs to delete the *i*-th data block m_i (here $1 \le i \le n$). The specific deletion protocol process is as follows:

(1) The user first sends a data block deletion request information "$update = (D,i)$" to the cloud server.

(2) When the cloud server receives the data deletion request information from the user, it uses a search algorithm to find the node $h(H'(m_i))$ associated with the data block m_i and all corresponding sub-block nodes $h(H(m_{i,j}))$. After retaining the auxiliary authentication information Ω_i of $h(H'(m_i))$, the server deletes the node m_i and its parent node, as well as all corresponding sub-block nodes of m_i. It then takes the sibling nodes of m_i as the new parent node and updates the hash values of all nodes along the path from this modified parent node to the root node. The new root R' is calculated, and σ_i is deleted from Φ' and $\sigma_{i,j}$ is deleted from Φ. Finally, the result $\{\Omega_i, H'(m_i), Sig_{sk}(H(R), R')\}$ is sent to the user.

(3) When the user receives the deletion operation proof, they first use the verification algorithm to check whether the data block is the one intended for deletion. Then, they generate the root value R using $\{\Omega_i, H'(m_i)\}$. By verifying the equation, they confirm the authenticity of the auxiliary authentication information (AAI) and the root value R. If the equation verification fails, the result is FALSE. Otherwise, the user signs the new root with $sig_{sk}(H(R'))$ and sends the new signature back to the cloud server. Upon receiving the new root signature information, the cloud server saves it and deletes the old root signature file $sig_{sk}(H(R'))$. The user then performs an integrity verification of the data block or sub-block. If the result is TRUE, it indicates that the cloud server has completed the user's data deletion operation; otherwise, it indicates that the cloud server has not completed the user's data deletion operation.

3.2.3 Modification of Sub-blocks or Blocks

we will describe the modification of a block, which is similar for sub-block modifications. Suppose the user needs to modify the *i*-th data block m_i to m^*_i. The specific data block modification protocol process is as follows:

(1)The user first decomposes the data block m^*_i into *J* sub-blocks, represented as $(m^*_{i,1}, m^*_{i,2}, ..., m^*_{i,J})$. Each sub-block $m^*_{i,j}$ is further decomposed into *K* basic blocks, resulting in $(m^*_{i,1,1}, m^*_{i,1,2}, m^*_{i,1,k}, m^*_{i,2,1}, ..., m^*_{i,J,K})$. The user then signs the data block m^*_i and its constituent data sub-blocks $\sigma_i^* = (H'(m_i^*) \cdot \prod_{k=1}^{k} u_k^{\sum m^*_{i,j,k}})^a$ and $\sigma^*_{i,j} = (H(m^*_{i,j}) u^{m^*_{i,j}})^a$. Next, the user constructs a Merkle Hash Tree (MHT) with m^*_i as the root node and sends the data modification request information "$update = (M, i, m_i^*, \sigma_i^*, \sigma_{i,j}^*)$" to the server. Here, *M* represents the modification operation request, which updates the request to the server.

(2) Upon receiving the update request, the cloud server runs the VerifyUser and ExecUpdate algorithms. After verifying the data block m^*_i, it performs the modification operation: first, the data blocks m^*_i, $\sigma^*_{i,j}$, and the corresponding layer I leaf nodes $h(H(m^*_i))$ are stored. Then, using a search algorithm, the server finds node $h(H'(m_i))$in the hierarchical MHT and retains its auxiliary authentication information Ω_i. The server then updates the hash values of all nodes along the path from $h(H'(m^*_i))$ (the modified node) to the root node at layer I, and modifies node $h(H'(m_i))$ to be an MHT with $h(H'(m^*_i))$ as the root node. This results in a new root value R', and signatures σ_i^* and $\sigma_{i,j}^*$ are added to Φ' and Φ respecti vely. Finally, the cloud server responds to the user's operation by sending $P_{update} = \{\Omega_i, H'(m_i), Sig_{sk}(H(R), R')\}$ back to the user.

(3) After the user receives the proof of the modification operation from the cloud server, they first use the verification algorithm to check whether the data sub-block is the one intended for modification. They then generate the root value R using $\{\Omega_i, H'(m_i)\}$. The authenticity of the auxiliary authentication information (AAI) and the root value R is verified through the verification equation $e(sig_{sk}(H(R)), g) = e(H(R), v)$. If the equation verification fails, FALSE is output; otherwise, the user verifies whether the server has faithfully executed the data modification operation by further using $\{\Omega_i, H'(m^*_i)\}$ to compute a new root value. The newly calculated root value is compared with R'. If the values are not equal, FALSE is output; otherwise, TRUE is output. The user then signs the new root node R' with $sig_{sk}(H(R'))$, and sends the signature of the new root $sig_{sk}(H(R'))$ to the cloud server. Upon receiving $sig_{sk}(H(R'))$, the server saves it and deletes the old root signature file. At the same time, the server deletes the original signature of the data block σ_i and all signatures of the corresponding sub-blocks $\sigma_{i,j}$. Finally, the user performs an integrity verification on the data, treating it

as blocks or sub-blocks. If the output result is TRUE, the stored information $\{Sig_{sk}(H(R')),P_{update},m^*_i,\sigma^*_i,\sigma^*_{i,j}\}$ is deleted by the user's client.

4. SECURITY ANALYSIS

High security is always a primary concern. For this solution, while improving overall efficiency, the main concerns for users are whether data integrity and privacy can be adequately assured. For cloud storage servers, the key concerns are the correctness and completeness of the data submitted by users. The following security analysis addresses these concerns based on the protocol. The security analysis of the protocol is based on the Diffie-Hellman problem and the Discrete Logarithm Problem (DLP).

4.1 Data Correctness and Integrity for Users

When users send outsourced data files to a remote cloud storage server, they must ensure the correctness, consistency, and integrity of the data $\{F,t,\Phi,\Phi',sig_{sk}(H(R))\}$ and its associated parameters. The signature scheme ensures that the data is unforgeable, meaning $H'(m_i)$ and $H(m_{i,j})$ cannot be forged. Additionally, since $spk,g,v,\{w_k\}_{1\le k}$, $\{u_k\}_{1\le k}$ is public, if $\sigma_i/\sigma_{i,j}$ or $m_i/m_{i,j}$ provided by the user are incorrect, they will not satisfy the bilinear mapping equation. Due to the difficulty of the discrete logarithm problem, it is infeasible for users to forge $\sigma_i/\sigma_{i,j}$ and $m_i/m_{i,j}$ simultaneously, which guarantees their consistency. For the user's data to pass verification by the cloud storage server, the user must submit authentic metadata and signatures to the cloud storage server.

4.2 Privacy Protection

During the verification process, if μ'/μ'' is directly exposed to the Third Party Auditor (TPA), the TPA can collect enough combinations of identical blocks or sub-block equations, which would allow them to easily solve a set of linear equations to obtain the data file block $\{m_i\}_{i\in I}$ or sub-block $\{m_{i,j}\}_{1\le i\le I,1\le j\le J}$. Therefore, it is crucial to ensure that μ'/μ'' information is not exposed to the TPA during the verification process. In the verification proof $\{\bar{\sigma},\mu_j,Q_k,(H'(m_i),\Omega_i)_{i\in I}/(H(m_{i,j}),\Omega_{i,j})_{i\in I,j\in I_2},sig_{sk}(H(R))\}$ sent to the TPA, since μ_j is the value obtained after masking the information μ'/μ'' with the element O_j randomly selected by the server, the TPA knows Q_k but, due to the difficulty of the discrete logarithm problem, O_j remains unknown. Thus, the privacy of μ'/μ'' can be assured through μ_j. Based on the assumptions of the Diffie-Hellman problem, the TPA cannot obtain the value of μ'/μ''. Therefore, privacy protection is guaranteed [6].

5. CONCLUSION

This paper first proposes a flexible data integrity verification scheme that supports multiple granularities based on the different requirements of verification operations. It introduces various data verification methods for users, further enhancing the overall verification efficiency. Additionally, under a new security model where the cloud server, TPA, and users are in an untrusted environment, it adds verification processes for the cloud server and root verification. This enables timely and efficient detection of threats to users. Furthermore, for data updates, it introduces the option for dynamic operations, allowing both sub-blocks and blocks to be dynamically updated. This scheme increases users' control over data verification, enabling cloud storage to provide better and more tailored services to users. At the same time, it integrates the features of public auditability and privacy protection. Finally, through analysis and comparison, the proposed scheme is demonstrated to be feasible.

COMPETING INTERESTS

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

FUNDING

This research was funded by the Scientific and Technological Research Program of Chongqing Municipal Education Commission (Grant No. KJQN202302403, KJQN202303111).

REFERENCES

[1] Yue Ru. Research on Data Storage and Management Technology in Cloud Computing Environment. Science and Technology Information, 2023, 21(21): 29-32.
[2] Dhakad N, Kar J. EPPDP: An Efficient Privacy-Preserving Data Possession With Provable Security in Cloud Storage. IEEE Systems Journal, 2022, 16(4): 6658-6668.
[3] Li Hao, Zhang Qian, Wang Lei. A Hybrid Approach for Cloud Data Integrity Verification Based on Blockchain Technology. Future Generation Computer Systems, 2022, 128: 542-556.

[4] Sun Xize, Zhou Fucai, Li Yuxi, et al. A Database Encryption Scheme Based on Searchable Encryption Mechanism. Journal of Computer Research and Development, 2021, 44(4): 806-819.
[5] Ding Y, Li Y, Yang W, et al. Edge Data Integrity Verification Scheme Supporting Data Dynamics and Batch Auditing. Journal of Systems Architecture, 2022, 128: 128.
[6] Hou Huiying, Ning Jianting, Huang Xinyi, et al. Verifiable Attribute-Based Timed Signature Scheme and Its Application. Journal of Software, 2023, 34(5): 2465-2481.

Journal of Computer Science and Electrical Engineering
ISSN: 2663-1946
DOI: 10.61784/jcsee3011

EXPLORATION OF AI EMPOWERED EXPERIMENTAL TEACHING REFORM IN COMPUTER SCIENCE

Yu Zhang
Beihai Campus of Guilin University of Electronic Technology, Beihai 536000, Guangxi, China.
Corresponding Email: 76474853@qq.com

Abstract: In the context of rapidly advancing information technology, Artificial Intelligence (AI) has profoundly impacted various industries, presenting new challenges and opportunities for higher education, particularly in computer science experiment teaching. Despite covering fundamental topics such as programming basics and algorithm design, current computer science experiment courses often suffer from a disconnect between content and real-world applications, with outdated materials that fail to keep pace with industry developments. This gap leaves students ill-prepared to navigate rapidly evolving technological landscapes. Additionally, traditional teaching methods and assessment models limit students' opportunities for independent exploration and innovation, while outdated laboratory facilities further hinder the quality of experimental teaching. To address these challenges, this study proposes AI-enabled reforms in experimental teaching. The strategies include establishing a "multi-dimensional, practice-oriented" curriculum system, implementing a "data-driven, precision-guided" teaching model, promoting "self-directed, flexible progression" learning paths, and building a "collaborative innovation and open-sharing" experimental teaching environment. These reforms aim to enhance students' practical skills, increase the practical relevance of courses, and comprehensively improve the quality of experimental teaching, ultimately preparing students to meet the demands of future industry needs.
Keywords: Artificial intelligence; Computer science; Experimental teaching; Multi-dimensional Curriculum System; Personalized teaching

1 INTRODUCTION

In today's rapidly evolving information technology landscape, Artificial Intelligence (AI) is transforming industries at an unprecedented pace, posing both challenges and opportunities for higher education, particularly in the field of computer science[1-3]. As the demand for computer science professionals with advanced skills continues to grow, traditional experimental teaching methods are increasingly proving inadequate in cultivating students' practical abilities, innovative thinking, and adaptability to future technological environments. While computer science experimental courses at universities typically cover fundamental topics such as programming basics, algorithm design, and operating systems, these courses often remain confined to theoretical validation and reinforcement, lacking close integration with rapidly evolving real-world applications. This disconnect is particularly evident in cutting-edge areas such as AI, big data processing, and cloud computing, where the pace of content updates lags behind industry advancements, leaving students ill-prepared to respond effectively to technological changes[4-5].

The prevailing experimental teaching approach largely relies on the traditional "teacher-directed, student-executed" model, which emphasizes standardized procedural steps while neglecting the development of students' abilities for independent exploration and innovative thinking. This rigid teaching model limits students' flexibility in addressing complex problems, and the assessment methods, overly focused on final outcomes, fail to fully capture students' learning processes and cognitive strategies. Moreover, outdated laboratory facilities and limited environmental flexibility further constrain the scope and depth of experimental teaching. The current hardware and software configurations in laboratories are often insufficient to meet the demands of complex contemporary experiments, depriving students of authentic technical experiences during practical exercises[6-7].

Against this backdrop, the need for reform in experimental courses to enhance students' practical skills, increase the applicability of course content, and improve overall teaching quality has become a pressing concern in computer science education. This paper aims to explore how AI-enabled approaches can be leveraged to construct a multi-dimensional experimental curriculum system, optimize data-driven teaching models, design flexible learning pathways, and build a collaborative and innovative experimental teaching environment. These efforts are intended to comprehensively elevate the quality and effectiveness of computer science experimental teaching, ultimately producing high-quality graduates equipped with comprehensive application capabilities and innovative potential.

2 CURRENT STATE OF COMPUTER SCIENCE EXPERIMENTAL TEACHING

2.1 Lack of Integration with Real-world Applications and Delayed Content Updates

Computer science experimental courses hold a crucial position within university curricula, aiming to help students translate theoretical knowledge into practical skills through hands-on experience. However, there are several notable issues and shortcomings in the current design of these courses[8]. Firstly, in terms of course content, while most

universities offer experimental courses covering fundamental topics such as programming basics, algorithm design, computer networks, and operating systems, these courses tend to focus predominantly on the verification and reinforcement of theoretical knowledge. They often lack a close integration with real-world applications, particularly in rapidly evolving fields such as artificial intelligence, big data processing, and cloud computing. The pace of content updates in these areas is often slower than the advancements in the industry, leaving students inadequately prepared to cope with the fast-changing technological environment.

2.2 Monotonous Teaching Methods and Overemphasis on Final Outcomes

In terms of teaching methods, the current experimental courses largely rely on the traditional "teacher-directed, student-executed" model. Most courses emphasize standardized procedural steps, with students typically following the instructions in the laboratory manuals mechanically, without engaging in deep critical thinking or independent exploration[9-10]. This monotonous teaching approach not only limits the development of students' innovative thinking but also weakens their ability to address complex problems. Moreover, the assessment of experimental work often focuses heavily on the final results rather than providing a comprehensive evaluation of students' thought processes and problem-solving strategies. This emphasis on end results further reduces student motivation and engagement in the experimental learning process.

2.3 Outdated Laboratory Equipment and Lack of Flexibility in the Environment

Outdated laboratory facilities also pose a significant constraint on the quality of experimental courses. Many university computer science laboratories are equipped with outdated hardware and software that fail to meet the demands of current complex experiments. For instance, traditional laboratory environments are often incapable of supporting large-scale data processing or the training of sophisticated AI models, which prevents students from gaining authentic technical experience during their practical exercises. Additionally, the lack of flexibility in laboratory environments often makes it difficult to accommodate the specific needs of different experimental courses, further limiting the breadth and depth of the experimental curriculum.

2.4 Insufficient Connection with Industry Frontiers, Lack of Industry Cases and Practical Application Scenarios

The current design of experimental courses often lacks effective alignment with the cutting-edge advancements in the industry. As technology continues to evolve, the industry's expectations for computer science professionals, particularly in emerging fields such as artificial intelligence, the Internet of Things, and blockchain, are rising. However, the content of university experimental courses frequently lags behind the pace of industrial development, resulting in a disconnect between the knowledge and skills acquired by graduates and the demands of the workforce. Additionally, the availability of industry-related case studies and practical application scenarios within experimental courses is limited, depriving students of opportunities to practice and truly comprehend the application of the latest technologies in industrial settings.

3 STRATEGIES FOR EXPERIMENTAL COURSE TEACHING REFORM

3.1 Reform Objectives

The primary objective of the reform is to enhance students' practical abilities. Experimental teaching reform should focus on cultivating students' hands-on skills and problem-solving abilities by introducing more challenging and practically valuable experimental projects. These projects would enable students to deepen their understanding of theoretical knowledge through practice and learn to flexibly apply what they have learned to real-world scenarios.

Another core objective of the reform is to increase the practical relevance of the courses. The content of experimental courses needs to be closely aligned with current industry trends, particularly in cutting-edge technologies such as artificial intelligence, big data, and the Internet of Things. By incorporating real-world application cases of these technologies, students can not only learn the latest techniques and methods but also understand how these technologies are applied in different industries, thereby enhancing their competitiveness in the job market.

Furthermore, improving the overall quality of teaching is the ultimate goal of the reform. The enhancement of teaching quality is reflected not only in students' academic performance but also in the development of their initiative, creativity, and comprehensive abilities throughout the learning process. By optimizing course structure and content, and employing diverse teaching methods, experimental teaching will evolve from a mere knowledge transmission process to one that stimulates students' potential and promotes their holistic development.

3.2 Reform Content

3.2.1 Establishing a "Multi-dimensional, Practice-oriented" Experimental Curriculum System

In today's rapidly evolving technological environment, relying solely on knowledge from a single discipline is insufficient to address complex problems. Therefore, the reform of experimental courses should focus on establishing a "multi-dimensional, practice-oriented" curriculum system. The core of this system lies in integrating multidisciplinary

content, organically combining fields such as artificial intelligence, data science, and computer engineering to enable cross-disciplinary application of knowledge. Through this integration, students will not only grasp the fundamental principles of various disciplines but also develop the ability and innovative thinking required to solve real-world problems using cross-disciplinary knowledge.

In practice, this goal can be advanced by designing experimental projects that encompass multiple disciplines. For example, in AI-related experimental courses, students would not only be required to master the design and implementation of algorithms but also integrate data science knowledge by processing and analyzing large datasets, applying these data to train models. This multidisciplinary approach not only enhances the practical relevance of experimental courses but also helps students better understand and apply their knowledge in complex scenarios, thereby fostering comprehensive abilities to navigate dynamic industrial environments.

Simultaneously, introducing industry project practices is another crucial component of this reform. Based on current industry development needs, representative real-world projects should be selected as experimental case studies through collaboration with enterprises. By solving these real-world problems, students can gain a profound understanding of the application scenarios of AI technologies. For instance, in an experimental project within the field of intelligent manufacturing, students would engage in the entire process, from requirements analysis and algorithm design to system implementation. Such industry-connected experimental courses not only expose students to the latest technologies and practices during their academic studies but also lay a solid foundation for their future careers.

3.2.2 Implementing a "Data-supported, Precision-guided" Teaching Model

To enhance the targeting and effectiveness of experimental teaching, the reform should implement a "data-supported, precision-guided" teaching model. This model leverages intelligent learning assessment systems that use big data analysis and AI technologies to conduct real-time evaluations of students' learning progress and experimental outcomes, thereby providing personalized learning suggestions for each student.

Specifically, an intelligent learning assessment system can comprehensively analyze students' operational data, learning behaviors, and experimental results to identify potential difficulties and issues encountered during the learning process. For example, the system might detect that a student frequently makes errors in a particular experimental step, prompting it to automatically recommend relevant supplementary resources or provide further guidance. This data-driven precision guidance not only helps students promptly correct their learning deficiencies but also significantly enhances their learning outcomes.

By deeply analyzing students' learning data, instructors can quickly identify shortcomings in their teaching, dynamically adjusting the content and difficulty of experimental courses to better meet the diverse learning needs of students. For instance, based on students' overall performance and feedback, instructors may increase or decrease the difficulty of certain experimental tasks or rearrange the sequence of experiments to improve the coherence of the course and the efficiency of students' learning. This dynamic adjustment mechanism ensures that experimental courses maintain high standards while also being flexible enough to accommodate the personalized needs of students.

3.2.3 Promoting a "Self-directed, Flexible Progression" Learning Pathway

The core of this strategy lies in the modular design of experimental courses, where the curriculum is divided into multiple independent modules, each focusing on a specific area of knowledge or skill. Students can independently select and combine different modules based on their interests and career plans, thereby achieving personalized learning objectives. Figure 1 illustrates three experimental module groups, with arrows indicating overlapping areas. For example, a student interested in AI algorithm development might choose experimental modules that include algorithm design and model optimization. This self-directed learning pathway not only allows students to explore their areas of interest in greater depth but also better prepares them for their future careers.

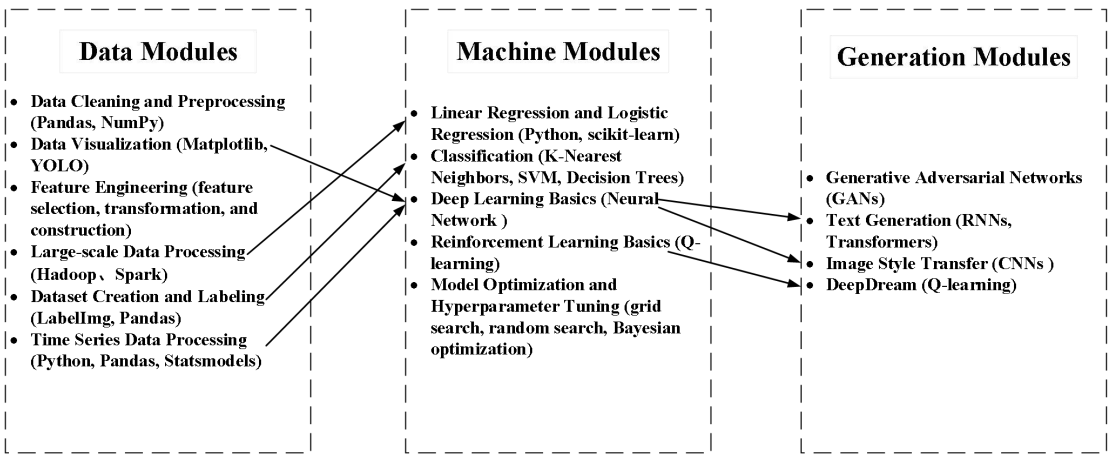


Figure 1 The Classification and Intersection of Experimental Module Groups

Supporting flexible scheduling of experiments is also a key component of this strategy. By introducing online experimental platforms and remote virtual lab environments, students can conduct experiments outside of regular class

hours and manage their progress more flexibly. For example, students may continue working on experiments beyond the lab's scheduled hours or revisit experiments after completion to reinforce their understanding. This flexibility not only enhances learning efficiency but also accommodates different learning paces, ensuring that each student can fully grasp the material. By offering flexible learning pathways and schedules, the reform effectively meets students' individual needs and promotes their overall development.

3.2.4 Building a "Collaborative Innovation and Open-sharing" Experimental Teaching Environment

To cultivate students' innovation capabilities and teamwork skills, the reform of experimental teaching should also focus on building a "collaborative innovation and open-sharing" teaching environment. The core of this environment is fostering teacher-student co-innovation through a project-driven teaching model, which encourages both teachers and students to collaborate in designing and implementing experimental projects. This collaboration not only stimulates students' creativity but also develops their ability to work collaboratively in solving problems.

In this model, teachers are no longer mere transmitters of knowledge but become collaborators with students. Together, teachers and students can determine the themes of experimental projects, design the experimental framework, and work closely throughout the execution process. This cooperative teaching approach not only engages students more actively in the experimental process but also helps them cultivate critical thinking and innovative skills in practice. For example, in an open-ended experimental project, teachers and students might jointly research a new algorithm, explore its applications in specific fields, and continuously refine it based on experimental results.

Strengthening university-industry collaboration is another critical element in building an open-sharing experimental teaching environment. By integrating industry experts and resources into the design and implementation of experimental courses, the curriculum can be more closely aligned with industry needs, ensuring that the content remains cutting-edge and practically relevant. University-industry collaboration not only provides students with more internship and employment opportunities but also offers new perspectives and resources for teachers' research and teaching. For instance, industry experts might be invited to participate in course design or organize student visits to corporate R&D departments to learn about the latest technologies and industry trends. This model of university-industry cooperation not only enhances students' practical skills but also boosts their competitiveness in the future job market.

4 CONCLUSION

Through the exploration of AI-enabled reforms in computer science experimental teaching, this study has proposed a series of specific strategies aimed at enhancing the quality of experimental teaching and improving its alignment with modern industry demands. Establishing a multi-dimensional, practice-oriented experimental curriculum system can cultivate students' comprehensive application skills and innovative thinking. The implementation of a data-supported, precision-guided teaching model makes the teaching process more personalized and efficient. Meanwhile, promoting self-directed, flexible progression learning pathways provides students with greater freedom in their studies, and building a collaborative innovation and open-sharing experimental teaching environment further stimulates students' creativity and teamwork abilities.

Overall, this study not only provides a theoretical foundation and practical guidance for current reforms in computer science experimental teaching but also points to future directions for educational reform. As AI technology continues to advance, the reform of experimental teaching must also progress to ensure that education and technology develop in tandem, ultimately producing more outstanding talents capable of leading future technological innovation.

COMPETING INTERESTS

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

FUNDING

This work is supported by the 2024 Guangxi Vocational education teaching reform research project (GXGZJG2024B202).

REFERENCES

[1] Wan Sha. Application of ChatGPT in College English Teaching. Science & Education Guide (Electronic Edition), 2024: 195-197.

[2] Ye Jinhu, Wang Hua, Liu Yanfang, et al. Constructing "New Medical Science": A Flipped Classroom Teaching Model of Medical Imaging in the Context of "AI 2.0 Era"—Taking Nuclear Medicine Imaging as an Example. Creative Education Studies, 2024, 12. DOI: 10.12677/ces.2024.126350.

