Scholar Publishing Group
International Journal of Big Data Intelligent Technology
https://doi.org/10.38007/IJBDIT.2026.070116
ISSN 2790-0932 Vol. 7, Issue 1: 131-139

Automated Operation Approach for Scalable Cloud Data Platform

Weiyao Ma

Robert H. Smith School of Business, University of Maryland, College Park, 20742, Maryland, USA

Keywords: Blockchain dynamic fragmentation model, revocable attribute based encryption, federated learning trust evaluation, zero knowledge proof authentication, edge computing outsourcing verification.

Abstract: The research background focuses on the core demands of secure and efficient transmission in the era of data sharing, revealing the three major challenges currently faced: economic value loss caused by cross organizational data silos, single point failure and privacy leakage risks in centralized platforms, and performance bottlenecks in large-scale node scenarios of blockchain. The research method innovatively proposes two technological paths: a blockchain dynamic sharding model based on trust management, which quantifies trust timeliness through spatiotemporal decay factors, achieves real-time sharding reassembly and load balancing, and solves the problem of traditional sharding security and performance imbalance; The revocable attribute based encryption scheme based on the blockchain adopts a private key/pre decryption key separation architecture and a computing outsourcing protocol to reduce computing overhead while ensuring revocation efficiency and adapt to edge computing scenarios. The research results have verified the superiority of the proposed scheme: experiments have shown that the system throughput significantly increases and latency decreases at the user scale; Effectively resist malicious node attacks and reduce the probability of failure through distributed ledger and trust management mechanisms; Compared to traditional attribute based encryption, key generation and encryption time overhead are reduced. The research conclusion emphasizes that this automated operation approach achieves full chain collaborative optimization through multidimensional technological innovation, supports the scalability and robustness of cloud data platforms in complex scenarios, and provides a secure and efficient technological paradigm for data sharing. Its value continues to be highlighted over time, supporting the development of enterprises, society, and government in multiple fields.

1 Introduction

In the era of data development, secure and efficient sharing of data has become an inevitable trend, and its value has become increasingly prominent over time, supporting the development of enterprises, society, and government in multiple fields. However, current data sharing faces three core challenges: firstly, the serious phenomenon of "data islands" formed by inter organizational data barriers. According to statistics, the economic value of open data sharing worldwide can reach 3 trillion to 5 trillion US dollars annually; Secondly, traditional centralized platforms have the risk of single point of failure and privacy breaches, while distributed systems partially alleviate this but face a wider attack surface; Thirdly, existing blockchain solutions have significant performance bottlenecks in actual large-scale node scenarios, with transaction processing per second (TPS)[1]dropping sharply to less than 20, and performance further declining in complex scenarios.

Existing research has shortcomings in addressing the aforementioned challenges. Although blockchain sharding technology is considered an efficient scaling solution, there are still shortcomings in terms of transaction trustworthiness, sharding security, and adaptive network load: different sharding nodes are prone to bias in determining transaction legitimacy, and the dynamic network environment is exploited by malicious nodes, resulting in a loss of transaction credibility; Random sharding allocation leads to trust imbalance, low trust sharding threatens system security, and existing systems have weak real-time perception of network load, making it difficult to dynamically adjust sharding scale. Attribute based encryption (ABE) [2] implements fine-grained access control, but there are issues with inefficient revocation mechanisms, high computational costs, low storage efficiency, and user consistency in attribute management. This study proposes two major improvement solutions to address the two core demands of "scalability" and "efficiency". Firstly, a blockchain sharding model based on trust management is designed [3], which considers trust difference, node number difference, and communication delay difference through node sharding strategy, maximizes system scalability and reliability, and proposes a trust management scheme for actively detecting and defending against internal attackers. A trust transfer chain is established to analyze trust decay from the spatiotemporal dimension; Secondly, design a blockchain based revocable attribute based encryption method that utilizes an indirect revocation mechanism to achieve user and attribute revocation, offloading complex bilinear operations to the proxy server and ensuring that it cannot obtain plaintext data. The experiment shows that this scheme is superior to similar schemes in terms of scalability, security, and efficiency: it supports optimized matching of new nodes and performs better in system throughput and latency under user scale; Effectively resist single point of failure and malicious node attacks through distributed ledger and trust management, with a lower probability of failure; Compared to traditional attribute based encryption, the cost of key generation, encryption, and update time is significantly reduced.