[3] Zhuang Qibin, Zhang Jiqin. Exploration of Application Paths of ChatGPT in University Teaching Process. Science & Technology Review, 2024(19).

[4] Jieun A, Jongbong L, Myeongeun S. ChatGPT in ELT: Disruptor? Or Well-Trained Teaching Assistant?. ELT Journal, 2024. DOI: 10.1093/elt/ccae017.

[5] Hu Yaochong, Huang Yiqing, Yuan Yiming. Teaching Research Integrating Internet and Artificial Intelligence in the Context of "New Engineering"—Taking Introduction to Automation as an Example. Science & Technology Review, 2024(2): 105-108.

[6] Xie Yongpeng, Xu Yinzhou. The Mechanism and Promotion Path of Artificial Intelligence Empowered Classroom Teaching Reform in Colleges and Universities. Journal of Neijiang Normal University, 2024, 39(3): 75. DOI: 10.13603/j.cnki.51-1621/z.2024.03.013.
[7] Cai Jincao, Wang Lei, Wang Jianbin, et al. Research on Teaching Design of Artificial Intelligence Introduction Course for Mechanical Manufacturing Major. Computer Knowledge and Technology, 2024(2).
[8] Hu Xiangyang. Exploration of the Application of ChatGPT in Python Programming Course Teaching. Education Informatization Forum, 2024(4).
[9] Tang Haoxia, Li Li, Liu Jinjiang. Research on the Application of ChatGPT in MySQL Database Teaching in Vocational Colleges. Internet of Things Technology, 2024(6).
[10] Wang Guicai, Fan Chao, Zhang Yuhong. Discussion on Artificial Intelligence Practical Teaching under the Mode of Industry-Education Integration—Taking the Artificial Intelligence Industry College of Henan University of Technology as an Example. Henan Education (Higher Education), 2023(8).

Journal of Computer Science and Electrical Engineering
ISSN: 2663-1946
DOI: 10.61784/jcsee3010

INTERCONNECTION DESIGN EQUIPMENT REQUIREMENT FOR IPPS IN PAKISTAN

Muhammad Farooq[1*], Haseeb Khalid[2]
[1]*Department of Computer Sciences and Information Technology, Superior University Lahore, Pakistan.*
[2]*Power China HDEC, Beijing 100037, China.*
Corresponding Author: Muhammad Farooq, Email: mfarooqawan2000@gmail.com

Abstract: In Pakistan, electrical energy is produced from a variety of sources, and the government is currently planning to combine renewable energy sources to provide end users with cheaper electricity. This would support Pakistan's manufacturing sector and may also result in more exports, which will reduce Pakistan's import-export imbalance and boost total reserves, which are now being depleted from the country's overall economy. These renewable energy projects have created new business opportunities in Pakistan, where investors are urged to become independent power producers (IPPs) and generate electricity using standardized technologies for both producing and selling electricity. Understanding the processes that IPPs must take in order to install a Power Plant is the goal of this publication. This paper will assist IPP investors in making better plans and managing projects within the allotted time frames and budgets. Additionally, it will provide an overall standard checklist for IEC or NTDC to choose better equipment and designs to run their plant more effectively.
Keywords: NTDC (National Transmission and Dispatch company); IEC (International Electro technical Commission); IPP (Independent Power Producers); PSP (Planning section of power system); MW (Mega Watt); MOU (Memorandum of understanding); EOI (Expression of Interest); NEPRA (National Electric Power Regulatory Authority)

1 INTRODUCTION

The power system in Pakistan is a combination of several sources that provide power to the end consumer through various interconnected systems. It is a positive development that the government, particularly the NTDC, has created a thorough plan to incorporate new energy sources, particularly renewable ones, in order to reduce the impact of existing energy costs on consumers and to boost Pakistan's economy especially in the field of manufacturing and exports to increase overall country revenue. According to statistics provided by the NTDC PSP department.

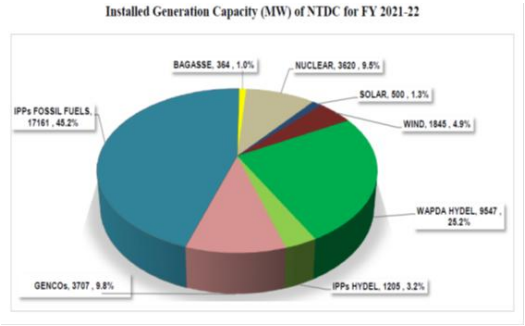


Figure 1 Showing Power Generation capacity of Pakistan [2]

Pakistan is capable of producing a total of 37,809 MW of electricity utilizing a variety of sources, the majority of which come from the combustion of fuel [1]. Pakistan, through NTDC PSP, wants to use renewable energy sources in the same way that the rest of the world is doing. This is definitely will boost overall economy of Pakistan and at the same time is a great chance for IPP's to come forward in the trade of renewable energy and earn fortune.
The government is undoubtedly advancing the renewable sector as it solicits EOIs, launches them, and signs MOUs to demonstrate its commitment to this industry.
The government intends to start 14,000 megawatts worth of solar power projects [4].The Pakistani PM declares that all government buildings would be solar-powered by April. [5]
The news like above is definitely and eye opener for potential upcoming IPP's, especially in solar and wind sector. The purpose of this document is to provide information about the approval process and the requirements for the approval of equipment.

2 RESEARCH METHODOLOGY

This paper's research methodology used a systematic approach to understand the needs and practices of IPPs in Pakistan. Here a detailed review of regulations, standards, and guidelines set by Pakistani organization like IEC and NTDC will be the part of this study. The study collected data from multiple sources, including Government reports, industry standards, and case studies of existing power plants.

3 LITERATURE REVIEW

A recent study has emphasized the importance of integrating different renewable energy resources into Pakistan's national grid, highlighting the need for standardized guidelines for independent power plants (IPPs). Reza et al. (2022) discussed the difficulties, challenges, and prospects of modern smart grid implementation in Pakistan, emphasizing the importance of new and well-designed interconnection framework [7]. Apart from this, some other research focuses on the technical aspects, advocating highly advanced technologies and reliable equipment to reduce the energy losses to some extent and ensure more stability of the system [8].
Another paper that examines the role of corporate social responsibility (CSR) in enhancing the overall performance of independent power pants (IPPs) in Pakistan [9], CSR practices, innovations, and environmental impacts in context of the entire thermal power generation sector in the country.

4 INRODUCTION TO PAKISTAN POWER STRUCTURE: GENERATION & DISTRIBUTION

Pakistan's power structure is a convoluted merger of departments, each of which is responsible for carrying out a certain duty, including the following.

4.1 NEPRA

It is also responsible for issuing licenses for generation, transmission and distribution of electricity, establishing and enforcing standards to ensure quality and safety of operation and supply of electric power to consumers; approving investment and power acquisition programs of utility companies; and determining tariffs for the generation, transmission, and distribution of electric power.

4.2 CPPA-G

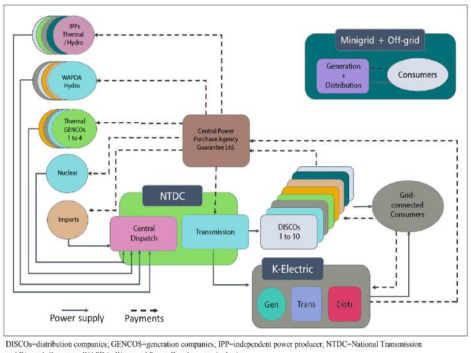
CPPA-G is Pakistan's market operator and is enabling the change from a single buyer to a competitive market for electricity.

4.3 NTDC

NTDC operates & maintains eighteen 500kV grid stations & Forty-Nine 220kV grid stations; it is also responsible for the approval of Projects in terms of Design, protection and communication.

4.4 NPCC

NPCC is the nerve center of Power Sector. Entire power generation and its transmission to the load centers is controlled and coordinated from its National and Regional Control Centers. NPCC is a contracting outfit and its charter includes management of Power Projects on turnkey basis i.e. Extra High Voltage Transmission Lines, Distribution Networks, Substations, Power Generation Plants, Industrial Electrification, External Lighting of Housing Complexes etc.
Figure 2 depicts the division of work among these departments and their standing in relation to job descriptions established by the Ministry of Energy and Power. The CPPA-G is in charge of calculating load demand and allocating the required values to various power production sectors, while NTDC assists in power transmission to various Power distributors sectors that distribute power to end customers, as can be seen in the figure. IPPs, WAPDA, and other generation stations receive payment from CPPA-G for the amount of power they have produced.



DISCOs=distribution companies; GENCOS=generation companies; IPP=independent power producer; NTDC=National Transmission and Dispatch Company; WAPDA=Water and Power Development Authority.

Figure 2 Showing Power Structure of Pakistan[6]

5 INTERCONNECTION DESIGN EQUIPMENT

The architecture of the plant, the choice of the equipment, and the equipment's type, all fall under the category of interconnection design equipment for power plants. Furthermore, the type of interconnection equipment that is chosen depends greatly on the generation source.
Interconnection equipment, also known as BOP, must be approved by regulatory agencies in accordance with certain IEC and NTDC standards.
These interconnection equipment are "Air Circuit Breaker, Batteries, Bus Support Insulators, Capacitor Bank, Circuit Switchers, Concrete Foundation, Conduits, Control House, Control Panels, Control Wires Current Transformers, Disconnect Switches, Distribution Bus, Duct Runs, Frequency Changers, Grounding Resistors, Grounding Transformers, High-Voltage Underground Cables, Lightning Arresters, Manholes, Metal-clad Switchgear, Meters, Microwave, Oil Circuit Breakers, Potential Transformers, Potheads, Power-line Carrier, Power Transformers,Rectifiers,Relays,SF6 Circuit Breakers, Shunt Reactors, Steel Superstructures ,Supervisory Control, Transmission Bus, Vacuum Circuit Breakers" .The equipment mentioned falls under a peculiar category in terms of design perspective and have been segregated by the regulatory agencies for approval which is covered in the lateral section of this document.

6 INTERCONNECTION DESIGN EQUIPMENT REQUIRMENT

Interconnection equipment has been segregated in three major categories:
1. Primary Design Equipment
2. Secondary Design Equipment
3. Communication Devices – Telecom Equipment

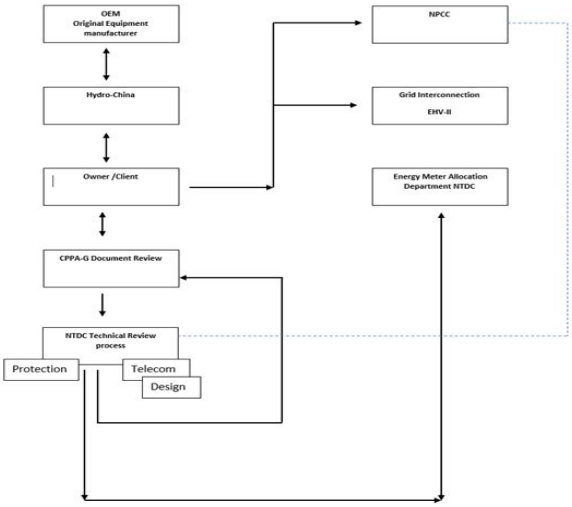


Figure 3 Revealing Flow Chart for the Approval Process at NTDC

In order to install a power plant for generation and specially to sell Electricity to CPPA-G, it is mandatory that the interconnection equipment should be approved by the regulatory authority and in case of selling to public sector, NTDC plays a vital role in approval.

7 OVER VIEW OF NTDC

National Transmission & Despatch Company (NTDC) was incorporated on 6th November, 1998 and commenced commercial operation on 24th December, 1998. [3] It was organized to take over all the properties, rights and assets obligations and liabilities of 220 KV and 500KV Grid Stations and Transmission Lines/Network owned by Pakistan Water and Power Development Authority (WAPDA).
National Transmission and Despatch Company (NTDC) links Power Generation Units with Load Centers spread all over the country (including Karachi) and thus establishes and governs one of the largest interconnected Networks.
The Company is responsible for evacuation of Power from the Hydroelectric Power Plants (mainly in the North), the Thermal Units of Public (GENCOs) and Private Sectors (IPPs) (mainly in the South) to the Power Distribution Companies through primary (EHV) Network.

8 TRANSMISSION NETWORK OPERATOR (TNO)

Operation and Maintenance of the 500/220kV Network: Planning, Design, and Construction of the New 500/220kV Systems, and Strengthening/Up gradation of Existing One [3].

8.1 System Operator (SO)

Arranging non-discriminatory and non-preferential economic dispatch to ensure a safe, secure, and reliable supply [3].

8.2 Wire Business

- Transmission Planning
- Design and Engineering
- Project Development and Execution
- Operation and Maintenance of Transmission Assets

8.3 System Operation & Despatch

- Generation Despatch
- Power System Operation and Control

9 NTDC ORGANOGRAM W.R.T TO DIFFERENT DEPARTMENTS

9.1 Design Department

- Reviewer Office
- Manager office
- Manager Dispatch office
- Chief office
- Chief Dispatch office
- GM office
- GM Dispatch office
- MD Office
- MD Dispatch office

9.2 Protection Department

- Reviewer Office
- Manager office
- Manager Dispatch office
- Chief office
- Chief Dispatch office
- GM office
- GM Dispatch office

9.3 Telecom Department

- Reviewer Office
- Manager office
- Manager Dispatch office
- Chief office
- Chief Dispatch office
- GM office
- GM Dispatch office

9.4 Transmission Design Office

9.5 TSG – Technical Service Group

9.6 Reports Review office

9.7 HR office

10 DEPARTMENTS & EQUIPMENTS

With respect to approval process as mentioned in figure 2 three departments of NTDC plays a vital role namely

1. Project design office
2. Protection and control Office

3. Telecom department

It should be taken in consideration that these standards are revised by NTDC from time to time in order to better surveille the equipment.

10.1 Design Department

List of equipment for designing department:

Table 1 List of Equipment for Designing Department

Sr#	Submission Item	IEC Standards	NTDC Specifications
1	SLD	N/A	N/A
2	Substation Layout	N/A	N/A
3	AIS Equipment Circuit Breaker	IEC 62271-100 IEC 62271-1 IEC 60273	P-193:2010
4	Type Test, FAT and Routine Tests of Circuit Breaker	IEC 62271-100	P-193:2010
5	AIS Equipment Line Disconnectors ,AC Bus Disconnectors & Earthing Switch	IEC 62271-102 IEC 60417 IEC 60273	P-128:2011
6	Type Test & Routine Tests of Line Disconnectors ,AC Bus Disconnectors & Earthing Switch	IEC 62271-102	P-128:2011
7	Instrument Transformer Protection Current Transformer	IEC 60044-11 IEC 60273 IEC 60044-1	P-90:2012
8	Routine Test ,Factory Acceptance Test &Type Test Protection Current Transformer	IEC 60044-1 IEC 60296 IEC 60273	P-90:2012
9	Instrument Transformer Metering Current Transformer	IEC 61869-1 IEC 61869-2 IEC 62155 IEC 60273	P-205:2020
10	Routine Test , Factory Acceptance Test &Type Test Metering Current Transformer	IEC 61869-1 IEC 61869-2 IEC 62155 IEC 60273	P-205:2020
11	Instrument Transformer Metering Voltage Transformer	IEC 60044-2 IEC 60385 IEC 296	P-206:2005
12	Routine Test , Factory Acceptance Test &Type Test Metering Voltage Transformer	IEC 358 IEC600044-22 ANSI C93.2	P-206:2005
13	Surge Arrestor	IEC 60099-4	P-181:2012
14	Type Test, FAT, Routine Test	IEC 60068 IEC 60099-4	P-181:2012

	Report of Surge Arrestor		
15	Grid Hardware	---------	P-176:2011 P-187:2010 P-142:1996 P-143:1996 P-188:1996
16	Metering Panel	---------	P-199:08 P-202:12
17	Energy Meter	--------	P-202:12

10.2 Protection Department

Table 2 List of Equipment for Protection Department

Sr#	Submission Item	IEC Standards
1	SLD	For Protection following specification are incorporated for the approval of submission items • NTDC specs for 132kV Relay and Control Panels-> P-151:2008 • NTDC specs for Auxiliary AC/DC panels: P48:81 • NTDC Specs for Metering system: P199:08 & P202:12 • P-204-08 is for the relays and protection system usually conforms with the IEC standard relays
2	PMSLD	
3	Relay Ordering Codes	
4	Trip Logic and Interlocking Logic	
5	AC DC SLD	
6	Signal List	
7	Relay and Control Panel Schematic	
8	Relay Settings	
9	FAT of Protection Panels	

10.3 Telecom Department

Table 3 List of Equipment for Telecom Department

Sr#	Submission Item	IEC Standards
1	OPGW	IEEE Standard 1138
2	Optical Fibre	IEC 60794 Optical
3	Protection Class of Cabinets/Cubicles	Fiber Cables EN 60529
4	line traps	IEC 60353
5	coupling devices for power line carrier systems	IEC 60481
6	Coupling Capacitors	IEC 60044
7	Recommended values for characteristic input and output quantities of single sideband power line carrier terminals	IEC 60495 IEC 60663
8	Planning of (single side band) power line carrier systems	IEC 60834-1
9	Tele-protection equipment of power systems – performance and testing	

11 CONCLUSION

In this paper, we have identified a way for IPPs investors to better plan and manage projects within the allotted time frame and budget. Furthermore, this study can help IEC or NTDC operate their plant more efficiently. However, this study also provides an overall standard checklist for choosing the best equipment and design to run. It is obvious from the previous description that the road map for constructing a power plant is laborious for IPPs. But the aforementioned steps are designed to improve the electricity system. However, the Government may make it easier for IPP investors by creating an online system. Since typical operating procedures call for the printing of millions of pages, many of which are ultimately thrown away, creating an online system could prevent this waste and make job easier.

COMPETING INTERESTS

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

REFERENCES

[1] National Transmission & Despatch Company Limited (NTDC). "Power System Statistics 47th Edition December 2022." Report, 2022. Available at: https://es.scribd.com/document/709410151/47th-Edition-Final-Version.
[2] National Electric Power Regulatory Authority (NEPRA). "State of Industry Report 2022." NEPRA Report, 2022. Available at: https://nepra.org.pk/publications/State%20of%20Industry%20Reports.php.
[3] National Transmission & Despatch Company (NTDC). "Official Website." NTDC, 2024. Available at: https://ntdc.gov.pk/.
[4] Tribune Express. "Official Website." Tribune Express, 2024. Available at: https://www.tribuneexpress.com/.
[5] Saur Energy. "Official Website." Saur Energy, 2024. Available at: https://www.saurenergy.com/.
[6] Central Power Purchasing Agency (CPPA-G). "Official Website." CPPA-G, 2024. Available at: https://www.cppa.gov.pk/.
[7] Muhammad Amir Raza, Muhammad Mohsin Aman, Abdul Ghani Abro, et al. Challenges and potentials of implementing a smart grid for Pakistan's electric network, Energy Strategy Reviews, 2022, 43, 100941. DOI: https://doi.org/10.1016/j.esr.2022.100941. https://www.sciencedirect.com/science/article/pii/S2211467X22001353
[8] PITCO Pakistan. "Projects." PITCO Pakistan. Accessed August 9, 2024. https://www.pitcopk.com/projects.
[9] Dunnan, Liu, Gul, Rana, Jaffri, Nasif, et al. Independent Power Producer (IPPs) of Pakistan and CSR to Enhance Performance. 2020. 1-5. DOI: 10.1109/iCoMET48670.2020.9073918.

Journal of Computer Science and Electrical Engineering
ISSN: 2663-1946
DOI: 10.61784/jcsee3012

THREAT ASSESSMENT OF AIR TARGETS BASED ON FAHP-ICRITIC COMBINATION WEIGHTING

Haobo Wang*, Jiantao Liu
Zhengzhou Campus of Army Academy of Artillery and Air Defense, Zhengzhou 450000, Henan, China.
Corresponding Author: Haobo Wang, Email: 475111565@qq.com

Abstract: Aiming at the problem of target threat assessment in air defense operations, a target threat assessment model based on Fuzzy Analytic Hierarchy Process (FAHP) and Improved Criteria Importance Although Intercrieria Correlation (ICRITIC) is proposed. On the basis of establishing the threat assessment index system, the FAHP and ICRITIC methods are used to determine the subjective and objective weights of the threat indicators, the combination weights are based on the principle of maximizing the deviation, and the target threat degree is evaluated based on the TOPSIS method (Technique for Order Preference by Similarity to an Ideal Solution, TOPSIS). The weights and target threat degree obtained by different weighting methods are analyzed to verify the rationality and effectiveness of the proposed method.
Keywords: Air defense operations; Threat assessment; Fuzzy analytic hierarchy process (FAHP); Improved criteria importance though intercrieria correlation (ICRITIC); Combination weighting

INTRODUCTION

The previous information local wars since the Gulf War show that air strikes have become the main form of war initiated by combat opponents. With the continuous development of military theory, military technology and military practice, there are more types of air-raid weapons, with better performance and stronger functions. In the future air defense operations, air strikes and air defense confrontation are more intense, and the "OODA" cycle in the combat process is significantly accelerated, and the complex and changeable battlefield situation requires commanders to make quick, accurate and efficient decisions. Scientific evaluation of the threat degree of various incoming targets is the premise of scientific decision-making, which can provide a basis for the subsequent reasonable allocation of anti-aircraft firepower and maximize the effectiveness of weapons and equipment.

The main contents of target threat assessment include: extraction of threat elements, quantification of threat elements, and determination of threat level[1]. When determining the target threat level, it is necessary to determine the weight of each threat element and rank the degree of the target threat.

A single method or a combination method can be used to determine the weight of threat indicators. Single empowerment method includes both subjective empowerment method and objective empowerment method. Subjective empowerment methods include expert scoring method, Delphi method, hierarchical analysis method, etc. Document [2] determines the weight of threat index of penetrating air combat aircraft, and improves the rationality of the scale method and consistency inspection method of hierarchy analysis to improve the efficiency of target optimization; Document [4] uses the fuzzy consensus matrix in the evaluation to meet the requirements of the relative ambiguity of the evaluation object and overcome the difficulty of consistency inspection. Objective empowerment methods include entropy value method, principal component analysis method, CRITIC method and so on. Literature [5] uses the entropy method to determine the subjective weight value of each attribute of the sea surface target to make the empowerment more objective; literature [6] uses the principal component analysis method to determine the threat index weight to reduce the subjective influence of determining the index weight; literature [7] uses the CRITIC method to determine the attribute weight, considers the data volatility and conflict of each attribute information, and makes the weight allocation more reasonable. The combination empowerment method can consider the subjective preference and the characteristics of evaluation element data, and the common method is linear weighting method[8], Multiplicative synthesis and normalization method[9], The vector similarity method[10]class. The linear weighting method can use the game theory[11], least square method[12]Maximize the difference[13].To determine the combined weight coefficients of the different empowerment methods.

Based on the above analysis, because the CRITIC method underconsiders the correlation coefficient and the degree of data dispersion between the indicators, this paper uses FAHP and improved CRITIC method to determine the subjective and objective weights of the indicators, makes linear weighted combination based on the idea of difference maximization, and uses TOPSIS method to evaluate the threat degree of aerial targets.

1 SELECTION OF THE INDEX SYSTEM

The threat degree of air target generally refers to the possibility of the enemy's successful air strike on our defended targets and the degree of possible damage caused after the successful air strike, which is a reflection of the enemy's air strike intention and the combat ability of air strike weapons. The factors affecting the target threat degree mainly include the target type, flight conditions, flight time, etc. The threat assessment indicators selected in the relevant

literature are shown in Table 1.

Table 1 Threat Assessment Indicators Selected from Relevant Literature

Literature number	Selected threat indicators
Literature [14]	S hortcuts to routes ; A ltitude ; T ime of approach ; N umber of targets
Literature [15]	Target type ; S peed ; T ime of approach ; A ltitude ; P arameter update time
Literature [16]	Target type ;Electronic jamming capability ;T ime of approach ; Altitude ; Maneuvering haracteristics ;Shortcuts to routes
Literature [17]	Target type ;Electronic jamming capability ; H eading angle ; A ltitude ; S peed
Literature [18]	Target type ; S peed ; T ime of approach ; Shortcuts to routes ; Electronic jamming capability ;
Literature [19]	S peed ; Characteristics of mobility ; Electronic jamming capability ; A ltitude ; H eading angle ; Distance from own position

Based on relevant research and practical experience, Target type, Altitude, Speed, Time of approach (TOA), Shortcuts to routes (STR), Electronic jamming capability (EJC) and others were selected as threat assessment indicators.

1.1 Target Type

Different target types, their mobility performance, carrying load, attack mode are also different, and the degree of threat to the defended target is also different. The target type threat degree was quantified by the expert scoring method, as shown in Table 2.

Table 2 Threat Affiliation of Target Types

Target type	Bomber	F ighter	A rmed helicopter	Cruise missile	Reconnaissance aircraft
quantified value	0.7	0.6	0.4	0.3	0.2

1.2 Target Height

When the target carries out low altitude and ultra-low altitude mobile penetration, due to the influence of terrain, earth curvature, radar blind area and other factors, the probability of being found by our intelligence reconnaissance system has decreased, and it is easier to achieve the surprise of air attack. Usually, the lower the target height, the greater the threat level. The threat membership function for the target height can be expressed as:

$$y(h)=\begin{cases}1 & 0\le h\le 1\\ e^{-k_h(h-1)^2} & 1\le h\le 20\end{cases} \tag{1}$$

Where h is the target height in k m, $k_h=10^{-2}$.

1.3 Target Speed

The higher the target speed, the easier it is to break through our interception, and the greater the threat degree. The threat membership function of the target speed can be expressed as:

$$y(v)=1-e^{k_v v} \tag{2}$$

Where v is the target speed in m / s, $k_v=-6\times10^{-3}$.

1.4 Fly Time

The shorter the target flight time, the shorter the planning and decision time and the response time of the weapon system, and the greater the threat degree. The threat membership function of flight time can be expressed as:

$$y(t)=e^{-k_t t^2},t\ge 0 \tag{3}$$

Where, t is the flight time and is measured in s, $k_t=2\times10^{-5}$.

1.5 Road Shortcuts

The smaller the shortcut of the target route, the stronger the intention to attack on us, and the greater the threat. The threat membership function of the shortcut can be expressed as:

$$y(d)=e^{-k_d d^2},-30\le d\le 30 \tag{4}$$

Where, d is the air route shortcut, in km, $k_d=8\times10^{-3}$.

1.6 Electronic Interference Capability

Air raid weapons can interfere with and suppress our detection and tracking equipment by carrying an electronic warfare system. The stronger the electronic interference ability of the target, the greater the threat degree. In this paper, the electronic interference ability is quantified as no, weak, weak, strong and strong, and the quantified values are 0,0,0,0 and 0..2.4.6.8

2 TARGET THREAT ASSESSMENT STEPS

2.1 The Subjective Weights were Determined based on the FAHP Method

FAHP determines the relative importance of each index by comparing them separately, establishes the fuzzy complementary matrix and the fuzzy consensus matrix, and then obtains the index weight.

2.1.1 Establish the fuzzy complementary matrix

.1-0.9The "0 scale method" is used to assign the relative importance of the index (the scale significance is shown in Table 3) to establish a fuzzy complementary matrix. Matrix A meets the following conditions: $A=\left(a_{ij}\right)_{m\times m}$

（1） $a_{ii}=0.5, i=1,2,\cdots,m;$

（2） $a_{ij}=1-a_{ji}, i,j=1,2,\cdots,m$

Table 3 Meaning of the “0.1-0.9 Scale”

Relative importance	Meaning	Description
0.5	E qually important	Indicators i and j are equally important
0.6	Slightly important	Indicator i is slightly more important than j
0.7	Significantly important	Indicator i is significantly more important than j
0.8	Strongly important	Indicator i is more strongly important than j
0.9	Extremely important	Extreme importance of indicator i over j
0.1,0.2,0.3,0.4	Anti-comparison	

2.1.2 Establish a fuzzy consensus matrix

The matrix A is converted into a fuzzy consensus matrix according to Equation (5) and (6). $B=\left(b_{ij}\right)_{m\times m}$

$$B=\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mm} \end{bmatrix}$$

$$b_i=\sum_{j=1}^{m}a_{ij} \qquad i=1,2,\cdots,m \tag{5}$$

$$b_{ij}=\frac{b_i-b_j}{2(m-1)}+0.5 \quad i,j=1,2,\cdots,m \tag{6}$$

2.1.3 Calculate the weight of each index

The weight of the i th indexω_iFor[20]：

$$\omega_i=\frac{1}{m}-\frac{1}{2\alpha}+\frac{1}{m\alpha}\sum_{j=1}^{m}b_{ij} \qquad i,j=1,2,\cdots,m \tag{7}$$

$\alpha\alpha$Where, (m-1) / 2, when = (m-1) / 2, the index weight difference is the largest.

2.2 Determine the Objective Weights Based on the I CRITIC Method

CRITIC The method determines the index weight according to the volatility and conflict of the data. Volatility is expressed as the standard deviation of the data under a certain index. On the basis of calculating the conflict degree between a certain index and the other index. According to the volatility and conflict, the information contained in a certain index is found, so as to get the weight of the index. The assignment method has the following deficiencies: (1) when the absolute value of the correlation coefficient of the two indicators is the same but the positive and negative values are different, the correlation reflected is the same; (2) the method mainly considers the volatility and correlation of the data, not the dispersion of the data, so it can be improved in combination with the entropy method[21].

CRITIC The formula for solving the weight method is (data standardization method and solution process, reference [7]):

$$\omega_i = \frac{S_i \sum_{j=1}^{m} \left(1 - r_{ij}\right)}{\sum_{i=1}^{m} S_i \sum_{j=1}^{m} \left(1 - r_{ij}\right)} \tag{8}$$

Where, m represents the number of evaluation indicators, and S_iRepresents the standard deviation of the i th indicator, r_{ij}Represents the correlation coefficient of the two indicators, i and j, using the Pearson correlation coefficient.
There are two main points for the improvement of the CRITIC method: (1) the information entropy value is added to the volatility calculation, and the improved volatility calculation formula is:

$$S_i^{'} = S_i + e_i \tag{9}$$

Where is e_iIs the information entropy of index i (solution process reference [22]).
(2) For the correlation coefficient, r_{ij}Taking the absolute value, the improved conflict calculation formula is:

$$R_i^{'} = \sum_{j=1}^{m} \left(1 - \left|r_{ij}\right|\right) \tag{10}$$

The improved weight calculation formula is:

$$\omega_i = \frac{\left(S_i + e_i\right) \sum_{j=1}^{m} \left(1 - \left|r_{ij}\right|\right)}{\sum_{i=1}^{m} \left(S_i + e_i\right) \sum_{j=1}^{m} \left(1 - \left|r_{ij}\right|\right)} \tag{11}$$

2.3 Determine the Combination Weight Based on the Difference Maximization Idea

The basic idea of difference maximization is that for different evaluation objects, if the evaluation value of each evaluation object is quite different under a certain attribute, then the attribute makes a large contribution to the evaluation results and should be given a large weight; if the difference is small, a small weight should be given.
The subjective weight vector =(was obtained using the F AHP method $\alpha\alpha_1, \alpha_2, \ldots, \alpha_m)^T$, Using the ICRITIC method to obtain the objective weight vector, $\beta = (\beta_1, \beta_2, \ldots, \beta_m)^T$, Combined the two linearly to obtain the combined weight vector, $W = (\omega_1, \omega_2, \ldots, \omega_m)^T$. The W can be expressed as:

$$W = \lambda_1 \alpha + \lambda_2 \beta \tag{12}$$

In the formula,λ_1、λ_2For the combined weight coefficient,$\lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_1^2 + \lambda_2^2 = 1$.
The total deviation of all the evaluated objects can be expressed as:

$$D = \sum_{j=1}^{m} \sum_{i=1}^{n} \sum_{p=1}^{n} \omega_j \left|y_{ij} - y_{pj}\right| \tag{13}$$

Where, n is the number of the evaluated objects, and $y_{i\,j}$Is the j th index value of the i th evaluation object after the standardized treatment.
Based on the difference maximization idea, the optimization model is constructed, as shown in (14). The combined weight coefficients are solved by constructing the Lagrangian function[23] .

$$\begin{cases} D = \sum_{j=1}^{m} \sum_{i=1}^{n} \sum_{p=1}^{n} \omega_j \left|y_{ij} - y_{pj}\right| = \\ \sum_{j=1}^{m} \sum_{i=1}^{n} \sum_{p=1}^{n} \left(\lambda_1 \alpha_j + \lambda_2 \beta_j\right) \left|y_{ij} - y_{pj}\right| \\ s.t. \quad \lambda_1^2 + \lambda_2^2 = 1 \\ \lambda_1 \geq 0, \lambda_2 \geq 0 \end{cases} \tag{14}$$

$$\begin{cases} \lambda_1 = \dfrac{\sum_{j=1}^{m} \sum_{i=1}^{n} \sum_{p=1}^{n} \alpha_j \left|y_{ij} - y_{pj}\right|}{\sqrt{\left(\sum_{j=1}^{m} \sum_{i=1}^{n} \sum_{p=1}^{n} \alpha_j \left|y_{ij} - y_{pj}\right|\right)^2 + \left(\sum_{j=1}^{m} \sum_{i=1}^{n} \sum_{p=1}^{n} \beta_j \left|y_{ij} - y_{pj}\right|\right)^2}} \\ \lambda_2 = \dfrac{\sum_{j=1}^{m} \sum_{i=1}^{n} \sum_{p=1}^{n} \beta_j \left|y_{ij} - y_{pj}\right|}{\sqrt{\left(\sum_{j=1}^{m} \sum_{i=1}^{n} \sum_{p=1}^{n} \alpha_j \left|y_{ij} - y_{pj}\right|\right)^2 + \left(\sum_{j=1}^{m} \sum_{i=1}^{n} \sum_{p=1}^{n} \beta_j \left|y_{ij} - y_{pj}\right|\right)^2}} \end{cases} \tag{15}$$

The final combined weight vector, W, is obtained after substitution of equation (12) and normalized to the combined

weight coefficient$_c$=(ω_{c1},ω_{c2},…,ω_{cm})^T .

2.4 Evaluate the Target threat Degree Based on the TOPSIS Method

The TOPSIS method is a comprehensive evaluation method of distance. The central idea is to determine the optimal ideal value and the worst ideal value of each evaluation index, and then find the Euclidean distance between each scheme and the positive and negative ideal value, thus obtaining the relative proximity between each scheme and the optimal scheme. The main calculation steps are performed as follows.
(1) Data standardization processing;
(2) Calculate the weighted standardization matrix;
(3) Determine the positive and negative ideal solution;
(4) Calculate the Euclidean distance from the ideal solution;
(5) Calculate the closeness between each scheme and the optimal scheme.

3 EXAMPLE ANALYSIS

It is assumed that in an air defense operation, a batch of air targets are found, and based on the comprehensive analysis of air situation data and various intelligence information, the attribute information of each target is shown in Table 4.

Table 4 Initial attribute values of t argets

Target number	Target type	Altitude	Speed	Time of approach	Shortcuts to routes	Electronic jamming capability
1	Fighter	8	400	170	22	0.8
2	Fighter	3	315	60	20	0.8
3	Cruise missile	0.1	400	150	15	0.6
4	Cruise missile	0.02	600	100	9	0.5
5	Armed helicopter	0.03	120	90	2	0.9
6	Bomber	10	240	140	28	0.5

The raw data was quantified according to the membership function of each threat indicator to obtain the initial decision matrix X:

$$X=\begin{bmatrix} 0.6000 & 0.6126 & 0.9093 & 0.5610 & 0.0208 & 0.6000 \\ 0.6000 & 0.9608 & 0.8489 & 0.9305 & 0.0408 & 0.4000 \\ 0.3000 & 1.0000 & 0.9093 & 0.6376 & 0.1653 & 0.0000 \\ 0.3000 & 1.0000 & 0.9727 & 0.8187 & 0.5231 & 0.0000 \\ 0.4000 & 1.0000 & 0.5132 & 0.8504 & 0.9685 & 0.2000 \\ 0.7000 & 0.4449 & 0.7631 & 0.6757 & 0.0019 & 0.4000 \end{bmatrix}$$

The X was standardized to obtain the standardized decision matrix Y:

$$Y=\begin{bmatrix} 0.7500 & 0.3022 & 0.8620 & 0.0000 & 0.0196 & 1.0000 \\ 0.7500 & 0.9294 & 0.7306 & 1.0000 & 0.0402 & 0.6667 \\ 0.0000 & 1.0000 & 0.8620 & 0.2073 & 0.1691 & 0.0000 \\ 0.0000 & 1.0000 & 1.0000 & 0.6974 & 0.5392 & 0.0000 \\ 0.2500 & 1.0000 & 0.0000 & 0.7833 & 1.0000 & 0.3333 \\ 1.0000 & 0.0000 & 0.5438 & 0.3104 & 0.0000 & 0.6667 \end{bmatrix}$$

3.1 Solve the Index Weights Using the FAHP Method

By comparing the threat indicators, the fuzzy complementary matrix A:

$$A=\begin{bmatrix} 0.5 & 0.4 & 0.7 & 0.8 & 0.2 & 0.8 \\ 0.6 & 0.5 & 0.8 & 0.9 & 0.4 & 0.8 \\ 0.3 & 0.2 & 0.5 & 0.6 & 0.2 & 0.6 \\ 0.2 & 0.1 & 0.4 & 0.5 & 0.1 & 0.5 \\ 0.8 & 0.6 & 0.8 & 0.9 & 0.5 & 0.9 \\ 0.2 & 0.2 & 0.4 & 0.5 & 0.1 & 0.5 \end{bmatrix}$$

Convert the matrix A to a fuzzy consensus matrix:

$$B=\begin{bmatrix} 0.5000 & 0.4400 & 0.6000 & 0.6600 & 0.3900 & 0.6500 \\ 0.5600 & 0.5000 & 0.6600 & 0.7200 & 0.4500 & 0.7100 \\ 0.4000 & 0.3400 & 0.5000 & 0.5600 & 0.2900 & 0.5500 \\ 0.3400 & 0.2800 & 0.4400 & 0.5000 & 0.2300 & 0.4900 \\ 0.6100 & 0.5500 & 0.7100 & 0.7700 & 0.5000 & 0.7600 \\ 0.3500 & 0.2900 & 0.4500 & 0.5100 & 0.2400 & 0.5000 \end{bmatrix}$$

A subjective weight vector = $(0.1827,0.2067,0.1427,0.1187,0.2267,0.1227)\alpha^{T}$.

2.3 The weights were Solved Using the ICRITIC Method

In the normalized matrix Y, the separate calculation, the standard deviation and information entropy of each indicator data,
S_j=[0.4306,0.4406,0.3609,0.3849,0.4001,0.4037],e_j=[0.7229,0.8619,0.8876,0.8238,0.5837,0.7376],S_j'=[1.1535,1.3025,1.2485,1.2087,0.9838,1.1413] .
correlation matrix R'For:

$$R'=\begin{bmatrix} 1.0000 & 0.8964 & 0.4114 & 0.1068 & 0.7904 & 0.8949 \\ 0.8964 & 1.0000 & 0.1968 & 0.4942 & 0.8597 & 0.8124 \\ 0.4114 & 0.1968 & 1.0000 & 0.1691 & 0.0338 & 0.3401 \\ 0.1068 & 0.4942 & 0.1691 & 1.0000 & 0.4237 & 0.2114 \\ 0.7904 & 0.8597 & 0.0338 & 0.4237 & 1.0000 & 0.6323 \\ 0.8949 & 0.8124 & 0.3401 & 0.2114 & 0.6323 & 1.0000 \end{bmatrix}$$

The objective weight vector is $\beta = (0.1202,0.1243,0.2634,0.2382,0.1219,0.1320)^{T}$.
The weight of the threat index obtained by each empowerment method is shown in Figure 1.

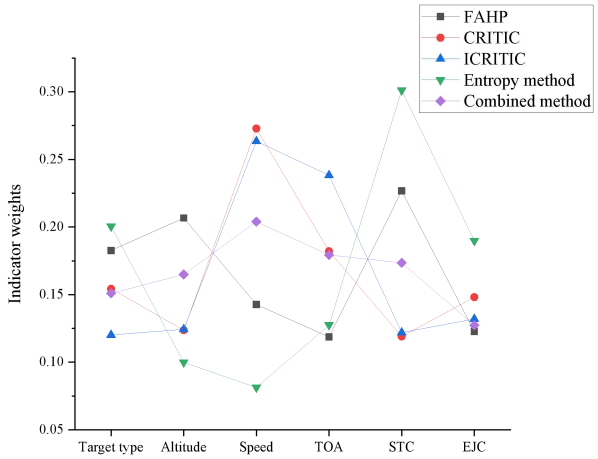


Figure 1 Comparison of Indicator Weights Obtained by Different Weighting Methods

3.3 Combination Empowerment

According to Equation (15) λ_1=0.6967, λ_2.7174=0, obtain the combined weight vector W_c=$(0.1510,0.1649,0.2039,0.1793,0.1735,0.1274)^{T}$.

3.4 Target Threat Level Assessment

The threat degree of each target was evaluated using the TOPSIS method, and the target threat degree under different empowerment methods is shown in Table 5 and Figure 2.

Table 5 Target Threat Level under Different Empowerment Methods

Target Number	FAHP	ICRITIC	Combined Method
1	0.4384	0.4965	0.4746
2	0.5627	0.7045	0.6288
3	0.4451	0.4845	0.4636
4	0.5420	0.6302	0.5813
5	0.6022	0.4678	0.5299

6	0.4000	0.4471	0.4251

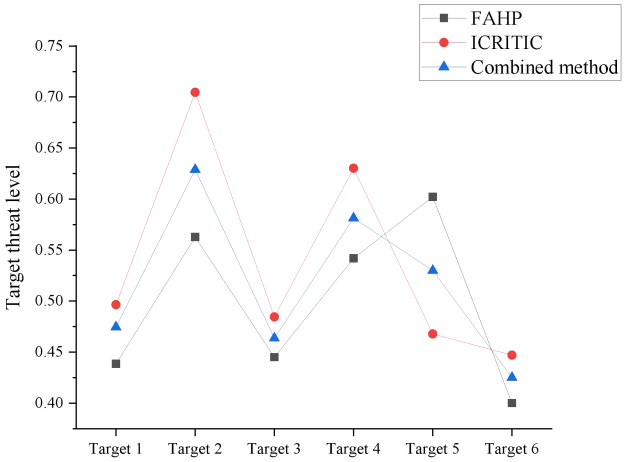


Figure 2 Threat Level of Target under Different Empowerment Methods

3.5 Results Analysis

From the empowerment, using the FAHP method to find the index importance order is: route shortcut, height, type, speed, interference ability, flying time, consistent with the evaluation in the matrix A, reflect decision makers experience and preferences, but the interference ability and flying time weight value is low, failed to fill combined with the actual situation and the use of objective information. CRITIC It is more important to find the target speed and flying time, Mainly because the standard difference of each index is not large, While these both are less correlated with other indicator data, Therefore, give a greater weight; The entropy method considers only the degree of index variation, In the data presented here, Give more weight to the route shortcuts, But if there are more target types, Large differences in indicators such as flight speed, The results are different; ICRITIC Method CRITIC, method with the entropy method, The obtained results were similar to those of the CRITIC, More comprehensive consideration of the volatility, correlation, discreteness of the data, However, the proposed method relies heavily on the data information, It is lacking in considering the purpose of evaluation. Based on the thought of maximizing FAHP method and ICRITIC method, find the objective weight coefficient is slightly larger, get each index importance order: speed, flying time, road shortcut, height, type, interference ability, on the results of the two evaluation methods are complementary, the data is relatively scattered, more easy to decision-making, rationality is stronger.

From the perspective of the target threat degree, using the FAHP method, the ICRITIC method, and the combined empowerment method, the threat degree ranking of the target 1 to 6 is S, respectively$_z$=（5,2,4,3,1,6）、S$_k$=（3,1,4,2,5,6）、S$_c$=（4,1,5,2,3,6） .It can be seen that the targets with high threat are low-altitude and ultra-low altitude, but the ranking of target types is different. The evaluation results of the combined empowerment method are more comprehensive considering the decision-making factors and more usable.

4 CONCLUSION

This paper proposes the method of evaluating the air target threat based on FAHP-ICRITIC combination empowerment. Firstly, the FAHP and ICRITIC methods are used to determine the threat index weight, and secondly, the combination weight coefficient of the target threat degree by TOPSIS method. Finally, the threat index weight and target threat degree under different methods are compared and analyzed to verify the effectiveness of the proposed method. The results show that the combined empowerment method can reasonably determine the subjective and objective weight coefficient, comprehensively consider the decision experience and the information provided by the data itself, effectively reduce the uncertainty of the commander in the subjective decision process, and provide a better basis for the command decision.

COMPETING INTERESTS

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

REFERENCE

[1] Tang Zhili, Zhang An. Research on the theory and methods of battlefield threat estimation. Fire and Command and Control, 2011,36 (09): 1-4.

[2] Bao Junchen, Han Daowen, Cheng Li, et al. Assessment of penetrating air combat aircraft based on Hierarchical analysis . Modern defense technology, 2023, 1-7.

[3] Yang Xiao, Wang Yumei. Study on operational target optimization based on improved hierarchical analysis. Computer Applications and Software, 2018,35 (04): 28-32.
[4] Zhao Manyun, Zhang Lei. Evaluation method of missile brigade of composite FAHP. Fire and Command & Control, 2022, 47(11): 170-174.
[5] Wang Guangyuan, Li Haomin, Zhu Dacheng, et al. Assessment of sea surface target threat degree based on entropy weight-gray correlation method. Command & Control and Simulation, 2023, 45(04): 57-61.
[6] Huang Liwei, Mu Yang. Escort Operations based on Principal Component Analysis. Fire and Command and Control, 2018, 43(08): 147-150 + 155.
[7] Xu Yuheng, Cheng Siyi, Pang Mengyang. Dynamic radiation source threat assessment based on CRITIC-TOPSIS. Journal of Beijing University of Astronautics, 2020, 46(11): 2168-2175.
[8] Fang Chengzhe, Kou Yingxin, Xu An, etc. Evaluation of VIKOR air combat threat based on AHP-CRITIC combination empowerment. Electro-optic and control, 2021, 28(02): 24-28.
[9] Zhou Hongbo, Zhang Jincheng. Gray target threat assessment based on combination weights. Fire and Command & Control, 2018, 43(10): 143-147.
[10] Lu Yunfei, Li Linlin, Zhang Zhuang. Research and application of the empowerment method of decision index combination. Computer Engineering, 2018, 44(01): 84-90.
[11] And Yelong, Zhang Pengfei, Zhao Yongjuan, et al. The TOPSIS method based on combined empowerment of CRITIC and AHP . Journal of Artillery Launch and Control, 2023, 1-9.
[12] Deng Bao. Research and application of index weight determination method based on portfolio empowerment method. Electronic information confrontation technology, 2016, 31(01): 12-16.
[13] Hua-you Chen. Combination assignment method based on difference maximization in multi-attribute decision making. Systems Engineering and Electronic Technology, 2004, (02): 194-197.
[14] Peng Tao. Construction of a target threat assessment model based on electric-fire integrated air defense. Fire and Command & Control, 2020, 45(03): 100-104.
[15] Yang Junjia, Wu Cuixia, Xiao Jian. Dynamic assessment of the SPA method. Fire and Command and Control, 2021, 46(05): 6-11.
[16] Wang Siyuan, Wang Gang, Zhang Jiarui. Evaluation method of air defense target threat based on the TOPSIS method. Journal of Attack and Guidance, 2019, 39(06): 171-176.
[17] Li Quen, Zhou Zhongliang, Zhang Xiaojie, et al. Aerial target threat assessment based on integrated empowerment and TOPSIS method. Electro-optic and Control, 2022, 29(05): 39-42 + 47.
[18] Yin Ruijie, Yang Zuobin, Wu Cuixia, et al. The TOPSIS method of threat based on grey theory. Ship Electronic Engineering, 2021, 41(11): 107-110.
[19] Kong Shangping, Zhang Hairui, Liao Xianping, et al. Aerial target threat assessment method based on AHP and entropy weight method. Tactical Missile Technology, 2018, (01): 79-84.
[20] Lv yuejin. Sort of fuzzy hierarchy analysis based on fuzzy consensus matrix. Fuzzy System and Mathematics, 2002, (02): 79-85.
[21] Song Dongmei, Liu Chunxiao, Shen Chen, et al. Multi-objective and multi-attribute decision-making method based on the subjective and objective empowerment method. Journal of Shandong University (Engineering Edition), 2015,45 (04): 1-9.
[22] Zhang Chihui, Zhang Yi, Li Jianning. Target selection model based on improved hierarchical analysis and entropy weight method. Journal of Probe and Control, 2022,44 (06): 117-121.
[23] Li Jingfeng, Xiang Huach, Yan Yarong, et al. Combination empowerment evaluation method based on distance maximization and its application. Fire and Command and Control, 2020, 45(06): 18-23 + 29.

Journal of Computer Science and Electrical Engineering
ISSN: 2663-1946
DOI: 10.61784/jcsee3013

CLOUD BASED SLEEP QUALITY MONITORING AND EVALUATION SYSTEM

Xu Chen, DeQuan Kong*
School of Information and Electronic Technology, Jiamusi University, Jiamusi 154007, Heilongjiang, China.
Corresponding Author: DeQuan Kong, Email: 893907285@qq.com

Abstract: To facilitate the monitoring and assessment of sleep conditions, this paper designs a cloud-based sleep quality monitoring and evaluation system. The system comprises a signal acquisition and processing module and a cloud platform. The signal acquisition and processing module connects to the cloud platform via Bluetooth, enabling the collection, reception, processing, and storage of various signals during sleep. The cloud platform employs a Naive Bayes sleep staging algorithm to classify sleep stages and assess sleep quality.
Keywords: Non-contact sleep monitoring; Naive bayes sleep staging algorithm; Smart device terminal; PVDF

1 INTRODUCTION

Sleep is a crucial natural physiological need. A survey in the "2022 China National Health Sleep White Paper" found that over half of the respondents have experienced sleep problems, which have gradually become a significant factor impacting people's work and daily life. This indicates that sleep issues are increasingly prevalent among residents in China. However, fewer than 6% of hospitals nationwide have established sleep monitoring facilities, and the quality of existing sleep medicine centers varies, failing to meet the public's demand for sleep quality services. Sleep staging is fundamental to sleep research and a key step in assessing sleep quality. Therefore, designing an affordable and user-friendly sleep quality monitoring and evaluation system is of great practical significance both nationally and globally at this stage[1]. This paper aims to develop an intelligent monitoring system capable of assessing sleep quality, utilizing non-intrusive sensor technology to monitor subjects during normal sleep[2]. The hardware of the system performs real-time acquisition of physiological signals, while the software handles data processing and analysis[3], ultimately generating a sleep quality evaluation report.

2 SYSTEM OVERALL PLAN

The overall architecture of the cloud-based sleep quality monitoring and evaluation system is shown in Figure 1.
The system's smart device terminal, heterogeneous network transmission, cloud server, and web client are each designated for specific functions: sleep assistance, monitoring of sleep physiological data, intelligent data analysis, anomaly detection and alerting, and customized intelligent features.