2 Correlation theory

2.1 Integration analysis of blockchain sharding and attribute based encryption technology

Blockchain sharding technology divides the network into interconnected small systems to enhance processing capabilities in order to address scalability issues. Network sharding forms logically independent sharding groups through node allocation algorithms, each responsible for internal transactions and block maintenance, achieving resource isolation and reducing node load. It requires dynamic allocation based on factors such as node reputation and geographic location, and designing efficient cross shard communication protocols (such as relay node mechanisms) to avoid data silos. Transaction sharding allocates transactions based on type or account address to improve parallelism, and requires coordinated operations across shards to ensure consistency. State sharding divides the global state (such as account balance), requiring collaborative maintenance of consistency and the design of a fast query mechanism to ensure integrity during state migration.Scalability analysis evaluates system performance under high load, including transaction throughput (on chain transaction volume per unit time) and latency (average time from transaction initiation to consensus). It is necessary to address challenges such as decreased consensus efficiency and increased network latency, providing theoretical basis for architecture design.Attribute based encryption (ABE) implements encryption based on attribute sets and is divided into two types: key policies[4](KP-ABE, policy embedded private key) and ciphertext policies (CP-ABE, policy embedded ciphertext), supporting fine-grained access control. Bilinear mapping satisfies bilinear, non degenerate, and computable properties, and is used in cryptographic schemes. The linear secret sharing scheme transforms access policy verification into a linear equation system solvability determination. Proxy re encryption allows semi trusted agents to convert ciphertext public keys, making it impossible for agents to obtain plaintext. It has the characteristics of unidirectionality, non interactivity, and minimal trust. It can be proven that security theory binds algorithm security to difficult problems (such as large integer factorization) through mathematical reduction, rigorously proves security under specific attack models, describes the attacker's ability range (such as choosing plaintext attacks), and reduces attack problems to solving difficult problems to indirectly verify security.

2.2 Design of blockchain sharding model based on spatiotemporal feature perception trust

Aiming at the frequent changes in network topology caused by dynamic access of network nodes and the difficulty of balancing security and network throughput in traditional blockchain architecture, a blockchain sharding model based on spatiotemporal feature aware trust algorithm (BSM-SFT) is proposed. This model constructs a trust calculation framework in both spatial and temporal dimensions, combined with an exponential decay function to achieve node trust evaluation and efficient sharding. Model introduces multi factor decision algorithm[5]to optimize node allocation: calculate communication delay difference, as shown in formula

$$D_{jb} = (t_h - t_0)/2$$

By minimizing the differences in communication latency, node count, and trust between shards, the system's security and scalability can be improved. The trustworthiness of sharding is calculated by weighting the trust values of nodes, and the trust levels are divided into high trustworthiness, trustworthiness, low trustworthiness, and untrusted. The trust update cycle is dynamically adjusted - low trustworthiness sharding uses smaller update cycles to detect malicious nodes in a timely manner, while high trustworthiness sharding uses larger cycles to balance resource overhead.The node allocation algorithm process is as follows: After initializing the sharding system, the new node calculates the sharding factor through communication with the sharding leader and assigns it to the shard with the smallest factor. When the system reaches the trust update cycle, the shard executes the trust evaluation algorithm. If the proportion of trusted nodes is below the threshold (such as 2/3), node reallocation is triggered - transferring nodes from the shard with more nodes to the current shard to maintain security and stability. Experiments have shown that the model effectively improves the transaction throughput of blockchain systems, reduces transaction latency and system failure probability, and achieves efficient, secure, and stable shard node allocation through spatiotemporal feature perception trust evaluation [6] and multi factor optimization.

3 Research method

3.1 Performance Evaluation Experiment of BSM-SFT Blockchain Sharding Model Based on OMNeT++

The experiment is based on the OMNeT++6.1 platform to simulate the blockchain network layer. The total number of nodes is 4000, the shard array is {4, 16}, and each shard contains 250 nodes. The communication bandwidth is 20mbps, the number of wireless interfaces is 1, the number of malicious nodes is {200, 1200}, and the trust value range is malicious nodes {0, 3}, trusted nodes