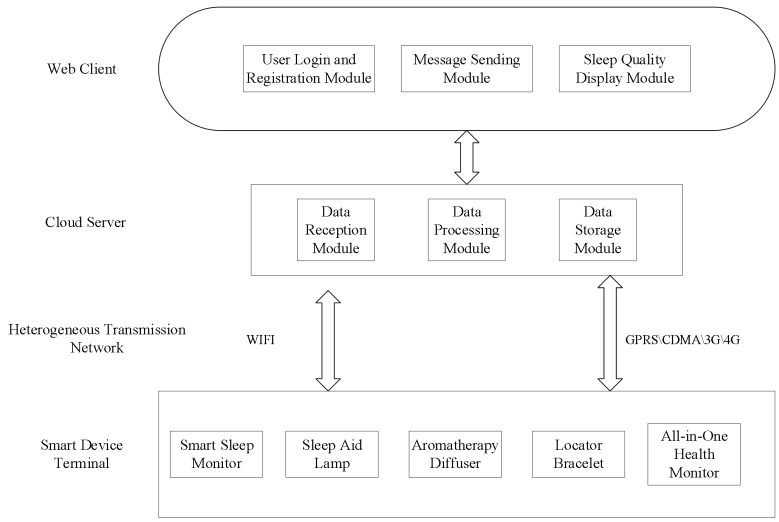


Figure 1 Overall Architecture of the Sleep Quality Detection System

To reduce software development complexity and make application development more flexible and efficient, a three-tier software architecture, as shown in Figure 2, has been designed. This architecture consists of a user interface layer, a business processing layer, and a data analysis layer. The distributed three-tier architecture aligns with the software development principle of "low coupling, high cohesion" [4]. This design separates interface display, business processing, and data support.

The user interface layer, which is the web page, allows users to interact directly with the application. The business processing layer is the core of the software; this project will use an improved Naive Bayesian sleep staging algorithm [5] to analyze physiological sleep data within a 1-minute window, determine the current sleep phase of the user, and generate a corresponding sleep quality report after a sleep monitoring session is completed. The data service layer is primarily responsible for collecting information from the smart device terminal, handling database operations, storage, and processing data requests from the business processing layer.

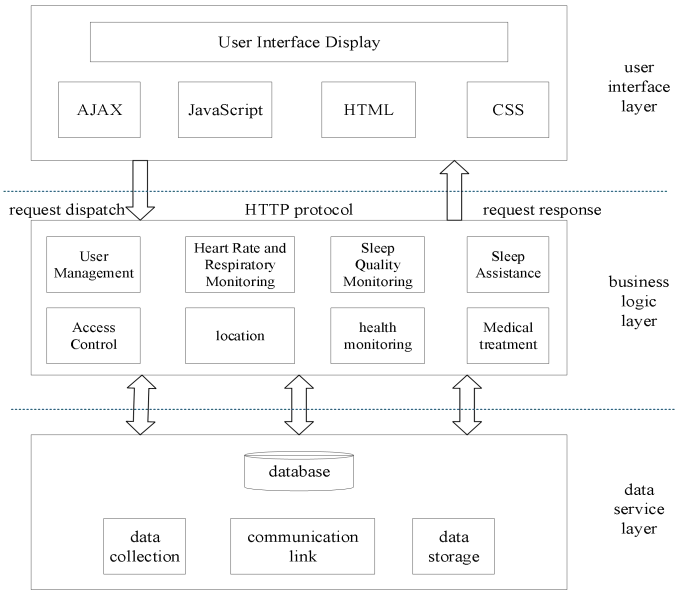


Figure 2 Sleep Monitoring Software Architecture

3 IMPROVEMENTS TO THE NAÏVE BAYES SLEEP STAGING ALGORITHM

In traditional Naive Bayes sleep staging computations, heart rate, respiratory rate, and heart rate variability are used [6]. Building upon this, the Naive Bayes classification method evaluates different types of data across four sleep stages and selects the most prevalent type as the final evaluation metric. However, there is a significant probability that the classification of the four stages may occur in pairs, presenting a dilemma for traditional Naive Bayes sleep staging methods, which may not provide intelligent classification.

To ensure the traditional Naive Bayes sleep staging algorithm functions correctly and to avoid issues such as those caused by high heart rate leading to respiratory rate problems, low heart rate or low respiratory intensity, and 0.5:0.5 sleep phase distribution due to high heart rate and respiratory intensity, improvements have been made. The improved Naive Bayes sleep staging algorithm process is shown in Figure 3.

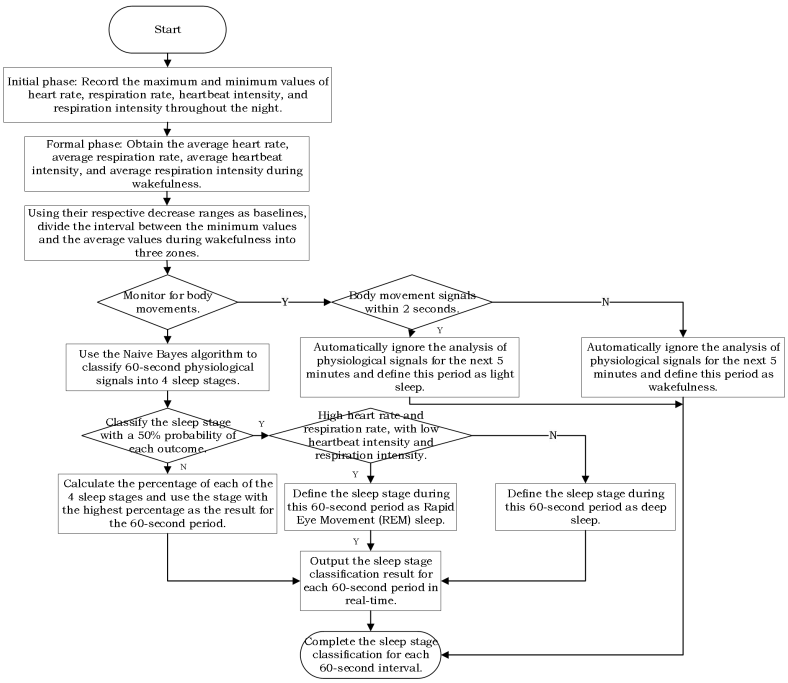


Figure 3 Naive Bayes Sleep Stage Classification Algorithm Flowchart

4 DESIGN OF THE CLOUD SERVER

The main functions of the cloud server are to receive, store, and process the sleep physiological information sent by the smart sleep monitor. The cloud server is primarily composed of three components: reception, storage, and processing. The data flow within the cloud server is illustrated in Figure 4.

Figure 4 Data Flow to Cloud Server

4.1 Data Processing Logic of the Data Reception Module

After the Gateway Worker module of the cloud server receives real-time sleep physiological data sent by the smart sleep monitor [7], it processes the data primarily through three event callback functions in `Applications/Your App/Events.php`. When the smart sleep monitor establishes a Socket connection with the cloud server, the `on-Connect` event is triggered. The callback function for this event displays the login information of the smart device on the cloud server's control panel. When the smart sleep monitor disconnects from the cloud server, the `on-Close` event is triggered, and the `Close` function displays the device's logout information on the cloud server's control panel. Each time the smart sleep monitor sends physiological data to the cloud server, an `on-Message` event is triggered. The processing logic for this callback function is illustrated in Figure 5.

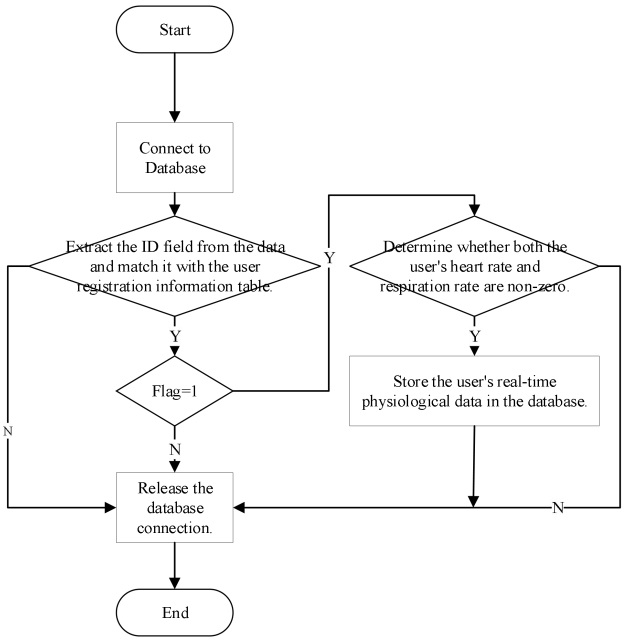


Figure 5 Data Processing Logic Diagram for the Data Reception Module

4.2 Design of the Data Storage Module

The data storage module is primarily used to store user information, such as basic user details and sleep quality assessment parameters.

4.3 Design of the Data Processing Module

Since the core module of the sleep monitoring software is the sleep monitoring module, the focus is on implementing the improved Naive Bayes sleep staging algorithm. This algorithm analyzes the user's physiological sleep data every minute to determine the current sleep phase and generates a sleep quality report at the end of each sleep monitoring session. The relevant data is then stored in a historical records table. Due to development challenges and proficiency issues, the backend data processing module of the system is implemented using PHP, with Sublime Text 3 as the development tool. The relationship between the sleep monitoring software and the data processing module is illustrated in Figure 6.

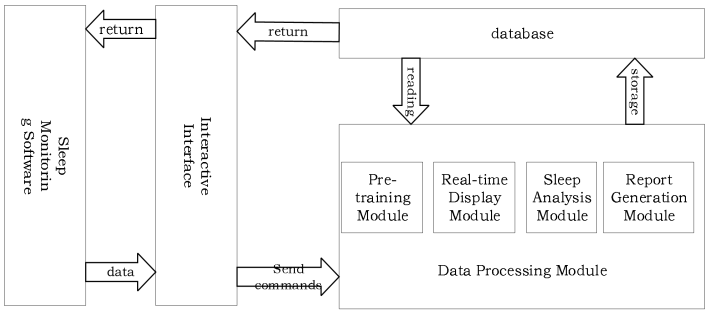


Figure 6 Relationship between Sleep Monitoring Software and Data Processing Module

(1) The functions of the preliminary training phase are mainly implemented by the preliminary training module. This involves collecting the maximum and minimum values of parameters such as heart rate, respiratory rate, and respiratory intensity throughout the entire sleep period of the subject. Measurements of these maximum and minimum values are taken, and the average heart rate, respiratory rate, heart intensity, and respiratory intensity for the first 5 minutes are assessed.
(2) The functions of the official sleep monitoring phase are accomplished through the real-time display module. This module acquires real-time physiological sleep data from the user, analyzes their sleep state, and tracks statistics on their movements, such as turning over, breathing pauses, getting out of bed, and waking times. Additionally, when determining the user's body movement status, the current sleep stage is identified based on the duration of body movements, without considering the physiological sleep signals from the subsequent 5 minutes. Based on this, the user's current state is assessed according to their actual conditions and the set alarm thresholds.
(3) The function of detecting body movement signals and the period beyond the 5 minutes after body movement cessation is mainly handled by the sleep analysis module. Using the Bayesian algorithm, the module determines the sleep phases within 1-minute segments of body movement signals, and performs statistical analysis to obtain the user's real-time sleep phase.
(4) The report generation module produces a complete sleep quality report after each sleep monitoring session. It also stores the relevant data from the sleep quality report in a historical records table, allowing users to easily query their past sleep data.

5 SYSTEM FUNCTION TESTING

Based on the actual application scenarios and the development progress of the sleep monitoring software, the testing process for this software has been designed to include three main phases: unit testing, integration testing, and system testing. The details are as shown in Table 1.

Table 1 Software Testing Process Design Table

Phase	Testing Method	Style
1	Unit Testing	Unit testing verifies whether the smallest units of software design—such as program modules or functional modules—contain errors, correctly implement their functions, and meet performance and interface requirements. Unit testing is typically performed using white-box testing techniques.
2	Integration Testing	Integrate the tested unit modules into a system or subsystem and then perform testing, focusing on the interfaces between different modules. Integration testing is typically conducted using black-box testing as the primary approach, supplemented by white-box testing.
3	System Testing	After completing unit and integration testing, conduct joint testing with other system components, such as data collection gateways and smart device terminals. The main goal of system testing is to verify the correctness of data interactions and response times between different components

Testing of various functional modules of the software is conducted on the established testing platform. During the functional testing, issues such as interface display and browser compatibility are also checked. For example, real-time monitoring is tested, as shown in Table 2.

Table2 Real-time Monitoring Test Case Table

Test Case Name	Real-time Monitoring
Test Objective	1.Verify that the system can display the user's current sleep physiological data in real-time.

Test Steps	2.Verify whether the system can receive anomaly alerts based on user-defined alarm thresholds. 3.Verify whether the system can assess the overall quality of the user's sleep segment when the sleep monitoring session ends. 1.Start the Gateway Worker module on the cloud server to receive data sent by the smart sleep monitor. 2.Log in using the sleep test subject's account and navigate to the "Real-time Monitoring" page. 3.Click the "Start Monitoring" button and compare the real-time data displayed on the page with the raw data on the Gateway Worker control interface. Check if data exceeding the alarm thresholds is highlighted in red. 4.Click the "End Monitoring" button and verify if the sleep quality analysis section on the page generates the appropriate chart report.

When the user is ready to start sleep monitoring, the Gate-way Worker module on the cloud server is activated to receive data sent by the smart sleep monitor. At this point, the user logs into the Web client using their sleep monitoring account, navigates to the "Real-time Monitoring" page, and clicks the "Start Monitoring" button. The start time is recorded on the page, the status changes to "Monitoring," and the data on the page is updated in real-time to match the update speed of the raw data on the cloud server's Gate-way Worker control interface. Any real-time data exceeding the custom alarm threshold is highlighted in red. When the user's sleep session ends, clicking the "End Monitoring" button generates a corresponding chart report in the sleep quality analysis section of the page, which matches the data recorded in the database table `sm_sleep_report`. This functionality test is passed.

This project provides a comprehensive introduction to the overall architecture of the sleep monitoring system, presenting a detailed design for a non-contact sleep quality monitoring system based on a cloud platform. It also explains the design concepts of each functional component in relation to practical conditions. The accuracy and timeliness of monitoring sleep quality are directly related to the detection quality of heart rate and respiratory rate [8]. The design described in this paper meets the demand for an affordable and user-friendly sleep quality monitoring device, offering excellent application prospects and high market value.

COMPETING INTERESTS

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

FUNDING

The project was supported by Sleep Quality Monitoring and Evaluation System Based on Cloud Platform (202210222139).

REFERENCES

[1] Zeng Jin, Lao Zhihui, Xu Mianqu. Design of an Embedded Sleep Posture Monitoring System Based on Infrared Cameras. Electronics Technology & Software Engineering, 2019(15): 168-170.

[2] Yang Jiafeng, Tong Jijun, Jiang Lurong, et al. Design and Implementation of a Non-Intrusive Sleep Respiratory Monitoring System. Computer Engineering and Applications, 2022, 58(05): 302-308.

[3] Liang Chao, Wang Peng, Cao Beibei, et al. Wearable Intelligent Sleep Quality Detection System. Journal of Electronic Measurement and Instrumentation, 2018, 32(05): 159-167.

[4] Yu D, Chander A, Islam N, et al. JavaScript instrumentation forbrowsersecurity. Acm Sigplan Notices, 2007, 42(1): 237-249.

[5] Jennifer Sloboda, Manohar Das. A Simple Sleep Stage Identification Technique for Incorporation in Inexpensive Electronic Sleep Screening Devices. Aerospace and Electronics Conference, Proceedings of the 2011 IEEE National, 2011: 21-24.

[6] Wu Taofeng, Gan Shuqi, Ouyang Chengjun, et al. Portable Sleep Quality Monitoring Device Based on Arduino. Electronics Technology & Software Engineering, 2021(18): 61-62.

[7] Zhong Linqian, Huang Jun, Liu Can, et al. Design and Implementation of Cloud-Based Sleep Monitoring Software. Information Communication, 2019, (03): 100-103.

[8] Dai Fengzhi, Lu Peng, Zhu Yuxuan. Design of a Multi-Sensor Sleep Monitoring and Assessment System. Foreign Electronic Measurement Technology, 2022, 41(04): 126-133.

Journal of Computer Science and Electrical Engineering
ISSN: 2663-1946
DOI: 10.61784/jcsee3014

EXPLORATION OF TEACHING REFORM IN ELECTRONIC TECHNOLOGY PRACTICE

WenMing Wang*, Ji Huang
Experimental Training Teaching Management Department, West Anhui University, Lu'an 237012, Anhui, China.
Correspondence Author: WenMing Wang, Email: hellowwming@foxmail.com

Abstract: Electronic technology practice is a fundamental and practical technical course, one of the compulsory courses for science and engineering students, and an important practical teaching link for shaping high-quality and versatile talents in the new era. This article elaborates on the exploration and implementation measures of the teaching reform of the electronic technology practice course in our school from the perspectives of blended online and offline teaching, virtual and real simulation experiments, CDIO engineering practice teaching, school enterprise collaborative education, and ideological and political education. The implementation effect of the course is good, and students have gained practical engineering experience from experiments, exercised their ability to solve problems, consolidated their professional knowledge, and established correct values.
Keywords: Electronic technology practice; CDIO(Conceive-Design-Implement-Operate); Teaching reform

1 INTRODUCTION

Electronic technology practice is an important part of practical teaching in applied undergraduate colleges. This practical teaching has low requirements for teaching conditions, small investment, and quick results, but it is of great significance for improving students' practical ability, innovation ability, and teamwork ability. The course focuses on engineering practice training, emphasizing the cultivation of students' practical abilities, the ability to independently acquire knowledge, and the ability to propose, analyze, and solve practical problems. It establishes a process foundation platform for students and enables them to acquire modern electronic process knowledge [1-5].

Electronic technology practice is a fundamental and practical technical course, which is one of the important compulsory courses for students majoring in science and engineering at our university. It is also an important practical teaching link for shaping high-quality and versatile talents in the new era. Through training in electronic technology practice, students can gain initial exposure to the practical production of electronic products; Understand and master electronic process knowledge and practical skills; Master the knowledge of safe use of electricity; Cultivate students' hands-on ability and innovative consciousness. Through practice, students are trained in learning ability, practical ability, innovation ability, and teamwork ability, laying a foundation for engineering practice in their future work.

The Electrical and Electronic Experiment Center of our school's Experimental Training Teaching Management Department undertakes the teaching work of "Electronic technology practice" for our science and engineering majors every year, and has been established for a total of 20 years. Based on the actual situation of our school's various professional training programs and the teaching needs of different majors, the electronic technology practice courses are divided into two categories: "Electronic Technology Practice A" and "Electronic Technology Practice B". Among them, "Electronic Technology Practice A" is designed for students majoring in non- electrical engineering at our university, and the teaching content is basic electronic practical experiments; "Electronic Technology Practice B" is designed for students majoring in electrical engineering at our university, with a focus on comprehensive and design oriented electronic practical experiments.

2 COURSE TEACHING OBJECTIVES

Through the study of this course, students will establish a sense of safe operation and establish basic practice operation norms. By standardizing operations and reporting requirements, guide students to prioritize integrity. In different practice content explanations and project training, based on current events and trends, combined with course and professional characteristics, introduce character introductions, case analysis, etc., while imparting course knowledge to students and establishing correct values. Increase labor education for students, organically integrate ideological and political elements, cultivate students' professional skills and innovative consciousness, and cultivate their socialist core values of "patriotism, dedication, and friendliness", enhancing their cultural soft power and core competitiveness.

Enable students to master the knowledge of safe use of electricity, learn the essentials of safe operation, cultivate a rigorous scientific style and a careful and pragmatic work style; Enable students to understand the categories, models, circuit symbols, main performance, and general selection knowledge of commonly used electronic components; Familiarize students with the basic knowledge and principles of electronic product assembly and welding, proficiently master manual electronic welding techniques, and assemble, weld, and debug electronic small products, and master the use of commonly used electronic instruments; Learn the correct usage methods of commonly used electricians, electronic instruments, and electronic devices, as well as measurement techniques for basic parameters of unit circuits; Enable students to master the SMT reflow soldering process and complete the assembly, soldering, and debugging of

the minimum system board for microcontrollers. At the same time, they should also master the use of the professional software Altium Designer to complete circuit schematics, printed circuit board diagrams, and circuit board design and production.

By combining theory with practice and hands-on practice, deepen the understanding of theoretical knowledge such as electronic circuits, and improve the ability to identify, analyze, and solve problems; Master the basic knowledge, methods, and skills of electrical experimental internships; Practice evidence-based academic discussions, write complete experimental records in a standardized manner, accurately process experimental data, scientifically and rigorously present results, and cultivate a scientific attitude of seeking truth from facts, being rigorous and trustworthy.

In different practice content explanations and project practices, based on current events and hot topics, combined with course and professional characteristics, introduce character introductions, case analysis, etc., while imparting course knowledge to students and establishing correct values. Increase labor education for students, organically integrate socialist core values and excellent traditional Chinese culture education, and establish the "Four Confidences" of socialism with Chinese characteristics. While cultivating students' professional skills and innovative consciousness, we should guide and educate them to establish a Marxist worldview and methodology, cultivate students' rigorous and pragmatic scientific attitude and academic ethics, cultivate students' scientific spirit of continuous exploration, courage to climb new heights, perseverance, and innovation, cultivate students' patriotism, and strive for the ideals and beliefs of realizing the Chinese Dream.

3 COURSE TEACHING DESIGN

3.1 Integrating Online and Offline Blended Teaching Practices

Establish an online course resource library using the Chaoxing Learning Platform. Complete practice project content in offline laboratory, provide course resources online, publish course assignments, upload assignments, and complete assignment review. Online teaching has a better effect on understanding knowledge and skills. In terms of cognitive level, it is relatively effective in knowing, comprehending, and applying knowledge. Offline teaching can better promote the comprehensive development of morality, intelligence, physical fitness, aesthetics, and labor skills. In terms of cognitive level, it is relatively effective for analysis, evaluation, and creation. Differentiated teaching content and objectives are necessary to maximize the unique advantages of blended learning, both online and offline.

Starting from the autumn semester of 2021, this course has undergone a small class online and offline smart classroom electronic technology practice teaching reform. In teaching, we use the Learning Pass mobile teaching platform to record the classroom teaching process, including check-in, questioning, investigation, in class discussion, uploading experimental phenomena, etc., to create an online classroom. Using Chaoxing Learning Platform for online intelligent teaching, the number of course activities has reached 64826, the cumulative number of course visits has reached 416096, and the cumulative number of interactions has reached 1191 (data statistics as of August 1, 2024).

3.2 Combining Virtual and Real to Ensure Efficient Implementation of Internship Practice

The electronic product production training workshop has a Surface Mounted Technology (SMT) production line, but students have limited practical workstations, high equipment operating costs, difficult observation of equipment component operation, and difficulty in hands-on operation. Through SMT virtual manufacturing simulation experiments, diversified teaching methods such as scenario simulation and interactive experience are adopted to enable students to have a sensory understanding of the SMT production line production scene, comprehensively understand the production process, equipment and operation process of the production line, display the operation of each component of the equipment in a 360 degree manner, complement the real teaching content in reality, combine reality and virtuality, enable students to operate on the real production line, cultivate their practical and innovative abilities, expand teaching capacity, and improve teaching efficiency.

The SMT process virtual simulation experiment teaching project, as shown in Figure 1. For the full process simulation of SMT electronic technology, virtual simulation experiments of the entire process of SMT electronic technology production line can be conducted online. By simulating the circuit board production workshop and corresponding equipment, and simulating the real circuit board mounting environment, students can have a clear understanding of current production technology and learn the experimental operation content of corresponding machines, providing strong support for future hardware production.

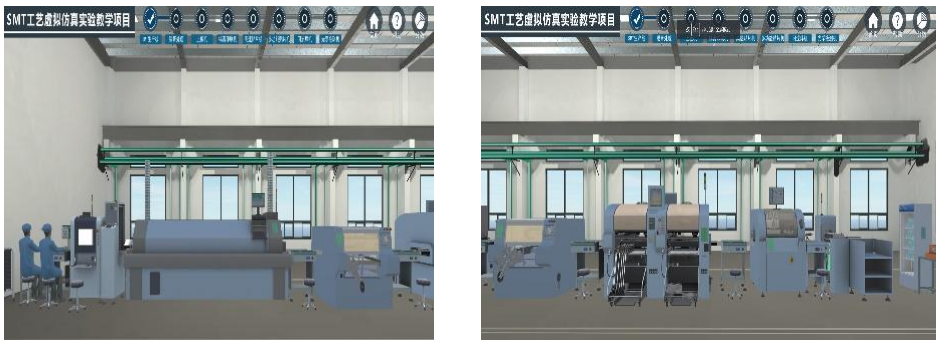


Figure 1 The SMT Process Virtual Simulation Experiment

3.3 Based on CDIO Engineering Practice Teaching

Constructing a CDIO project practical course teaching model for electronic technology practice courses guided by the new engineering concept. Using practical projects as the carrier, students as the main body, and teamwork as the form of cooperation, problem-based learning is adopted, and various teaching methods such as case teaching, discussion based learning, and experiential learning are used to effectively combine theoretical knowledge with practical application, guiding students to explore and innovate independently in the process of project implementation, closely integrating course content with actual production, and increasing the practicality of course content. At the same time, emphasis should be placed on the intersection, infiltration, and integration of curriculum, innovation and entrepreneurship education, and subject competitions, to achieve multi-party collaborative education in practical teaching [6]. The practical teaching process of CDIO project is shown in Figure 2.

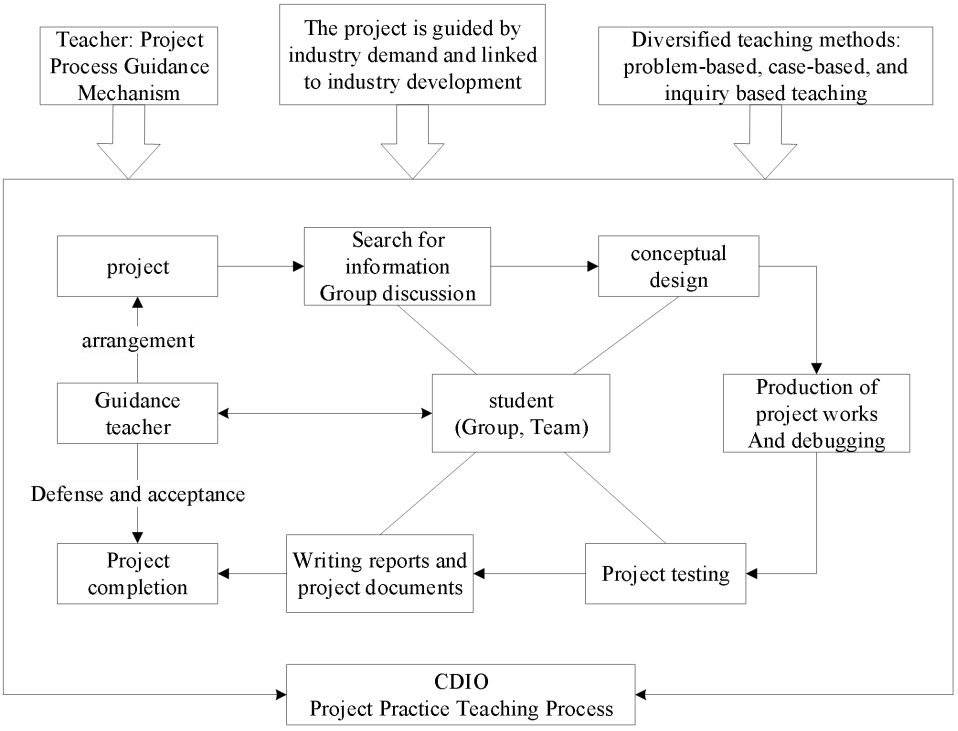


Figure 2 CDIO - Project Practice Teaching Process

3.4 Carry out School Enterprise Cooperation and Collaborative Education Model

Actively contact enterprises, introduce enterprise technical personnel to train students' professional skills, and carry out remote engineering training teaching; Combining industry trends and contacting enterprises to improve course teaching content and teaching models. Improve students' practical ability to independently solve problems, increase their new understanding of engineering consciousness, stimulate their innovative spirit based on professional knowledge, achieve the purpose of electronic process internship teaching, and meet the overall requirements of course teaching.

3.5 Combining Theory with Practice, Integrating Ideological and Political Education into the Classroom

Through the study of this course, students will establish a sense of safe operation and establish basic internship operation norms. By standardizing operations and reporting requirements, guide students to prioritize integrity. In the explanation of different internship contents and project implementation, based on current events and trends, combined with course characteristics and professional features, we introduce national craftsmen, character introductions, case analysis, etc., while imparting course knowledge to students and establishing correct values.

3.6 Actively Conducting Questionnaire Surveys

In the process of reforming, exploring, and implementing electronic technology practice teaching, to understand students' learning outcomes and related suggestions in a timely manner, we insist on conducting questionnaire surveys, student symposiums, and feedback meetings for each class of students to understand the impact of relevant teaching reforms on students' learning outcomes. The satisfaction survey results of students towards this course in the past three years are shown in Table 1.

Table 1 Student Satisfaction with this Course in the Past Three Years

Year	Very satisfied	Basically satisfied	commonly	dissatisfied
2021	93.3%	4.5%	2.2%	0
2022	95.5%	2%	2.5%	0
2023	96.2%	2.1%	1.7%	0

According to Table 1, it can be seen that the vast majority of students are satisfied with the setting of the practice program and the course content. Some students feel that the teaching effect of the course is average, and no student is dissatisfied with this course.

3.7 Statistical Analysis of Student Grades

The course grades are evaluated based on the completion status of students' practice projects at each stage, as well as their learning attitude and discipline, mastery of skills in daily teaching, completion of practice tasks, practice diaries, and summary of practice reports. The total score is synthesized according to different weights. Practice attitude and discipline (accounting for 20%), practical ability and level (accounting for 50%), practice diary and work summary (accounting for 30%), a total score is obtained, which is divided into five levels: excellent, good, medium, passing, and failing to give the overall score. Table 2 shows the distribution of course grades for two classes in the autumn semester of 2023.

Table 2 Distribution of electronic technology practice course grades for the autumn semester of 2023

Curriculum	Number s	Excellent/%	Good/%	Medium/%	Passing/%	Failed/%
Electronic Technology Practice A	34	6/17.6%	16/47.1%	9/26.5%	3/8.8%	0/0%
Electronic Technology Practice B	52	10/19.2%	16/30.8%	13/25%	13/25%	0/0%

From Table 2, it can be seen that the ratio of excellent to good is around 60%, indicating that during the practice process, the vast majority of students have a serious and upright attitude towards the practical course of electronic technology practice, which is mainly based on hands-on practice. They are particularly interested in the production of electronic small products during the practice, such as making radios, microcontrollers, PCBs, etc. During the practice, they not only learned relevant theoretical knowledge and process knowledge, but also cultivated good habits and innovative abilities of carefulness and patience, successfully completing the practice tasks at each stage and achieving the training objectives of electronic technology practice.

4 CONCLUSION

To meet the requirements of national engineering education certification and cultivate qualified engineering and technical talents, the Electronic Technology Practice course needs to continuously reform teaching content, improve teaching methods, closely follow the direction of technological development, combine with production reality, improve teaching methods, and strengthen the innovative elements of the Electronic Technology Practice course.

The teaching work of the Electronic Technology Practice undertaken by the Electrical and Electronic Experiment Center of our school's Experimental Training Teaching Management Department plays an enlightening and foundational role in cultivating students' hands-on operation, scientific research, new technology application, and innovative abilities. It also lays a solid foundation for the subsequent learning of related courses for our school's science and engineering students.

COMPETING INTERESTS

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

FUNDING

This work was supported by the Collaborative Education Project of the Ministry of Education of China (No.230800245304252), the Natural Science Key Project of West Anhui University (No. WXZR202310) and West Anhui University Quality Engineering Project (No. wxxy2020057, No. wxxy2021099).

REFERENCE

[1] Chai Zhijun, Wang Jixiang, Wang Yunxia. Reform and Exploration of Electronic Technology Practice Course. Heilongjiang Education (Theory and Practice), 2023(07): 69-71.
[2] Xu Jianjun, Wang Xilian, Huo Jingyi. Exploration of Ideological and Political Education in Electronic Technology Practice Course under the Background of Innovation and Entrepreneurship. Research and Practice of Innovation and Entrepreneurship Theory, 2020, 3(08): 72-74.
[3] Wang Jianhua, Sun Yaxing. Exploration of Ideological and Political Education in Electronic Technology Practice Course. China Education Technology Equipment, 2024(03): 61-64.
[4] Xie Zaijin, Liao Jihai, He Pu, et al. Exploration of blended online and offline teaching mode for electronic technology training. University Physics Experiment, 2023, 36 (02): 144-149.
[5] Zhao Shuling, Chen Yan. Exploration of Ideological and Political Education in Electronic Technology Practice Course under the Background of New Engineering Construction. Science and Education Literature Review, 2021(32): 127-129.
[6] Wang Wenming, Zhao Jiangdong, Huang Xiaojiao, et al. Exploration of Teaching Reform in Electronic Technology Practice Based on CDIO Innovation Model. Science and Education Guide - Electronic Edition, 2022(8): 128-130.

Journal of Computer Science and Electrical Engineering
ISSN: 2663-1946
DOI: 10.61784/jcsee3015

RESEARCH ON DESIGN AND QUALITY TRACEABILITY OF AERO-ENGINE PRODUCTION BASED ON CBR

JunShi Luo[1], Qiong He[1*], Shuang Yu[2]
[1]*School of Management Information and Engineering, Beijing Information Science and Technology University, Beijing, 100192, China.*
[2]*Aero Engine Academy of China, Beijing, 101399, China.*
Corresponding Author: Qiong He, Email: joanrun@126.com

Abstract: Productivity and traceability play an important role in aero-engine production. Aero-engine production is complex and there are many factors affecting its quality. CBR plays a significant role in this field. CBR model construction includes key technologies such as case entry, retrieval, revision and update. In terms of productive application, CBR can provide historical data reference, comprehensively evaluate performance indicators, consider cost factors, assist in evaluating manufacturability and reduce design risks. In the application of traceability, CBR is helpful to trace the source and use of parts, the history of production technology, the historical data of quality inspection, the history of failure and maintenance, and the performance and changes of raw materials. Through CBR technology, the productive design can be carried out in the product development stage, and the whole process of engine production can be effectively monitored and traced.
Keywords: CBR case-based reasoning; Aero-engine manufacturing; Productivity; Traceability

1 INTRODUCTION

In the global competitive environment, quality has become one of the key factors of national competitiveness, and building a quality power is an important strategic measure to promote the high-quality development of China's economy and enhance the comprehensive strength of the country. As a representative of high-end manufacturing industry, the quality level of aero-engine directly affects the competitiveness of China's aviation industry in the international market. The production of aero-engine involves a variety of parts and components, the production process is quite complicated, and there are many interference factors in the production process. At the same time, the operation process covers different links. The environment of the production workshop, the implementation of the production plan and the working environment of personnel will have an impact on the overall efficiency of the enterprise, which requires the design of product productivity in the product development stage[1]. Productive design originated from the technical document "Productive system guidelines-the five steps to success" published by the US Navy in 1999, in which the definition of productiveness is: measuring the relative ease of manufacturing products through indicators such as quality, time and cost is an ability to face the whole production process and consider whether the manufacturing process, materials, labor and cost meet the production standard products under certain circumstances. In addition, there are many factors that affect the quality of products, and every link from product design to delivery to customers will affect the final quality. In the past, it was difficult to effectively monitor the whole process of product manufacturing by simply relying on manual quality information collection and transmission. At the same time, with the increasing requirements of ISO9000 quality system and product quality supervision departments in China, it is of great significance to explore how to realize the traceability of engine production[2].

2 CBR APPLICATION STATUS

Case-based reasoning (CBR) is a kind of reasoning method based on empirical knowledge, which is suitable for fields where it is difficult to establish accurate mathematical models but there are a large number of case records. In 1982, the theory of dynamic memory structure put forward by Schank scholar was regarded as the embryonic form of this theory. After more than 30 years of development, it has been successfully used in many engineering fields[3]. The basic idea of CBR is to integrate historical successful cases into a case base and use information technology to retrieve similar cases to solve current problems[4]. It simulates human thinking, facilitates knowledge acquisition, and greatly improves the speed and quality of reasoning[5]. The reasoning of CBR usually includes four processes: case retrieval, matching/reuse, modification and storage. With the help of CBR technology and its theory, it is helpful to realize the manufacturability and traceability of aero-engines. Therefore, this paper explores the application of CBR in the two characteristics from this theory.

The initial target areas of CBR are problem diagnosis, decision order and strategic planning. However, with the continuous expansion of the understanding of CBR, CBR now has a very wide range of applications, which mainly cover: emergency decision-making, environmental monitoring, risk identification, fault diagnosis, product design, medical diagnosis, mapping and many other research directions.. Relying on a large number of cases to establish a case database, whether in the product design stage or in the production process traceability, can be traced. At the same time, with the continuous discovery of new problems in the production process, the database can be continuously expanded.

The work steps of CBR are mainly divided into four steps. First, the past cases are classified and integrated, and the database is established by relying on the past cases. Then, in the process of production or design, the relevant past cases are queried through data matching, and then the current plan is modified by combining the past cases. Finally, the modified scheme and modification process are stored in the database to further improve the database. The process is shown in the following figure1:

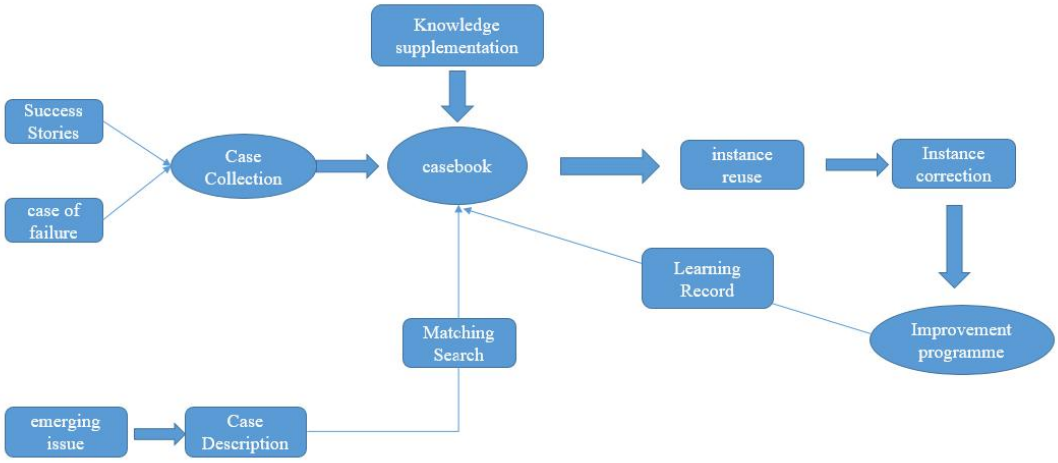


Figure 1 CBR Model

Enter the past success and failure cases into the case base, and then combine with expert opinions and national standards to improve. When new problems appear in the process of production and design, first describe the problems, then match them with the cases in the case base, then reuse the past solutions on the existing problems, and then modify and improve the problems to get the final solution, and finally enter the improved case plan into the case base again.

Key steps:

(A) case entry

The historical fault case data in the aero-engine production process are obtained, and they are preprocessed, classified and feature extracted in turn, so that the corresponding feature attributes of each category are obtained, and the fault case database is established accordingly.

Fault data collection: the collection of fault cases mainly comes from the collection and arrangement of troubleshooting and maintenance records in the engine production process, including fault maintenance manuals and maintenance service field work minutes. The fault maintenance manual covers descriptions and solutions of common engine faults by equipment parts manufacturers and users, including fault phenomena, fault equipment, fault causes and fault location. On-site work summary of maintenance service is the fault description and solution issued by equipment parts manufacturers and users when they carry out on-site maintenance for engines that have not occurred. In order to improve the quality and integrity of case data, the collected fault data are unified and eliminated. Then classify the fault data: divide the fault case data into mechanical system, electronic system and hydraulic system according to the production design system. Secondly, fault case feature extraction: define the feature attributes of each fault case according to the standard, and the feature attributes of each fault case can be expressed as pi = (pi1, Pi2, Pi3, Pi4), (I = 1, 2, ..., n), where n is the number of fault cases. Specifically, Pi1 represents the operating environment, with multi-category definitions, including temperature, humidity, salinity, altitude, etc. Recording this information can provide a basis for quickly matching similar faults triggered in the same operating environment. Pi2: indicates the fault, which is represented by three levels, and can be represented as component level, part level and component level. Each fault needs to record its component, part and component, and if it is a part or component level fault, the low level need not be recorded. For example, the hydraulic leg fails, and its components are hydraulic cylinder seals, parts are hydraulic cylinders, and parts are hydraulic legs. This representation method can quickly establish the correlation relationship for the subsequent failures of the same level and the same category. Pi3: indicates the fault phenomenon. By extracting the characteristic words of continuous information data, the keyword of the fault phenomenon can be obtained, and the number of this attribute is not limited. This attribute is the core attribute of this system and the basis of analysis and reasoning. Pi4: indicates the troubleshooting method, and this attribute corresponds to the Pi3 attribute. Finally, the fault classification is entered into the case base: after the fault case features are extracted, the fault features are saved into the case base according to the production system type.

(B) Case retrieval

After the case collection is completed, the RBF neural network is trained by using the case data in the fault case base to obtain the trained RBF neural network. RBF neural network calculates the similarity between the vector of target case feature elements and the vector of known case feature elements through the excitation function of hidden layer, and RBF neural network is equivalent to a similarity calculation network[6]. In this system, the fault cases are segmented by Chinese word segmentation algorithm (jieba word segmentation method), and then the features of the cases are extracted by TF-IDF keyword extraction algorithm, and converted into vectors according to the word frequency weights.

In this embodiment, the BoW(Bag of Words) word bag model is adopted for the digitization of feature words. The vectorized feature words are used as the input of RBF neural network. In RBF network, the input vector is x = [x1, x2, …xn]T, which is generated by the fault phenomenon in the fault case. The output vector is y = [y1, y2, ... ym] t, which is the cause of the fault.

In this study, radial basis function is used, and its weighted network output is:

$$y_j(x) = \sum_{i=1}^{h} \varpi_{ij} g_i(x) = \sum_{i=1}^{h} \varpi_{ij} \cdot \exp(\frac{-\left\|x - c_i\right\|}{\sigma_i^2}), (i = 1,2,...h; j = 1.2....m) \tag{1}$$

Where, it is the output of the j-th node of the output layer, the output of the i-th node of the hidden layer, the weighting coefficient from the hidden layer to the output layer, the center and variance of the Gaussian function of the i-th node, ||||||| is the distance between the input x and, and n, h and m are the number of nodes in each layer. In RBF network, the center, variance and weighting coefficient of the hidden layer function are obtained through the learning and training process of the network to the actual case base. In the process of unsupervised learning, the center and variance of hidden layer basis function are solved; In the process of supervised learning, the weights between hidden layer and output layer are solved. The specific calculation process is as follows: $y_j(x)$ is the output of the jth node of the output layer, $g_i(x)$ is the output of the ith node of the implicit layer, c_i 、 σ_i is the center and variance of the Gaussian function of the ith node for ϖ_{ij} , $\|\cdot\|$ is the distance between the input x and c_i , n、 h、 mis the number of nodes in each layer. H centers are selected for clustering. For the radial basis of Gaussian kernel function, the variance is solved by the formula:

$$\sigma_i = \frac{c_{\max}}{\sqrt{2h}}, (i = 1,2,...h) \tag{2}$$

Where is the maximum distance between the selected center points. The weighting coefficient is directly calculated by the least square method, that is, the partial derivative of the loss function is solved to make it equal to 0, and the calculation formula is: $c_{\max}$ is the maximum distance between the selected center points. ϖ_{ij} is directly calculated by the least squares method to obtain, that is, the partial derivative about is solved for the loss function so that it is equal to 0. The formula is:

$$\varpi = \exp\left(\frac{h}{c_{\max}^2}\left\|x - c_i\right\|^2\right), (i = 1.2,...h) \tag{3}$$

Step 3, feature extraction is carried out on the fault to be diagnosed, and corresponding key feature attributes and corresponding word vectors are obtained; According to this, the number of clustering center points is selected, and all relevant cases in the case base are clustered by K-means according to the word vector, so as to find cases with the same category as the target cases, that is, similar cases; The fault features of each similar case are transformed into word vectors and then input into the trained RBF neural network for retrieval and reasoning, and the similarity with each similar case is output.

After extracting the feature attributes of the fault to be diagnosed (that is, the target case), TF-IDF keyword extraction algorithm is used to extract the features of the case, mainly to extract the keywords from the case, to obtain the keywords with high word frequency in the case and low word frequency in other cases in the case base, and to vectorize them at the same time.

The TF-IDF value of keyword mountain in the case is:

$$TF - IDF_{\varpi} = TF_{\varpi} * IDF_{\varpi} = \frac{n_{\varpi}}{\sum n} * \frac{\left|D\right|}{\left|j_{\varpi} + 1\right|} \tag{4}$$

Where is the number of occurrences of the word ω in this case, which is the sum of the occurrences of all words in the case; Is the number of cases containing the word ω, and 1 is added to avoid the denominator being 0, which is the total number of cases in the case base. n_{ϖ} is the number of times the vocabulary ω appears in the case, $\sum n$ is the sum of the occurrences of all terms in the case; $\left|j_{\varpi} + 1\right|$ is the number of cases containing the vocabulary ω, adding 1 is to avoid a denominator of 0, $\left|D\right|$ is the total number of cases in the case base.

Case retrieval reasoning: query past cases through key feature attributes, that is, retrieve case base. After vectorizing the extracted key feature attributes, the number of clustering centers is selected, and all cases in the case base are clustered by K-means according to the word vector, so as to find cases with the same feature attributes as the target case, that is, similar cases. The fault features of each similar case are converted into word vectors and input into the trained RBF neural network, and the similarity between the target case and each similar case is output at the output layer.

(C) case revision and update

Select fault cases with high similarity, determine the corresponding fault causes and solutions, and update the fault case base. According to the output results, the fault cases with high similarity are selected to locate the fault, find out the cause of the fault and its solution. If the same source case as the target case is retrieved, a definite solution can be

obtained and the target problem can be solved. Because the target case already has the same case, the target case can be discarded to avoid redundancy. If the source case similar to the target case is retrieved, the suggested solution of the target case is obtained by the solution of the similar case, and the suggested solution is modified according to the actual situation, and then the definitive solution is obtained, and the target case is saved as a new case in the case base. The above case correction is mainly based on the solution of the actual target case, which is mainly the process that technicians reason the failure according to their actual experience, the test results of technical instruments and the relevant technical principles of equipment, and correct the case after the actual failure problem is solved, so as to ensure the follow-up study of personnel and reference maintenance when similar failures are found, thus further improving the maintenance efficiency.

3 DESIGN APPLICATION BASED ON CBR

Productivity considers the problems in the whole manufacturing process, which includes manufacturability, assemblability and detectability. Manufacturability, assemblability and detectability mainly consider the technological ability of equipment and process equipment from the technical point of view, and see whether a product can be manufactured and assembled easily and can be detected. Besides the meaning of manufacturability, assemblability and detectability, producibility also includes whether the enterprise can meet the requirements in terms of equipment production capacity and personnel capacity when the products are produced in batches and at the required time of listing. Relying on CBR technology, this quality characteristic can be realized.

First of all, it can provide historical data reference. For example, a certain type of engine has excellent fuel efficiency under specific working conditions, which can be compared with the current fuel system design. If the key parameters are similar, it can predict the fuel efficiency potential. Secondly, it helps to comprehensively evaluate performance indicators, such as thrust, weight and durability, and can retrieve excellent cases and compare them. For example, specific materials and processes can reduce weight and increase durability, and if the current schemes are similar, it is expected to achieve good results. Furthermore, considering the cost factor, reviewing the cost composition and control strategy can predict the current cost, and learning from previous optimization measures can control the cost. In addition, assist in evaluating manufacturability, retrieve manufacturing problems and solutions of similar structures, prevent in advance, and ensure smooth production. Finally, reduce design risks, provide innovative case results, increase confidence in success, and promote cautious improvement in failure. During the construction of design case base, we can input the construction case base from six aspects: design object, design purpose, detailed parameters, problems, actual cost and actual effect. When designing and developing an engine, we search from five aspects: environmental requirements, performance requirements, target requirements, risk prevention and cost requirements, match relevant cases, and then modify the current design according to past cases. Application of CBR in designability can be seen in Figure 2.

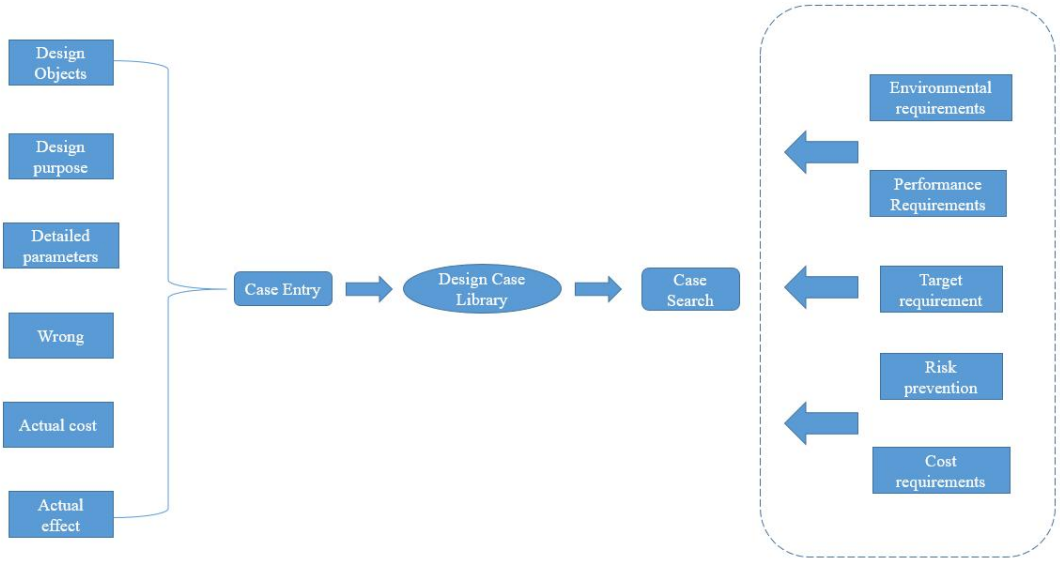


Figure 2 Application of CBR in Designability

4 QUALITY TRACEABILITY APPLICATION BASED ON CBR

Traceability means that every component, every process and raw material of military aero-engine can be clearly traced back to its source, production process and use. Once a quality problem occurs, the source and influence scope of the problem can be quickly and accurately located[7]. Depending on CBR technology, the process and parts can be traced back separately. First of all, CBR helps to trace the source and use of parts. Aero-engine is composed of many complex parts, each of which has its own specific supplier and production batch. Through CBR system, we can compare the information of parts used by the current engine with past cases. Secondly, CBR can trace and analyze the history of production technology. Different types of aero-engines may adopt similar production processes, but there may be

differences in details. Using CBR, we can review the effects of similar processes under different conditions. Furthermore, CBR helps to trace the historical data of quality inspection. In the process of aero-engine production, strict quality inspection will be carried out many times. Through CBR, the current test results can be compared with the results of similar engines in the same test link in the past. If it is found that the current test results are similar to a case with quality problems, further investigation and analysis can be quickly started to ensure the quality of the engine. In addition, CBR can also trace the failure and maintenance history. When the engine breaks down in use and needs maintenance, CBR can provide the handling methods and maintenance records of similar failures in the past. Finally, CBR also plays a role in tracing the performance and changes of raw materials. The performance of raw materials used in aero-engines, such as special alloys, may be affected by factors such as production batch and processing technology[8]. Through CBR, we can compare the performance data of raw materials currently used with those in past cases, so as to better predict and control the performance and reliability of the engine. As shown in the figure 3 below, the traceability function can be realized through the case list database:

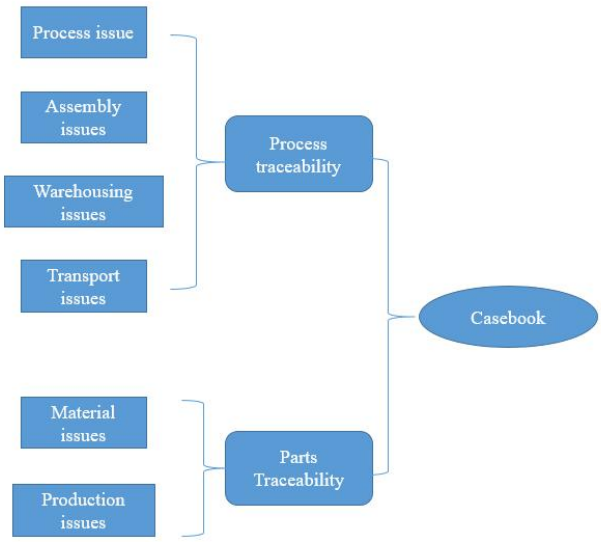


Figure 3 CBR Traceability Application

When building a case base, it should be entered from three parts: fault phenomenon, fault location and solution, as shown in Table 1:

Table 1 Case Entry Information Table

Fault phenomenon	Fault location	Knot solution
01 blade crack	Compressor blade	Strengthen raw material testing and manufacturing process control.
02 Turbine speed is unstable	turbine shaft	Check the wear of the shaft and repair or replace it.
03 insufficient air intake	air inlet	Clean the foreign body in the intake port and check the opening of the intake valve.
04 excessive engine vibration	Fan part	Balance the fan blades and check the installation and fixation.

5 SUMMARY

This paper aims to explore the application of case-based reasoning (CBR) in the manufacturability and traceability of aero-engines. As a representative of high-end manufacturing industry, the quality level of aero-engine is very important. In the production process, due to the variety of parts, complex production technology and many interference factors, it is necessary to carry out productive design in the research and development stage and realize effective monitoring and traceability of the whole process of product manufacturing[9]. The basic idea of CBR is to integrate historical successful cases into a case base to solve current problems. Its working steps include case collection, case reuse, case revision and learning record. In aero-engine production, the key technologies of CBR include case entry, case retrieval and case revision and update[10]. In productive application, CBR can provide historical data reference, compare current design with excellent cases, and predict fuel efficiency potential and performance effect; Consider the cost factor, estimate and control the cost; Assist to evaluate manufacturability and prevent production problems in advance; Reduce design risks and provide innovative case results. In the application of traceability, CBR is helpful to trace the source and use of parts, and compare the current parts information with past cases; Trace the history of production technology and analyze the effect of similar technology under different conditions; Trace the historical data of quality inspection and

compare the current inspection results with past cases; Trace the history of faults and maintenance, and provide treatment methods and maintenance records for similar faults; Trace the performance and changes of raw materials to better predict and control the performance and reliability of the engine. With the help of CBR technology, it is helpful to realize the producibility and traceability of aero-engines and is of great significance to enhance the competitiveness of China's aviation industry in the international market.

COMPETING INTERESTS

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

FUNDING

The research is funded by the Beijing Municipal Education Commission Research Plan General Project (grant number: KM202411232007)

REFERENCES

[1] Yang Jianjun. Product Design Manufacturability and Production System. Beijing: Aviation Industry Press, 2009.06.
[2] Xing Yichao. Quality Tracing Support System for Production Process of Multi-species and Small-lot Machining Workshop. Chongqing University, 2012.
[3] Fang Wenjuan, Li Shaojian. Research and application of case-based reasoning technology. Agricultural Network Information, 2005(1): 13-17.
[4] LIU Yong, LI Haichao, BOW Xizhong, et al. Fault diagnosis of aerospace measurement and control equipment based on case-based reasoning. Telecommunication Technology, 2017, 57(02): 236-242.
[5] GUO Haochen, WANG Han, LOU Lin, et al. Intelligent diagnosis technology of aircraft assembly process faults based on knowledge engineering. Aviation Manufacturing Technology, 2022, 65(07): 90-95+109. DOI: 10.16080/j.issn1671-833x.2022.07.090.
[6] Xia Shu. Slope reliability analysis based on radial basis neural network. Heilongjiang Transportation Science and Technology, 2024, 47(08): 32-35. DOI: 10.16402/j.cnki.issn1008-3383.2024.08.028.
[7] Meijian Xu, Yue Huang. Application of case-based reasoning technique in aviation maintenance. Science and Technology Innovation and Application, 2019(29): 182-183.
[8] Chen Xi. Construction of whole-process quality control system for aero-engine manufacturing based on digital transformation. Value Engineering, 2023, 42(15): 53-55.
[9] Cao Ming, Jin Quan, Zhou Jian, et al. Status, Challenges and Opportunities of Fault Diagnosis and Health Management of Civil Aviation Engines I: Fault Diagnosis and Prediction of Airway, Mechanical and FADEC Systems. Journal of Aeronautics, 2022, 43(09): 9-41+2.
[10] Huang Y. Construction of product quality evaluation system based on the whole life cycle. Electronic Quality, 2023(08): 64-67.

Journal of Computer Science and Electrical Engineering
ISSN: 2663-1946
DOI: 10.61784/jcsee3018

COMPUTATIONAL BIG MODEL-BASED STUDY OF PRIVACY PROTECTION MECHANISMS AND PROBLEMATIC USAGE BEHAVIOUR IN DIGITAL MEDIA

Jin Lu[1], Ji Li[2*]
[1]*Guangdong Key Laboratory of Big Data Intelligence for Vocational Education, Shenzhen Polytechnic University, Shenzhen 518000, Guangdong, China.*
[2]*Research Management Office, Shenzhen Polytechnic University, Shenzhen 518000, Guangdong, China.*
Corresponding Author: Ji Li, Email: liji@szpu.edu.cn

Abstract: With the rapid development of information technology, digital media has become an important part of people's daily lives. From social media to online shopping, from cloud computing to artificial intelligence, digital media not only greatly enriches people's lifestyles, but also brings unprecedented data privacy protection challenges. In particular, the wide application of computational big models (e.g., deep learning models, natural language processing models, etc.) has further exacerbated the complexity of data privacy protection. This paper aims to explore the privacy protection mechanism of digital media based on computational big models and analyse the resulting problematic use behaviour, with a view to providing references for research and practice in related fields, delving into the development of the convergence of computational big models and digital media, analyzing the manifestation of problematic use behaviors and privacy issues in digital media, exploring the challenges and countermeasures to prevent the privacy issues, and presenting the challenges and countermeasures to prevent privacy issues in digital media privacy protection mechanism under the computational big model. Successful privacy protection experiences and lessons of problematic use behaviors leading to privacy issues are summarized through case studies. Finally, the research results are summarized and future trends and research directions are outlined.
Keywords: Big model; Privacy protection mechanisms; Usage behaviour; Digital media

1 INTRODUCTION

In the social media environment, leakage sources such as the user himself, other Internet users, platforms and third parties may cause the leakage of user privacy information. The characteristics of digital media's multi-subject participation (users, friends, and platforms) make user privacy protection strategies more complex. Traditional online privacy protection is mainly achieved through corporate privacy practices, industry self-regulation and government regulation, however, the effectiveness of such corporate privacy protection strategies is widely questioned in the era of big data. Based on the shortcomings of existing privacy protection strategies, the EU proposes an approach that integrates engineering and strategic management to achieve selective and continuous minimization of information system privacy risks through technical and regulatory controls, a strategy to minimize privacy risks through technology and controls from a system design perspective[1]. With the popularity of mobile devices, people are increasingly using various mobile applications in their daily lives, and the competition among mobile application developers is becoming more and more intense. For developers, the collection and use of user data has huge business potential and is the basis for realizing personalized services, so developers will ask users for authorization to collect some information during the process of opening or using the app. However, with the rapid changes in IT, which makes the data diffusion method hard to be predicted, users are not sure of the consequences of authorization and are worried about the misuse of their private information[2-3], and due to the frequent leakage of online users' information and privacy in recent years, users have become more and more cautious about authorizing information collection behaviors[4]. Therefore, it has been explored how to design and optimize the mobile APP authorization interface, and how to collect and use privacy information, so as to provide better services for users, which is crucial for both APP developers and users[5]. Data leakage will not only cause economic loss to enterprises, but also damage consumer satisfaction and purchase intention[6]. Therefore, the problem of remediation after data leakage has become an important challenge for the Internet industry. However, service remediation of privacy data leakage is different from traditional service remediation in that it is characterized by a large number of victims and an immeasurable loss[7]. As a result, companies are increasingly interested in how best to conduct data breach remediation.

With the emergence and popularity of smartphones, mobile social networks have become part of an individual's daily life, providing new ways of communicating, socializing and maintaining friendships, thus improving people's life experience. However, more and more people have started to show some tendencies of problematic use in uncontrollable ways in untimely scenarios. Problematic use refers to unplanned and impulsive use of digital media, which usually has a negative impact on the user[8]. Research has shown that problematic use of digital media is detrimental to people's real-life relationships[9], job performance[10], academic achievement[11], and subjective well-being[12], among others. Much of the literature has examined the triggers of problematic behaviors, focusing on individual psychological characteristics, but neglecting the role of technological media. Studies have shown that media characteristics play an

important role in influencing user behaviour[13]. Therefore, there is a need to investigate the role of media technologies so that users and practitioners can better understand problematic use of digital media.

Privacy-preserving mechanisms are essential in the digital media landscape to protect user data and prevent problematic usage behaviors. With the increasing use of mobile devices, monitoring users and building detailed profiles for behavioral advertising has become common practice[14]. Digital trace data collection allows for the analysis of content usage on social media platforms, shedding light on textual and audio-visual content[15]. The use of the internet and social media has changed consumer behavior, leading to eWOM overload that can be mitigated by new tools and mechanisms[16]. In the healthcare sector, digital transformation has led to the development of privacy-preserving smart healthcare frameworks based on blockchain technology[17]. The National Strategy to Advance Privacy-Preserving Data Sharing emphasizes the importance of federally funded research in networking and information technology to protect user data[18]. Privacy in targeted advertising on mobile devices is a growing concern, with research focusing on usage behavioral patterns and interest-based ads targeting[19]. In the digital sphere, privacy-preserving mechanisms are crucial for protecting private attributes and data flows related to media usage and marketplace behavior[20]. Big data privacy solutions, such as differential privacy and homomorphic encryption, have been developed to protect user data in content and social media analysis[21]. As digital platforms continue to impact news and journalistic content, research is needed to address privacy concerns and protect user data[22]. Overall, the intersection of privacy-preserving mechanisms and problematic usage behaviors in digital media highlights the importance of protecting user data and ensuring ethical practices in data collection and analysis. Solutions such as blockchain technology, differential privacy, and federated learning play a crucial role in safeguarding user privacy in the digital age[23].

This paper provides an in-depth study of the development of the convergence of computational big models and digital media. Chapter 2 analyses the behavioral manifestations of problematic use of digital media and privacy issues. Chapter 3 explores the challenges and countermeasures to prevent privacy issues and introduces the privacy protection mechanism of digital media under the computational big model. Chapter 4 summaries successful privacy protection experiences and lessons learned from problematic use behaviors leading to privacy issues through case studies. Finally, the research results are summarized and future trends and research directions are outlined.

2 THEORY AND PROGRAMME

The study of privacy-preserving mechanisms and problematic use behaviors of digital media based on computational macro-models is a complex topic covering multiple aspects, which involves a number of domains such as technology, ethics, law and social impact. While digital media satisfy people's needs, they also cause serious privacy leakage, and the process is defensible with complex features such as multiple sources, interaction, blocking, network externalizes and concealment. In order to gain access to digital media, gain the trust of online users or share life experiences, etc., people need to provide their personal information such as name, ID number, mobile phone number, geographic location, etc. online, and share life experiences and family photos, etc., and platforms and third parties may also obtain private information such as the user's interpersonal network, identity and preferences through association detection, data mining and other technologies. This study will examine the potential vulnerabilities and threats posed by the increasing volume and complexity of data in the context of big data analysis. Furthermore, the research will explore the impact of these vulnerabilities on the confidentiality, integrity, and availability of data, as well as the potential consequences for organizations and individuals. Moreover, the study will investigate the efficacy of current data security measures in mitigating these risks and propose recommendations for enhancing data protection strategies in the age of big data analysis. Additionally, the study will assess the role of data encryption, access controls, and authentication mechanisms in safeguarding data integrity and preventing unauthorized access in the context of big data analysis. Moreover, this research will also analyze the role of data governance frameworks and regulatory compliance in ensuring data security in the age of big data analysis. This study aims to provide a comprehensive analysis of the multifaceted challenges and implications of data security in the era of big data analysis, with a particular focus on the role of encryption, access controls, authentication mechanisms, data governance frameworks, and regulatory compliance in safeguarding data integrity and preventing unauthorized access. Furthermore, the research will delve into the intricate interplay between data security measures and organizational practices, aiming to offer a holistic understanding of the complexities involved in ensuring robust data protection in the era of big data analysis. This comprehensive analysis will shed light on the intricate interplay between data security measures and organizational practices, providing a holistic understanding of the complexities involved in ensuring robust data protection in the era of big data analysis. Next, several basic theories covered in this paper will be analyzed.

2.1 Computational Big Models in Digital Media

The application of computational big models in digital media mainly includes three aspects of content recommendation, automatic generation, speech recognition and image processing, which are shown in Table 1.

Table 1 Application and Classification of Computational Big Models in Digital Media

Element	Connotation Description
Content	One of the major applications of computational big models in digital media is content

Recommendations[24]	recommendation. Through in-depth analysis of user behaviour, content features and other data, the big model is able to recommend personalized content for users, such as news, videos, music and so on. This process requires the model to process a large amount of user data, including but not limited to browsing history, clicking behaviour, purchase records, etc.
Automatic Generation[25]	Computational Big Model also has the ability to automatically generate content such as press releases, articles, video scripts, etc. This not only reduces the cost of manual creation, but also increases the efficiency and diversity of content production. However, this automated generation process also involves processing and analysing large amounts of data such as text and images.
Speech Recognition and Image Processing[26]	In the field of speech recognition and image processing, computational big models play an equally important role. Through deep learning of speech signals and image data, the models are able to achieve high-precision speech-to-text, image classification, target detection and other functions. The realization of these functions relies on the processing of sensitive information such as the user's voice and facial features.

2.2 Digital Media Privacy Protection Mechanisms

The digital media privacy protection mechanism mainly includes three parts: the innovative application of encryption technology, the construction of intelligent defence system, and the formulation and implementation of privacy protection policy, as shown in Table 2.

Table 2 Categorical Description of Innovative Applications of Cryptography

Element	Connotation Description
Innovative applications of encryption[27]	In digital media, data encryption is an important means of protecting privacy. Traditional encryption methods, although effective, appear to be incompetent in the face of computationally large models. Therefore, it is necessary to develop more efficient and intelligent encryption models by combining deep learning techniques. For example, using deep learning algorithms to train efficient encryption models can improve the efficiency of data transmission and storage while ensuring data security. In addition, the introduction of techniques such as differential privacy can provide rich data resources for the training of computationally large models while protecting user privacy.
Construction of Intelligent Defence System[28]	Intelligent defence systems based on computational macro-models can monitor network traffic in real time, automatically identify and intercept malicious traffic, and effectively prevent network attacks. These systems are able to cope with increasingly complex and covert attack methods by continuously learning and optimizing their own defence strategies. In addition, a phishing detection solution built using encryption technology can analyse the content and structure of emails to identify potential phishing attacks and activate an early warning mechanism in the first instance.
Privacy Policy Development and Enforcement[29]	A sound privacy protection policy is the basis for safeguarding user privacy. In the age of digital media, the purpose of data collection and use must be clear and legitimate, and data subjects must grant explicit consent to the collection and use of their data. At the same time, data collectors must adopt appropriate security measures to protect personal data and prevent data leakage and misuse. Governments and relevant organizations should strengthen the regulation of data collection and use to ensure that data collectors comply with relevant regulations and policies.

2.3 Problematic Use Behaviour and its Coping Strategies

The problematic use behaviors can be classified into three categories: data misuse and leakage, lack of user privacy awareness, and privacy leakage under technological unconsciousness. These are outlined in Table 3, together with recommended coping strategies.

Table 3 Description of Problematic Use Behaviors and Their Coping Strategies

Element	Connotation Description	Response Strategies
Data Misuse and Compromise[30]	Data misuse and leakage are common privacy protection issues in digital media. Some unscrupulous elements or internal employees may take advantage of system loopholes or mismanagement to illegally obtain, use or leak user data. Such behaviour not only violates users' privacy, but may also cause property damage and mental harm to users.	● Enhance system security: Regularly check and fix system vulnerabilities to ensure system security and stability. ● Improve internal management: Strengthen security education and training for employees, and formulate strict regulations on data access and use to prevent internal leakage. ● Establish an emergency response mechanism: In the event of a data leakage

		incident, activate the emergency response mechanism in a timely manner to quickly identify the cause and take measures to reduce losses. ● Raise users' privacy awareness: Raise users' awareness of privacy protection through publicity and education, case sharing, etc., so that users understand the importance of personal privacy and how to protect it. ● Simplify privacy policies: Simplify complex privacy policies into easy-to-understand language to ensure that users can fully understand and agree to the content of the privacy policy. ● Introduce third-party certification mechanism: Improve user trust in digital media platforms by introducing a third-party certification organization to assess and certify the privacy protection capabilities of the platforms. ● Enhance technology transparency: Increase transparency in the use of technology so that users understand the process and possible risks of data collection, processing and use. ● Limit the scope of data collection: Clarify the purpose and scope of data collection and avoid collecting unnecessary data to minimize the risk of privacy breaches. ● Provide user control: Give users control over their data and allow them to view, modify or delete their data at any time.
Lack of user privacy awareness[31]	While enjoying the convenience of digital media, users often neglect the protection of their personal privacy. Some users casually disclose personal information due to a lack of awareness of privacy protection, or authorize third parties to access personal data without understanding privacy policies.	
Privacy breaches in the technological unconscious[32]	With the popularity of smart devices and the development of AI technology, a large amount of data is automatically exchanged between smart devices without the user's knowledge. This technological unconscious privacy leakage poses a great risk to users.	

2.4 Problematic

In order to solve the first problem raised in this paper, combined with the principal-agent theory, the permission request features (permission explanation, third-party authentication and permission relevance) are explored to influence user authorization behaviour by reducing the user's perceived uncertainty, and a research model, as in Figure 1, is established and examined through scenario-based experimental methods. In this case, permission explanation refers to explaining the purpose of information collection and how it is used to the user when asking for authorization. Permission relevance refers to the degree of relevance of the requested permission to the core functionality. Third-party authentication refers to authentication by an authority independent of the developer.

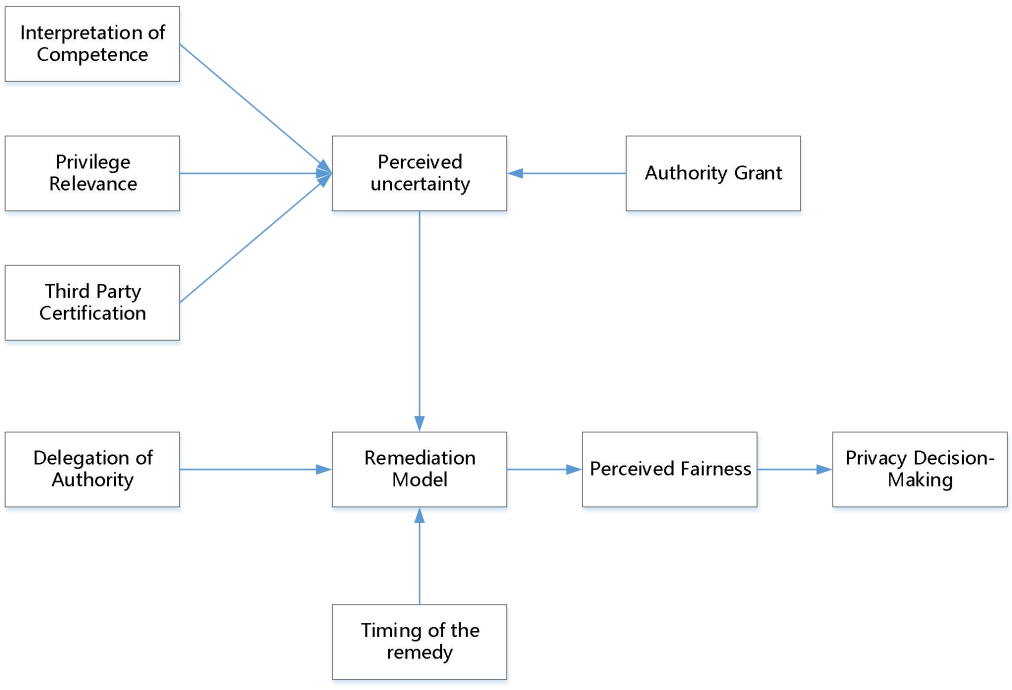


Figure 1 Model of the Effect of Authorization Interface Request Characteristics on Group Privacy Intentions

3 THE CONVERGENCE OF COMPUTATIONAL BIG MODELS AND DIGITAL MEDIA

In today's digital era, the rise of computational big models has brought profound changes to digital media. With its powerful data analysis and processing capabilities, Computational Big Model is able to quickly and accurately process large amounts of data such as text, images, audio and video, providing digital media with more accurate content recommendations, personalized services, and automatic generation of news, articles and videos. The emergence of computational big models not only improves the production efficiency and quality of digital media, but also provides users with a richer and more personalized media experience.

3.1 Application Areas of Computational Big Model in Digital Media

The application areas of computational big models in digital media include four sections such as content recommendation and personification, automatic generation of news, articles and videos, speech recognition and image processing, and sentiment analysis, as shown below.

3.1.1 Content recommendation and personalized services

Computing big models can provide personalized content recommendation and services for users by analyzing data such as users' historical behaviour, interests and social relationships. For example, a news client can recommend news articles of interest to users based on their reading history and interest preferences; a video platform can recommend personalized video content to users based on their viewing history and preferences.

3.1.2 Automatic generation of news, articles and videos.

Computational big models can automatically generate content such as news, articles and videos by learning from large amounts of text, image and video data. For example, some news organizations have started to use computational big models to automatically generate news stories, improving the efficiency and quality of news production.

3.1.3 Speech recognition and image processing

Computational Big Models can achieve functions such as speech recognition and image processing by learning a large amount of speech and image data. For example, voice assistants can provide users with voice interaction services through voice recognition technology; image recognition software can recognition information such as objects and people in images through image processing technology.

3.1.4 Sentiment analysis

Computing big models can achieve functions such as sentiment analysis by analyzing sentiment information in data such as text, voice and images. For example, social media platforms can use sentiment analysis technology to understand users' emotional inclination towards an event or topic, and provide decision-making references for enterprises and governments.

3.2 The Opportunities and Challenges of Computing Big Models for Digital Media

Through deep learning technology, the computational model can quickly generate high-quality content, such as news reports, articles, video scripts, etc., which greatly improves the production efficiency of media content. In terms of creative design, large models can generate illustrations, posters, advertisements and other visual elements, providing a rich source of material for the media industry.

3.2.1 Enhancing efficiency and innovation

The emergence of computational big models has brought unprecedented opportunities for digital media. Computational big models can process large amounts of data quickly and accurately, improving the production efficiency and quality of digital media. At the same time, computational big models can also provide digital media with more innovative services and functions, such as personalized recommendations and automatic content generation, to meet the changing needs of users.

3.2.2 Data security and privacy protection challenges

With the application of computational big models, digital media are facing increasingly severe data security and privacy protection challenges. Computational big models require a large amount of data for training and optimization, which may contain users' personal information and privacy. If these data are leaked or misused, it will have a serious impact on users' rights and interests and the stability of the society.

4 THE ANALYSIS OF THE BEHAVIOR AND PRIVACY OF THE DIGITAL MEDIA

Large models need to deal with a large amount of user data in the process of training and using, so how to protect user privacy and prevent data leakage has become an important issue.

4.1 The Definition of Digital Media Problematic Usage Behavior Based on Computational Big Model

4.1.1 Over-reliance on big models to generate content

Some digital media practitioners and users over-rely on computing big models to generate content, and lack of content review and control, resulting in a decline in content quality and damage to information authenticity.

4.1.2 Misuse of big models to spread false information

Some bad elements abuse big computing models to spread false information, mislead the public and undermine social stability.

4.2 Specific Manifestations of Problematic Use Behavior

1) The quality of content is reduced and the authenticity of information is impaired. The content generated by over-reliance on large models often lacks depth and thinking, and the quality is not high. At the same time, because the training data of large models may be biased, the generated content may have the problem of information authenticity.
2) Users' blind trust in big models. Some users blindly trust the content generated by large models, lack the ability to distinguish information, and are easily misled by false information.
3) The impact on traditional media creation. The emergence of large computing models has had a certain impact on the creation of traditional media. Some traditional media practitioners may lose their jobs because of the competition of big models, and at the same time, the creative methods and values of traditional media may also be challenged.

4.3 The Potential Threat of Big Models to Digital Media Privacy

4.3.1 Increased risk of data leakage

Computing large models requires a large amount of data for training and optimization, which may contain personal information and privacy of users. If these data are leaked, it will have a serious impact on the rights and interests of users.

4.3.2 Potential for misuse of personal information

Computing big model can understand the user's interests, habits and other information by analyzing the user's data. If this information is abused, it will violate the privacy of users.

5 CONCLUSION

With the popularity and development of digital media, users' personal information and privacy are at increasing risk. At the same time, the problematic use behaviors of digital media have become increasingly serious, such as over-reliance on big models to generate content, and misuse of big models for false information dissemination, etc. These behaviors not only affect the quality of content and authenticity of information in digital media, but also pose a threat to the rights and interests of users and the stability of the society. Therefore, it is of great practical significance and urgency to study the privacy protection mechanism and problematic usage behaviors of digital media under computational big models. The privacy protection mechanism of digital media based on computational big model is a complex and important topic. The level of privacy protection in digital media can be effectively enhanced through the innovative application of encryption technology, the construction of intelligent defence systems, and the formulation and implementation of privacy protection policies. However, the existence of problematic usage behaviors still poses challenges to privacy protection. In the future, it is necessary to further strengthen technological innovation and policy regulation, as well as to raise public awareness of privacy protection, so as to jointly build a safe and trustworthy digital media environment.

COMPETING INTERESTS

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

FUNDING

This work is partially supported by 2023 Shenzhen Education Science Planning Project "Research on the Mechanisms and Strategies of Cross-border Co-operation in the Industry-Teaching Integration Community of Vocational Undergraduate Colleges and Universities (yb23048)", 2024 Guangdong Province Education Science Planning Project (Higher Education Special Project) "Research on Smart Classroom Teaching Behavior Analysis Methods Based on Scene Semantic Understanding and Deep Learning Characteristic Representation (2024GXJK766)", 2023 Guangdong Provincial Higher Vocational Education Teaching Reform Research and Practice Project "Research on Online Teaching Quality Evaluation Method Based on Multimodal Affective State Analysis (2023JG277)", 2024 Shenzhen Polytechnic University Quality Engineering Project "Research on Classroom Scene Understanding and Behavior Analysis Method Based on Multimodal Attention Mechanisms (7024310268)", 2024 Higher Education Scientific Research Planning Project of the Chinese Society of Higher Education "Research on the Analysis of Teaching and Learning Deep Interaction Characteristics in Smart Classroom Environment Supported by Multimodal Data (24XH0407)".

REFERENCES

[1] Spiekermann S. Viewpoint The Challenges of Privacy by Design.Communications of the ACM, 2012, 55(7): 38-40.

[2] Bélanger F, Crossler E R,Hiller S J, et al. POCKET: A tool for protecting children's privacy online. Decision Support Systems, 2013, 54(2): 1161-1173.

[3] Berendt B, Günther O, Spiekermann S. Privacy in e-commerce. Communications of the ACM, 2005, 48(4): 101-106.

[4] Degirmenci K. Mobile users' information privacy concerns and the role of app permission requests. International Journal of Information Management, 2020, 50, 261-272.

[5] Spiekermann S. The challenges of privacy by design. Communications of the ACM, 2012, 55(7): 38-40.

[6] Spiekermann S, Acquisti A, Böhme R, et al. The challenges of personal data markets and privacy. Electronic Markets, 2015, 25(2): 161-167.

[7] Acquisti A, Brandimarte L, Loewenstein G. Privacy and human behavior in the age of information. Science, 2015, 347(6221): 509-514.

[8] Neves J, Turel O, Oliveira T. Privacy concerns in social media use: A fear appeal intervention. International Journal of Information Management Data Insights, 2024, 4(2): 100260-100260.

[9] S. J H , Taylor J P, Kara B, et al. Digital Media and Developing Brains: Concerns and Opportunities. Current Addiction Reports, 2024, 11(2): 287-298.

[10] Ofir T, Hamed S Q, Isaac V. Special Issue: Dark Sides of Digitalization. International Journal of Electronic Commerce, 2021, 25(2): 127-135.

[11] Turel, Qahri-Saremi. Problematic Use of Social Networking Sites: Antecedents and Consequence from a Dual-System Theory Perspective. Journal of Management Information Systems, 2016, 33(4): 1087-1116.

[12] M S, M A M, A F, et al. Impact of a targeted direct marketing price promotion intervention (Buywell) on food-purchasing behaviour by low income consumers: a randomised controlled trial. Journal of human nutrition and dietetics: the official journal of the British Dietetic Association, 2017, 30(4): 524-533.

[13] Norman V, Ahlqvist M, Mattsson T. Evaluation of scale invariance in fatigue crack growth in metallic materials. International Journal of Fatigue, 2024, 189, 108545-108545.

[14] D'Ambrosio S, De Pasquale S, Iannone G, et al. Privacy as a proxy for Green Web browsing: Methodology and experimentation. Computer Networks, 2017, 126: 81-99.

[15] Ohme J, Araujo T, Boeschoten L, et al. Digital trace data collection for social media effects research: APIs, data donation, and (screen) tracking. Communication Methods and Measures, 2024, 18(2): 124-141.

[16] Dwivedi Y K, Ismagilova E, Hughes D L, et al. Setting the future of digital and social media marketing research: Perspectives and research propositions. International journal of information management, 2021, 59: 102168.

[17] Stoumpos A I, Kitsios F, Talias M A. Digital transformation in healthcare: technology acceptance and its applications. International journal of environmental research and public health, 2023, 20(4): 3407.

[18] Ahammed M F, Labu M R. Privacy-Preserving Data Sharing in Healthcare: Advances in Secure Multiparty Computation. Journal of Medical and Health Studies, 2024, 5(2): 37-47.

[19] Delaney J, Ghazi B, Harrison C, et al. Differentially Private Ad Conversion Measurement. Arxiv preprint arxiv: 2403. 15224, 2024.

[20] Jain P, Gyanchandani M, Khare N. Big data privacy: a technological perspective and review. Journal of Big Data, 2016, 3(1): 1-25.

[21] Imdad U, Roksana B, S S K. Privacy in targeted advertising on mobile devices: a survey. International journal of information security, 2022, 22(3): 31-32.

[22] Aridor G, Che Y K. Privacy Regulation and Targeted Advertising: Evidence from Apple's App Tracking Transparency. 2024.

[23] Borenstein B E, Taylor C R. The effects of targeted digital advertising on consumer welfare. Journal of Strategic Marketing, 2024, 32(3): 317-332.

[24] Ran H. Improved content recommendation algorithm integrating semantic information. Journal of Big Data, 2023, 10(1).

[25] Pan Y, Wang M, Lu L, et al. Scan-to-graph: Automatic generation and representation of highway geometric digital twins from point cloud data. Automation in Construction, 2024, 166, 105654-105654.

[26] J. M R. Signals and Images: Advances and Results in Speech, Estimation, Compression, Recognition, Filtering, and Processing. Photogrammetric Engineering & Remote Sensing, 2020, 86(2): 77-78.

[27] Bartoletti I, Plantié S, Sambodaran A. Security and privacy risks in the blockchain ecosystem. Cyber Security: A Peer-Reviewed Journal, 2019, 3(3): 195-207.

[28] Jia W, Guangbin W, Heng L, et al. Intelligent Construction Activity Identification for All-Weather Site Monitoring Using 4D Millimeter-Wave Technology. Journal of Construction Engineering and Management, 2024, 150(11).

[29] Zhu R, Srivastava A, Sutanto J. Privacy-deprived e-commerce: the efficacy of consumer privacy policies on China's e-commerce websites from a legal perspective. Information Technology & People, 2020, 33(6): 1601-1626.

[30] Watson J, Lacey D, Kerr D, et al. Understanding the effects of compromise and misuse of personal details on older people. Australasian Journal of Information Systems, 2019, 23.
[31] Kani-Zabihi E, Helmhout M. Increasing service users' privacy awareness by introducing on-line interactive privacy features//Information Security Technology for Applications: 16th Nordic Conference on Secure IT Systems, NordSec 2011, Tallinn, Estonia, October 26-28, 2011, Revised Selected Papers 16. Springer Berlin Heidelberg, 2012: 131-148.
[32] Bilal A. Rise of technomoral virtues for artificial intelligence-based emerging technologies' users and producers: threats to personal information privacy, the privacy paradox, trust in emerging technologies, and virtue ethics. 2022.

Journal of Computer Science and Electrical Engineering
ISSN: 2663-1946
DOI: 10.61784/jcsee3019

DESIGN OF A CAT3512B DIESEL GENERATOR DATA ACQUISITION SYSTEM BASED ON STM32

YuanJu Zhou
Chengdu Jiushang Technology Co., Ltd., Chengdu 610000, Sichuan, China.
Corresponding Email: yjzhou.2008@163.com

Abstract: The CAT3512B generator set, renowned for its efficient and stable operation as well as robust power support, has found widespread application across various industries such as engineering machinery, petroleum, power industry, and chemicals. This article presents a data acquisition system designed utilizing embedded technology and IoT applications, centered on the STM32F103C8T6 microcontroller. This system is tasked with real-time acquisition, transmission, and application of operational parameters from both the engine and generator of the CAT3512B generator set. The objective is to facilitate real-time monitoring and early warning, enabling historical data retrieval of operational parameters. Leveraging big data and AI algorithms, it further aids in preventive and precise maintenance, providing management with data-driven, scientifically sound decision support.The data acquisition unit employs the STM32F103C8T6 chip, which strikes an excellent balance between performance and cost. Engine parameters are collected via the CAN bus, while electrical parameters from the generator are gathered using an Rs485 bus interface with an electrical parameter tester. The acquired data is packaged into the corresponding message format of the data platform and transmitted to the data processing center via the NB-IoT communication module for analysis, storage, and application.
Keywords: STM32F103C8T6; CAT3512B; CAN bus; Rs485 bus; NB-IoT

1 INTRODUCTION

The significance of real-time monitoring of equipment operational parameters and early warning information cannot be overstated in modern industrial production and maintenance. It plays a pivotal role in effectively preventing equipment failures, enhancing efficiency and energy consumption ratios, ensuring production safety, leveraging big data and AI algorithms for preventive and precise maintenance, facilitating data visualization and historical trend analysis, and ultimately providing data-driven, scientifically sound decision support for management.
The CAT3512B generator set, renowned for its efficient and stable operation coupled with robust power output, has found widespread application across industries such as engineering machinery, petroleum, power, and chemicals. Its exceptional performance and versatility have made it a highly sought-after product in the market, providing vital energy support for various industries' production and development.
To effectively utilize these generators' operational parameters for monitoring, diagnosis, maintenance, and control, real-time data acquisition and warning information, along with a seamless communication channel between the equipment and the data center, are crucial. This article designs a data acquisition system for CAT 3512B diesel generator sets based on the STM32F103C8T6 microcontroller, which collects engine and electrical parameters to provide data support for real-time monitoring of equipment operation and subsequent data analysis applications.

2 DESIGN PLAN

2.1 Overall Hardware Design Plan

The STM32F103C8T6 is a high-performance, low-power 32-bit microcontroller introduced by STMicroelectronics. Based on the ARM Cortex-M3 core, it features high performance and low power consumption. With a 32-bit bus width and a maximum clock frequency of 72MHz, it provides fast data processing capabilities. It has 64KB of Flash memory for storing program code and 20KB of SRAM for storing runtime data. The operating temperature ranges from -40°C to 85°C, ensuring stable operation in various industrial environments. This chip integrates a variety of on-chip peripherals, including USART, CAN, USB, and ADC, to facilitate data exchange with other peripherals or devices.
The STM32F103C8T6 is widely used in various embedded control and computing fields, including but not limited to: industrial control, automotive electronics[1], smart home, Smart Agriculture [2],and medical equipment.
This article designs a data collector for monitoring real-time parameters during the operation of the CAT3512B generator. Based on the working principle of the CAT3512B generator, the data can be divided into engine operating parameters and generator operating parameters. The data collector adopts an embedded design and is equipped with an Rs485 transceiver (Max3485) and a CAN transceiver (TJA1050). The main functions of the collector include: collection of engine operating parameters and alarm data, collection of generator operating parameters, and data transmission module. The specific design plan is shown in Figure 1:

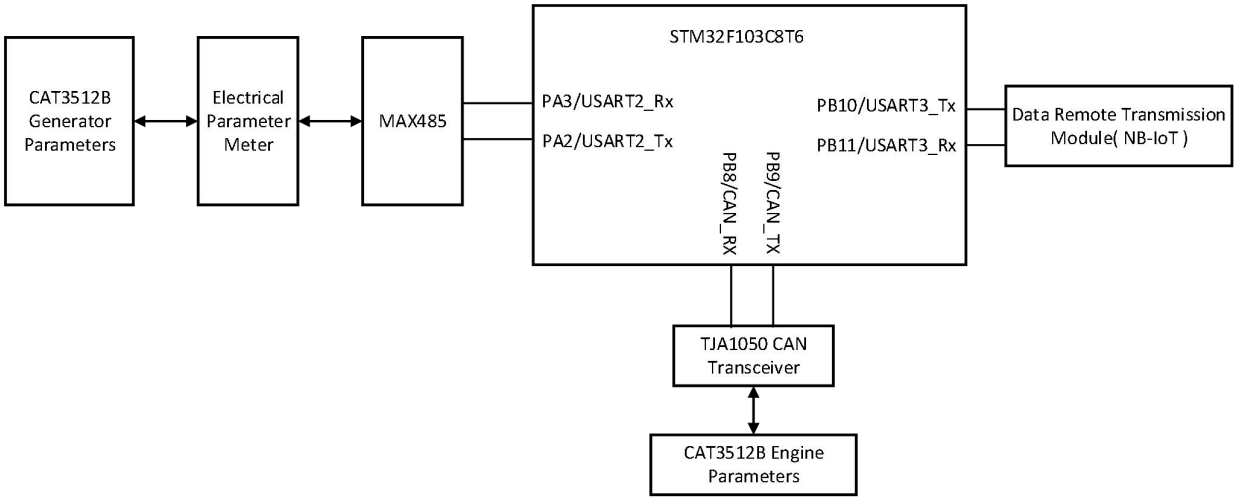


Figure 1 Hardware Design Plan

The engine data primarily encompasses engine load, fuel pressure, oil pressure, fuel level, coolant temperature, and alarm parameters during engine operation. This data is transmitted using the J1939 protocol[3][4], with the parameter message structure conforming to the CAN 2.0B protocol. The STM32F103C8T6 microcontroller, leveraging its on-chip peripherals, accomplishes real-time data acquisition and processing of the engine through the TJA1050 CAN transceiver.

Generator parameters include three-phase voltage, three-phase current, power factor, active power, reactive power, among others. These generator parameters are collected by an electrical parameter meter, and the acquisition unit retrieves these parameters from the meter via an RS485 communication port, utilizing the Modbus-RTU communication protocol.

The data remote transmission module mainly uses the NB-IoT communication module to report the real-time collected engine and generator data to the data center.

2.2 Overall Software Design Plan

The data collector uses a table lookup method for data caching, which effectively improves the execution efficiency of the microcontroller, reduces the consumption of computational resources, facilitates post-upgrade maintenance and expansion of the software, and to some extent, simplifies multi-business interaction. In terms of business design, a combination of interrupts and timer polling is employed. Engine data is received via CAN interrupts, and after comparing with the PGN through table lookup, it is cached as shown in the table; generator data is periodically polled from the electrical parameter instrument using the Modbus protocol, verified, and then stored in the generator parameter table; the data transmission module regularly queries the parameter table and reports data according to the protocol. The software process is illustrated in Figure 2:

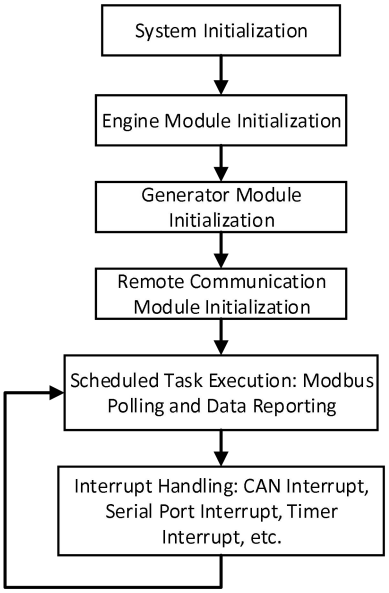


Figure 2 Software Design Process

3 ENGINE MODULE

3.1 Engine Parameter

The CAT 3512B engine, during its operation, transmits message parameters that adhere to the SAE-J1939 protocol. SAE-J1939 is a recommended standard established by the Society of Automotive Engineers (SAE) specifically designed for digital communication among electronic components in medium- and heavy-duty on-road vehicles as well as off-road vehicles (including construction machinery, agricultural machinery, rail vehicles, ships, etc.). Built upon the Controller Area Network (CAN) technology developed by Bosch in Germany, SAE-J1939 serves as the High-Level Protocol (HLP) for CAN, providing a standardized architecture to enable communication between electronic control units (ECUs) from different manufacturers. The parameters of the CAT 3512B engine during operation are listed in Table 1, with the following common parameter:

3.1.1 Engine Status Parameters

RPM (Revolutions Per Minute): Indicates the current rotational speed of the engine, crucial for monitoring engine performance.

Fuel Consumption Rate: Measures the amount of fuel consumed by the engine per unit time, essential for assessing engine efficiency and fuel management.

Coolant Temperature: Reflects the status of the engine's cooling system, vital for preventing engine overheating.

Oil Pressure and Temperature: Key parameters ensuring the proper functioning of the engine's lubrication system.

3.1.2 Performance Parameters

Torque Output: Indicates the amount of torque produced by the engine under current operating conditions, significant for evaluating engine power performance.

Power Output: The actual power output of the engine, typically measured in kilowatts (kW) or horsepower (hp).

Fuel Efficiency: Represents the engine's ability to convert fuel into mechanical energy, impacting both economy and environmental friendliness.

3.1.3 Fault Diagnosis and Warning Parameters

Fault Codes: Generated when the engine or related systems encounter faults, transmitted via the CAN bus to the monitoring system for diagnosis.

Sensor Status: Includes readings and status information from various sensors (e.g., temperature sensors, pressure sensors, speed sensors) to monitor engine operation.

Warning Information: Alerts operators or maintenance personnel when the engine or related systems are approaching critical conditions.

3.1.4 Control Parameters

Fuel Injection Control: Involves parameters such as fuel injection quantity and timing, crucial for precise control of the engine's combustion process.

Intake Control: Encompasses parameters like turbocharger control and intake air volume adjustment, optimizing the engine's intake efficiency.

Emission Control: Includes parameters for Exhaust Gas Recirculation (EGR) control, aftertreatment system control, and others, ensuring compliance with emission regulations.

Table 1 Engine Partial Parameter Information Table

PGN (Parameter Group Number)	Parameter Name	Data Range
0xF003	Engine Percent Load At Current Speed	0 to 250 %
0xF004	Engine Speed	0 to 8,031.875 rpm
0xFEDF	Engine's Desired Operating Speed	0 to 8,031.875 rpm
0xFEE5	Engine Total Hours of Operation	0 to 210,554,060.75 hr
0xFEE9	Engine Total Fuel Used	0 to 2,105,540,607.5 L
0xFEEE	Engine Coolant Temperature	-40 to 210 deg C
0xFEEF	Engine Oil Pressure	0 to 1000 kPa
0xFEF2	Engine Fuel Rate	0 to 3,212.75 L/h
0xFEF5	Barometric Pressure	0 to 125 kPa
0xFEF6	Engine Turbocharger Boost Pressure	0 to 500 kPa
0xFEF7	Battery Potential (Voltage), Switched	0 to 3212.75 V
0xFE06	Engine Exhaust GAS Temperature-Right Manifold	-273 to 1735 deg C
0xFE92	Unfiltered Engine Oil Pressure	0 to 1000KPa

3.2 Analysis of Engine Alarm Information

SAE J1939, at the application layer, defines various diagnostic messages (DM) and diagnostic trouble codes (DTC) for fault diagnosis and monitoring. Here are the common diagnostic information types within SAE J1939:

(1)Diagnostic Message 1 (DM1) – Active Diagnostic Trouble Codes: This message is used to transmit currently active fault codes that indicate issues with the vehicle's systems. Users typically utilize DM1 to read these active fault codes. DM1 corresponds to a Parameter Group Number (PGN) of 65226 or hexadecimal FECA.
(2)Diagnostic Message 2 (DM2) – Inactive/Historically Active Fault Codes: This message type carries fault codes that were once active but are no longer current, providing a historical record of faults.
(3)Diagnostic Message 3 (DM3) – Clear Inactive Fault Codes: This message is used to request the clearing of inactive fault codes from the system's memory.
The diagnostic trouble codes (DTCs) themselves are structured into three main components, Each DTC is composed of 4 bytes, with the bit allocation for the three components typically structured shown in Table 2:
(1)Suspect Parameter Number (SPN): This identifies the specific parameter or system associated with the fault.
(2)Failure Mode Identifier (FMI): This indicates the specific type of failure or issue related to the parameter identified by the SPN.
(3)Occurrence Count (OC): This counts the number of times the fault has occurred or been detected.

Table 2 DTC Representation in CAN Data Frame for DM1 (Byte 3 Closer to CAN Identifier)

DTC																															
Byte3 8 least significant bits of SPN (bit 8 most significant)								Byte4 send byte of SPN (bit 8 most significant)								Byte5 3 most significant bits of SPN and the FMI (bit 8 SPN msb and bit 5 FMI msb)								Byte6							
SPN																		FMI				CM	OC								
8	7	6	5	4	3	2	1	8	7	6	5	4	3	2	1	8	7	6	5	4	3	2	1	8	7	6	5	4	3	2	1

3.3 Hardware Design for Engine Parameter Acquisition

In this article, the bxCAN (Basic Extended CAN) controller of the STM32F103C8T6 is utilized to establish a connection with the CAN bus of the CAT3512B engine. The CAN transceiver TJA1050 converts the signals from the CAN controller into corresponding differential signals for the CAN bus, enabling communication with other CAN devices. The key features of the bxCAN controller include:
(1)Compatibility with CAN Protocols: The bxCAN supports both CAN protocol 2.0A and 2.0B in active mode, allowing it to be compatible with different versions of the CAN standard.
(2)Multi-level Depth FIFO: It features two receive FIFOs, each with a depth of three levels. This enables the storage of multiple complete messages, which are automatically managed by hardware, thereby reducing the burden on the CPU. Variable Filter Banks: It provides 14 filter banks (Standard CAN) or 28 filter banks (Extended CAN) with variable bit widths. This allows the software to configure the filter banks to select only the required messages while discarding others, enhancing communication efficiency.
(3)Receive Timestamp: It records the timestamp of the Start of Frame (SOF) moment, facilitating synchronization and analysis of the received data.
The TJA1050, as a high-performance CAN bus transceiver, is widely used in IoT (Internet of Things), industrial electronics, and automotive electronics applications. Its key attributes include:
(1)Wide Temperature Range: The TJA1050 supports a broad temperature range, making it suitable for industrial control and automotive applications.
(2)Strong EMI Resistance: Designed to minimize surge interference caused by electromagnetic radiation, it possesses robust electromagnetic interference (EMI) resistance, capable of withstanding electromagnetic noise from the external environment, including surges that may be induced.
(3)Protection Against Surge Damage: The TJA1050 integrates protection components such as transient voltage suppression (TVS) diodes, effectively safeguarding the chip from surge damage.
This combination of the STM32F103C8T6's bxCAN controller and the TJA1050 CAN transceiver enables reliable and efficient communication with the CAT3512B engine's CAN bus.

3.4 Software Design for Engine Parameter Acquisition

Engine parameter tables are divided into Engine Parameter Group Number (PGN) parameter tables and alarm information parameter tables. The PGN parameter table primarily includes parameters such as PNG, parameter location, parameter length, resolution, etc. It is used for looking up and comparing parameters, as well as caching parameters reported by the engine. The alarm parameter table, on the other hand, contains SPN (Suspect Parameter Number) and FMI (Failure Mode Identifier), which are utilized to check and compare whether alarms have occurred, and to cache alarm information.
Initializing the CAN peripheral involves several steps, including GPIO multiplexing, enabling the CAN clock, setting the CAN bus baud rate, and enabling CAN interrupts. To enhance CAN communication efficiency, CAN filter parameters can be calculated based on the PGN parameter table. The filter settings can be configured using a

combination of list mode and mask mode, allowing for more flexible and efficient filtering of CAN messages. The specific software process is illustrated in Figure 3:

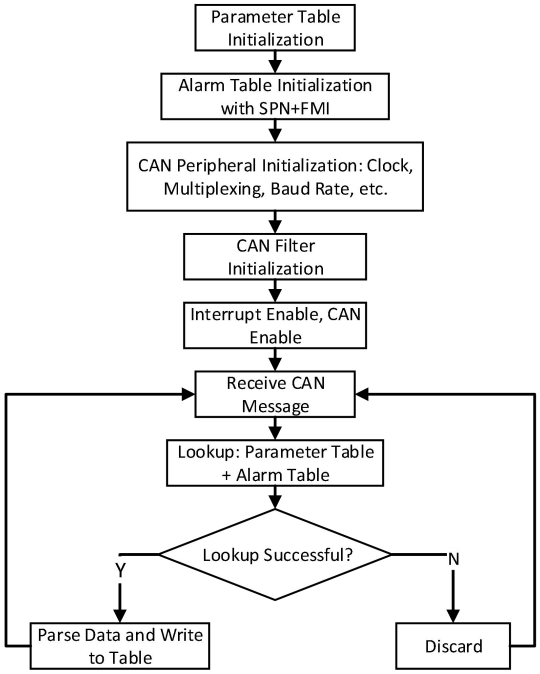


Figure 3 Engine Parameter Software Process Design

4 GENERATOR MODULE

The CAT3512B generator boasts exceptional power generation capabilities, with a rated power range between 1230 and 1500 kW, depending on its configuration and operating frequency. Its stable voltage and frequency output ensure the reliability of power quality, safeguarding the proper functioning of electrical equipment. Monitoring the generator's parameters is crucial for maintaining its safety, efficiency, and stable operation. It also facilitates the timely detection and resolution of potential issues, thereby safeguarding the reliability and continuity of power supply.

Based on the generator's real-time load conditions, its output power can be adjusted to ensure optimal operation within the desired load range, enhancing energy utilization efficiency. By optimizing the generator's operational parameters, such as fuel injection and rotational speed, it is possible to reduce fuel consumption and emissions, achieving energy-saving and emission-reduction goals.

This design incorporates an electrical parameter tester to collect various electrical parameters from the generator in real-time, including voltage, current, frequency, impedance, and power factor. The data collector retrieves the generator's parameters through an RS485 communication interface using the Modbus protocol, providing a reliable and efficient means of monitoring and managing the generator's performance.

4.1 Electrical Parameter Meter

An electrical parameter meter is an instrument used to measure and monitor various electrical parameters in circuits. It accurately measures parameters such as voltage, current, impedance, power factor, among others, assisting engineers and technicians in evaluating and optimizing circuit performance. Employing precise measurement technology, the electrical parameter tester boasts high resolution and low measurement error, catering to applications requiring high precision.

The electrical parameter meter covers a wide range of measurements, including voltage, current, frequency, impedance, power factor, and supports various measurement modes like real-time measurement, peak measurement, and root mean square (RMS) measurement to meet diverse application requirements.

The generator parameters are outlined in the following table 3:

Table 3 Generator Modbus Parameter Table

Parameter Name	Modbus Address	Data Length	Resolution	Description
Voltage (AC RMS)	40001	2 bytes	0.1 Volt	Root Mean Square AC Voltage
Current (AC RMS)	40003	2 bytes	0.1 Ampere	Root Mean Square AC Current
Frequency	40005	2 bytes	0.01 Hz	AC Frequency

Power Factor	40007	2 bytes	0.001	Power Factor
Impedance (Real Part)	40009	2 bytes	0.1 Ohm	Real Part of Impedance
Impedance (Imaginary)	40011	2 bytes	0.1 Ohm	Imaginary Part of Impedance
Active Power	40013	4 bytes	1 Watt	Active Power (Real Power)
Reactive Power	40017	4 bytes	1 VAR	Reactive Power

RS485 bus is a widely used serial communication protocol standard that employs differential signal transmission to effectively suppress common-mode interference, enhancing data transmission stability and reliability. Even in complex electromagnetic environments or high-noise scenarios, the RS485 bus maintains excellent communication quality. When using twisted pair cables as the transmission medium, the RS485 bus can achieve transmission distances of up to several kilometers, with a maximum distance of 15km at a transmission rate of 9600bps. The RS485 bus flexibly supports up to 32 transceivers, facilitating the establishment of device networks where bus control rights can be allocated through master polling or token passing, meeting communication needs in various scenarios and complex networks.

The Modbus protocol is a widely adopted communication protocol in the industrial control field. The communication network typically adopts a master-slave structure, where the master polls multiple slave devices to obtain their data. The Modbus protocol exhibits excellent fault tolerance and reliability, ensuring stable operation in complex industrial environments and transmission reliability.

4.2 Software Design for Generator Parameter Acquisition

A generator parameter table typically includes details such as the Modbus address, data length, and resolution for each parameter. These parameters are periodically acquired from the electrical parameter tester (EPT) via the RS485 interface using the Modbus protocol through a polling mechanism. The response message, once validated through data checks, is then written into the generator parameter table. The specific software process is illustrated in Figure 4:

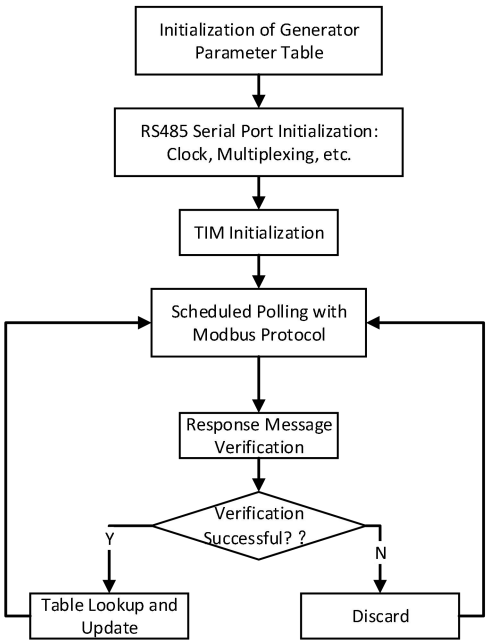


Figure 4 Generator Parameter Software Flowchart

5 DATA TRANSMISSION MODULE

NB-IoT, as a low-power, narrowband wireless communication technology, boasts broad application prospects and significant technical value in the Internet of Things (IoT) field. NB-IoT devices can achieve long-range transmission and large-scale connectivity with low power consumption[5], offering higher coverage gain than traditional GSM networks. Moreover, both the module cost and communication operating cost of NB-IoT are relatively low, facilitating the popularization and promotion of IoT. Employing multi-layer encryption mechanisms, NB-IoT ensures the security and privacy protection of data during transmission.

NB-IoT is widely applied in various sectors, including but not limited to logistics and warehousing, consumer devices, smart manufacturing, public utilities, healthcare, smart cities, and agricultural environments[6][7][8].

When integrating NB-IoT with an STM32F103 microcontroller, the two communicate via a serial port. This serial communication is used to initialize the communication parameters of the NB-IoT module, establish device connections, and connect the device to a data center or cloud platform. During operation, the device's sensors or collectors

periodically report engine and other relevant parameters to the system through the NB-IoT connection, ensuring timely and reliable data transmission. The software process for data remote transmission is illustrated in Figure 5:

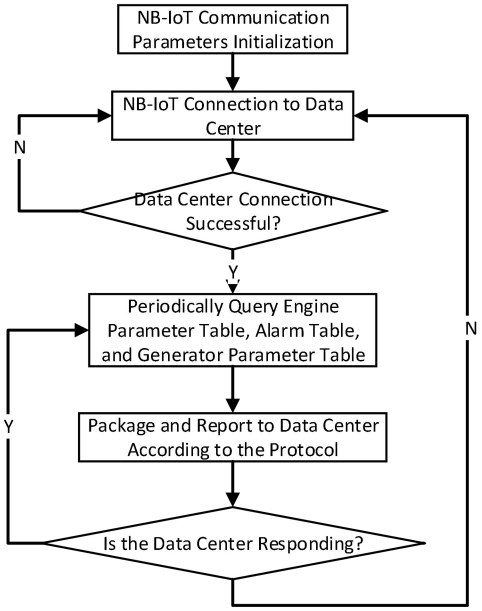


Figure 5 Data Remote Transmission Software Flowchart

6 CONCLUSION

This paper presents a design scheme for a CAT3512B diesel generator set data acquisition unit based on STM32, which realizes remote real-time monitoring of diesel generator equipment and provides data support for subsequent preventive maintenance. This scheme boasts low cost, high reliability, easy expandability, and low power consumption, making it widely applicable for similar equipment to achieve data collection, processing, and transmission. In the future, we will further optimize the performance and functionality of this device. By enabling real-time monitoring of operational parameters and warning information, it will not only prevent equipment failures but also play a vital role in enhancing equipment efficiency and safety, reducing maintenance costs, supporting remote operation and maintenance, and providing data visualization and decision support. These features collectively constitute an essential component of modern industrial production and equipment maintenance.

COMPETING INTERESTS

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

REFERENCE

[1] Hongwei Guo, Tuo Yang. Design and implementation of an intelligent car obstacle avoidance system based on deep learning. Electronics Science Technology and Application, 2023, 10(2).
[2] Xuhai Wang, Weiguo Li, Lili Wang, et al. Based on STM32F103 cowshed environment intelligent control system. IOP Conference Series: Materials Science and Engineering, 2020, 782(5).
[3] Alzahrani Ahmad, Shriya Makarand Wangikar, Vairavasundaram Indragandhi, et al. Design and Implementation of SAE J1939 and Modbus Communication Protocols for Electric Vehicle. Machines, 2023, 11(2) : 201-201.
[4] Xuan Shao, Xingwu Kang, Xuping Wang et al. Design of special vehicle condition monitoring system based on J1939. Journal of Physics: Conference Series, 2020, 1549(3) : 032092.
[5] Jun Wang, Ting Ke, Mengjie Hou, et al. The Design of Home Fire Monitoring System based on NB-IoT. International Journal of Advanced Computer Science and Applications (IJACSA), 2022, 13(5).
[6] Ernesto Sanz, Jorge Trincado, Jorge Martínez, et al. Cloud-based system for monitoring event-based hydrological processes based on dense sensor network and NB-IoT connectivity. Environmental Modelling and Software, 2024, 182: 106186-106186.
[7] Lin Qingyao, Gao Qiuhong. Electric bicycle battery management system based on NB-IoT. 2024.
[8] Jun Lin, Xiaobin Xu. Design of a textile storage environment fire detection system based on ZigBee and NB-IoT. Journal of Physics: Conference Series, 2024, 2797(1).

Journal of Computer Science and Electrical Engineering
ISSN: 2663-1946
DOI: 10.61784/jcsee3020

MLMM : MULTI-MODAL LANGUAGE MODELS BASED MULTI-AGENTS COLLABORATED SOLVING CV PROBLEMS

YanQiao Ji
Liaoning Equipment Manufacturing Vocational and Technical College, Shenyang 110161, Liaoning, China.
Corresponding Email: 491318458@qq.com

Abstract: To enhance the system's ability to interact with and comprehend the visual world, we integrate vision capabilities into a large language model (LLM) framework, specifically the AutoGen framework. By employing a multi-agent conversation methodology, our work aims to mitigate errors in image generation and improve the system's output to better meet user expectations. Our work is based on the Meta Llama 3 family of pretrained and instruction-tuned generative text models, specifically optimized for dialogue applications, LLaVA for image recognization and the stable diffusion model for image generation. Our work is efforting on addressing vision-related problems and the potential for further enhancements with the support of more sophisticated models.
Keywords: Software engineering; Artificial intelligence; Deep learning

1 INTRODUCTION

The journal of Upubscience Publisher gives preference to manuscripts of high scientific level, which have not been published, and are written not only for specialists but also for the general public interested in the questions of related fields. A large language model (LLM) is a type of artificial intelligence model that uses machine learning techniques to analyze and generate human-like text. It is trained on a massive amount of text data, allowing it to learn language structure, grammar, and vocabulary, and can be used for various natural language processing tasks such as language translation, text summarizing, and question answering. As the rapid development of LLMs, it' s becoming a powerful back-end of developing intelligence agent technology because of the ability of planning, memorizing, tool using, and adaptation to novel observations in a multitude of real-world tasks[1, 2]. Vision plays an important role in recognizing and interacting the worlds for humans and many other animals. A central aspiration of Artificial Intelligence (AI) is to devise AI agents capable of emulating the efficient perception and generation of visual signals, thereby enabling the inference about and interaction with the visual world. Examples of this include the recognition of objects and actions within scenes, as well as the creation of sketches and images for communicative purposes. Establishing foundation models endowed with visual capabilities represents a thriving research area in the pursuit of this goal.
A enduring ambition within the field of artificial intelligence is the creation of versatile assistants capable of adhering to users' (multimodal) directives to accomplish an extensive array of real- world duties[3]. Recently, there has been a surge in the community's focus on the development of foundational models that exhibit spontaneous capabilities for multimodal comprehension and production within open-world settings [4]. Whereas the methodologies for deploying Large Language Models (LLMs), such as GPTs[5], to engineer universal assistants for linguistic tasks have been validated, the approaches for constructing all-purpose, multi-modal aides for computer vision and vision-language assignments are yet to be fully charted.
In our work, we mainly made the following contributions:

■We integrate vision capabilities into our system, enabling it to interact with and comprehend the visual world.

■Our work based on a multi-agent conversation methodology to mitigate errors in image generation and improve the system's ability to deliver images that better meet the user's expectations.

■ We implement an iterative exchange between the image-generating agent and the evaluator to refine the image generation process, which enhance the alignment between the system's output and the user's specific requirements through the collaborative dialogue between the agents.

2 RELATED WORKS

2.1 LLM based Multi-Agents Communication Framework

In this part, we explore the application of a proposed multi-agent framework to enhance the functional- ity of existing AI models. We focus on several prominent multi-agent collaboration frameworks[6, 7] and assess how these models operate and their limitations.
Auto-GPT[6], a milestone in AGI development, demonstrates autonomous capabilities through its integration of internet access, memory management, and advanced text generation. It can execute a wide range of tasks with minimal user input. Our framework can be mapped onto Auto-GPT's architecture, treating its main agent as a node in a multi-agent system. This agent interacts with various plugins, which can be seen as external tools or services, and can create and

manage other agents. The introduction of a "Supervisor Agent" in our framework could address issues like looping, while the concept of co-agents could enable collaboration among multiple Auto-GPT instances.
BabyAGI[8] operates on a triad of LLM chains for task generation, prioritization, and execution. Our framework can model BabyAGI as a network of specialized agents, enhancing its structure and modularity. The inclusion of a feedback loop could facilitate learning and improvement over time. Gorilla[7]is a fine-tuned LLaMA model with advanced document retrieval and API interaction capabilities, represents a significant advancement in language modelling. It can be represented within our framework by a single agent with plugins for handling APIs, allowing for flexibility and adaptability to changes in API documentation and functionality. Camel[9] proposed a novel communicative agent framework in which agents communicate with each other in a role-playing method. This framework leverages inception prompting as a means to steer chat agents towards the fulfillment of tasks while ensuring alignment with human objectives. They introduced an innovative communicative agent framework and facilitate further research in the domain of communicative agents and beyond.
Microsoft developed AutoGen[10] in the year 2023, which is an open-source architecture that em- powers developers to create LLM-driven applications through interconnected agents capable of dialogue to achieve objectives. These agents are adaptable and capable of interaction, functioning across different modes that integrate LLMs, human participation, and various tools. Developers can also wield the flexibility to stipulate the interaction protocols for agents using AutoGen. The framework accommodates both natural language and programming code to establish versatile com- munication protocols tailored for various applications. Acting as a universal platform, AutoGen facilitates the construction of a wide spectrum of applications, differing in complexity and LLM capabilities. Experimental research validates the framework's efficacy across numerous exemplar applications, spanning fields such as mathematics, programming, Q&A systems, operations research, digital decision-making, and entertainment.

2.2 Vision-Language Modeling

General vision-language modeling. Building on successes in large language and vision models, recent years have seen a growing interest in large vision-language models (VLMs)[11, 12]. VLMs are adapt at simultaneously understanding both visual and textual content. This multidimensional understanding has enabled VLMs to be effectively applied to a diverse array of tasks, including visual question answering [13], image captioning [14], optical character recognition [15], and object detection[16].
The integration of images into these models takes various forms. For instance, [17]enhance pre- trained language models with a mechanism that allows them to attend directly to a single context image. This approach enables the model to focus on specific visual information relevant to the language task at hand. Frozen[18] take another notable method, in which the vision encoder parameters are optimized through back-propagation, while the language model component remains frozen. This method allows for the fine-tuning of the visual processing capabilities without altering the linguistic knowledge already encoded in the language model. This separation of concerns ensures that the model's proficiency in understanding and generating language is preserved while its ability to interpret visual content is enhanced.
Our work is a demo based on AutoGen Framework and composed with large multi-modal model LLaVa[19] and image generation model stable diffusion. LLava stands for Large Language and Vision Assistant, an end-to-end trained, comprehensive multi-modal model that integrates a vision encoder with an LLM to achieve broad-spectrum understanding of visual and linguistic content. In order to promote ongoing research into the adherence to visual instructions, we have developed two evaluation benchmarks encompassing a variety of demanding and applied tasks.The Stable Diffusion model is a deep learning model used for generating images from text, developed by Stability AI in collaboration with other research institutions. It is a type of diffusion model, which is a class of deep learning models that aim to produce data with a specific structure from random noise. In the context of image generation, diffusion models work by starting with a corruption of the input data and learning to de-noise it over time, eventually producing a clean, structured output.

3 APPLICATION

In this section, we will introduce our work which is primarily built upon the foundational framework of AutoGen. Our system integrates vision capabilities, enabling it to interact with and comprehend the visual world. Due to time constraints, we will present a straightforward sequence to elucidate the comprehensive workflow of our research.
We assume a circumstance in which user need to generate a picture according to a certain theme. In the conventional approach, one might directly employ an image-generation model, which may yield sub optimal results, such as an output that does not adequately fulfill the user's vision. Our emphasis lies in leveraging a multi-agent conversation methodology to mitigate these types of errors.
Specifically, one agent is responsible for generating images, while another agent assumes the role of evaluator, who assesses and scores the produced images, as well as the prompt words employed, ensuring that the output meets the desired standard. Upon encountering discrepancies, the evaluator provides feedback to the image-generating agent, prompting it to make necessary adjustments. This iterative exchange continues until the evaluator deems the image and prompt words to be satisfactory, at which point the process concludes.

The efficacy of this communication method lies in its ability to iterative refine the image generation process, with the goal of enhancing the alignment between the system's output and the user's specific requirements. By fostering a collaborative dialogue between the agents, the system is better equipped to deliver images that better meet the user's expectations.

3.1 Models Definition

As the back-end LLM of AutoGen, we use local-deployed LLaMA3-70b model using for conversation generation and summarization. Developed by Meta-AI, Meta Llama 3 family is a collection of pre- trained and instruction tuned generative text models. The Llama 3 instruction-tuned models have been specifically optimized for dialogue applications and have demonstrated superior performance against many open-source chat models in standard industry benchmarks. Moreover, in the development of these models, the optimization of helpfulness and safety has been prioritized to ensure their effective and responsible use. In the autoGen configuration file, setting the baseurl to the local host of ollama as shown in Algorithm 1.

```
[{
    "model": "llama3:70b",
    "api_key": "EMPTY",
    "tags": ["ollama"],
    "base_url": "http://localhost:11434/v1"
}]
```

Algorithm 1 Configure Setting of Ollama

In the task, there are two vision models will be used. One is responsible for reviewing and evaluating the picture. We deploy LLaVa-7b model locally using ollama framework. The other is focusing on generating an image according the message given by user. We employ the mature and dependable stable-diffusion-v1-5[20]model to transform text into imagery. This model is locally deployed through the diffuse library in Python, and it is selectively invoked by agents within the AutoGen framework.

3.2 Agent Creation

In AutoGen Framework, the core design principle is to optimize and integrate multi-agent workflows using multi-agent conversations, which greatly reduce the effort to create LLM-based applications on different LLMs in various frames and enlarge the re-usability of LLM-agents as well.
A conversable agent is an entity endowed with a particular role that facilitates the exchange of messages with other agents, enabling the initiation and continuation of conversations. These agents rely on the messages they send and receive to maintain an internal context. Additionally, they can be programmed with a suite of capabilities, which may be derived from large language models, integrated tools, or direct human input. These agents operate in accordance with predefined behavior patterns. The design of the agents is outline in Algorithm 2.

```
img_gen_assistant = AssistantAgent( name="te
    xt_to_img_prompt_expert",
    system_message="""
        You are a text to image AI model expert, you will use text_to_image_generation function to
            generate image with prompt provided, and also improve prompt based on feedback
            provided until it is 10/10.
        For image generation tasks, only use the function you have been provided with. Reply
            TERMINATE when the task is done.
    """,
    llm_config=llm_config_assistants,
)

img_critic_assistant = AssistantAgent( name="i
    mg_critic", system_message="""
        You are an AI image critique, you will use img_review function to review the image
            generated by the text_to_img_prompt_expert against the original prompt, and provide
            feedback on how to improve the prompt.
        For all tasks, only use the functions you have been provided with. Reply TERMINATE when
            the task is done.
    """,
   llm_config=llm_config_assistants,
```

Algorithm 2 Design of Agents

4 CONCLUSION AND DISSCUSSION

To address a vision-related problem, we have integrated a diverse range of models within our AutoGen framework. This integration enables a multi-agent system to possess both image generation and image comprehension capabilities. We anticipate that this collaborative approach can be readily extended to incorporate other models and functionalities, thereby addressing a wide array of challenges. Currently, the capabilities of the language models and the models utilized in our work are not yet optimal. We are confident that with the support of more sophisticated models, our system will be able to provide enhanced solutions. The relevant code can be found at this URL.

FUNDING

The project was supported by Research on the Construction of Virtual Simulation Training Base for Vocational Education (LZJG2023041).

COMPETING INTERESTS

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

REFERENCES

[1] Yao, S, Zhao, J, Yu, D, et al. ReAct: Synergizing Reasoning and Acting in Language Models, 2023. DOI: https://doi.org/10.48550/arXiv.2210.03629.
[2] Wang, L, Ma, C, Feng, X, et al. A Survey on Large Language Model based Autonomous Agents. Frontiers of Computer Science, 2023, 18. DOI: https://doi.org/10.48550/arXiv.2308.11432.
[3] Li, C, Gan, Z, Yang, Z, et al. Multimodal Foundation Models: From Specialists to General- Purpose Assistants. Foundations and Trends® in Computer Graphics and Vision, 2023, 16(1-2): 1-214. DOI: https://doi.org/10.48550/arXiv.2309.10020.
[4] Li, C, Liu, H, Li, L, et al. Elevater: A benchmark and toolkit for evaluating language- augmented visual models. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, A. Oh, eds., Advances in Neural Information Processing Systems, 2022, 35, 9287-9301. DOI: https://doi.org/10.48550/arXiv.2204.08790.
[5] OpenAI. GPT-4 Technical Report. 2023.
[6] Yang, H, Yue, S, He, Y. Auto-GPT for Online Decision Making: Benchmarks and Additional Opinions, 2023. DOI: https://doi.org/10.48550/arXiv.2306.02224.
[7] Patil, SG, Zhang, T, Wang, X, et al. Gorilla: Large Language Model Connected with Massive APIs, 2023. DOI: https://doi.org/10.48550/arXiv.2305.15334.
[8] Nakajima, Y. BabyAGI, 2023.
[9] Li, G, Hammoud, H A A K, Itani, H, et al. CAMEL: Communicative Agents for "Mind" Exploration of Large Language Model Society. In Thirty-Seventh Conference on Neural Information Processing Systems, 2023, 2264, 51991 - 52008. DOI: https://doi.org/10.48550/arXiv.2303.17760.
[10] Wu, Q, Bansal, G, Zhang, J, et al. AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation, 2023. DOI: https://doi.org/10.48550/arXiv.2308.08155.
[11] Li, LH, Yatskar, M, Yin, D, et al. Visualbert: A simple and performant baseline for vision and language, 2019. DOI: https://doi.org/10.48550/arXiv.1908.03557.
[12] Hu, X, Gan, Z, Wang, J, et al. Scaling up vision-language pretraining for image captioning. In 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, 17959-17968. DOI: 10.1109/CVPR52688.2022.01745.
[13] Zhou, LW, Hamid Palangi, Zhang L, et al. Unified vision-language pre-training for image captioning and VQA. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(07): 13041-13049. DOI: https://doi.org/10.1609/aaai.v34i07.7005.
[14] Gan, Z, Li, L, Li, C, et al. Vision-language pre-training: Basics, recent advances, and future trends. Foundations and Trends® in Computer Graphics and Vision, 2022, 14(3-4): 163-352. DOI: https://doi.org/10.1561/0600000105.
[15] Li, M, Lv, T, Chen, J, et al. Trocr: Transformer-based optical character recognition with pre-trained models. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(11): 13094-13102. DOI: https://doi.org/10.1609/aaai.v37i11.26538.
[16] Chen, T, Saxena, S, Li, L, et al. Pix2seq: A language modeling framework for object detection. 2022. DOI: https://doi.org/10.48550/arXiv.2109.10852.
[17] Alayrac, J-B, Donahue, J, Luc, P, et al. Flamingo: a visual language model for few-shot learning. 2022. DOI: https://doi.org/10.48550/arXiv.2204.14198.
[18] Tsimpoukelli, M, Menick, J, Cabi S, et al. Multimodal few-shot learning with frozen language models. NIPS'21: Proceedings of the 35th International Conference on Neural Information Processing Systems, 2021, 16, 200-212.
[19] Liu, H, Li, C, Wu, Q, et al. Visual instruction tuning. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, S. Levine, eds., Advances in Neural Information Processing Systems, 2024, 36, 34892-34916. DOI: https://doi.org/10.48550/arXiv.2304.08485.
[20] Rombach, R, Blattmann, A, Lorenz, D, et al. High-resolution image synthesis with latent diffusion models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, 10684-10695. DOI: https://doi.org/10.48550/arXiv.2112.10752.