{3, 5}, and highly trusted nodes {5, 10}. The evaluation criteria cover transaction throughput (TPS, formula 3.30), transaction latency (TD, formulas 3.31-3.32), and blockchain failure probability (Pf, formulas 3.33-3.34). The comparison schemes are Monoxide and Rapidchain; The experimental results show that in terms of transaction throughput, when the number of shards increases from 4 to 16, the TPS of the BSM-SFT scheme continues to outperform the comparison scheme - when the number of shards is 4, the BSM-SFT reaches 5549, which is 14.4% higher than Monoxide (4852) and 22.4% higher than Rapidchain (4533); when the number of shards is 16, the BSM-SFT reaches 12439, which is 6.5% higher than Monoxide (11672) and 66.7% higher than Rapidchain (4756); In terms of transaction latency, when the number of shards is 4, the BSM-SFT latency is 42.4 seconds, which is lower than that of Monoxide (48.7 seconds) and Rapidchain (46.2 seconds). When the number of shards is 16, the BSM-SFT latency is 23.4 seconds, which is 29.3% lower than Monoxide (33.1 seconds) and 25.7% lower than Rapidchain (31.5 seconds); In terms of blockchain failure probability, when the proportion of malicious nodes increases to 50%, the failure probability of BSM-SFT decreases by 39.7% compared to Monoxide and 32.8% compared to Rapidchain. This is because the trust evaluation mechanism can timely identify and eliminate malicious nodes, suppress collusion behavior, and improve system security; Analysis shows that BSM-SFT significantly outperforms the comparison scheme in terms of throughput, latency, and failure probability through multi factor optimization and spatiotemporal feature trust evaluation, verifying the effectiveness of efficient sharding and node trust evaluation.

3.2 Design and Analysis of Revocable Attribute Based Encryption Scheme Based on Blockchain

Aiming at the efficiency bottleneck and inefficient revocation mechanism of traditional attribute based encryption technology in dynamic permission management, an efficient revocable attribute based encryption scheme integrating blockchain technology is proposed. This scheme divides keys into user keys held locally by users and pre decryption keys distributed under controlled conditions through an improved key segmentation mechanism. By combining the user pre decryption key mapping list, efficient revocation of users and attributes is achieved, and the security of the scheme is proven under the deterministic bilinear Diffie Hellman assumption.The system model consists of six entities: authorization authority, data owner, data visitor, proxy server, InterPlanetary File System (IPFS)[7], and blockchain. The authorized agency is responsible for attribute management, key distribution, and maintaining revocation lists; The data owner defines encryption policies and encrypts the data, storing the ciphertext in a distributed network; Data visitors must meet the attribute requirements in the access policy in order to decrypt data; Proxy servers perform computation offloading and ciphertext conversion to reduce the computational overhead of data owners and visitors; IPFS is responsible for storing encrypted data and metadata, ensuring data persistence and verifiability through content addressing; Blockchain stores encrypted data tuples, including encrypted data keywords, hash values, timestamps, owners, and data types. The plan includes six stages: system startup, key establishment, data encryption, attribute authentication, ciphertext decryption, and user/attribute modification. The system startup phase is executed by the authorized agency, generating the system public key and master key; The key establishment phase is executed by the data owner and data visitor, generating public-private key pairs and pre decryption keys; The data encryption stage includes symmetric encryption, symmetric key pre encryption, and symmetric key re encryption algorithms; The attribute authentication stage is executed by the proxy server to verify whether the user attributes meet the access policy; The ciphertext decryption stage restores plaintext through pre decryption and local decryption algorithms; The user/attribute change phase supports revocation algorithms and attribute updates, dynamically adjusting the pre decryption key mapping table. The security model uses the game process between challengers and adversaries, and based on the Indifferentiation of Explicit Attacks (IND-CPA) [8], attacks the game to prove the security of the proposed solution. Experimental verification shows that this scheme significantly optimizes the computational overhead in key generation, encryption, decryption, and key update, effectively improving the efficiency and security of attribute based encryption in dynamic permission management.

3.3 Design and Performance Analysis of Revocable Attribute Based Encryption Scheme Based on Blockchain

This solution addresses the efficiency bottleneck and inefficient revocation mechanism of traditional attribute based encryption technology in dynamic permission management, and proposes an efficient revocable attribute based encryption framework that integrates blockchain technology. During the system startup phase, the bilinear group generation algorithm is used to output the system public key and master key; The key establishment stage adopts a user local key and pre decryption key segmentation mechanism, combined with a user pre decryption key mapping list to achieve efficient user/attribute revocation; The data encryption stage includes symmetric encryption, symmetric key pre encryption, and proxy re encryption algorithms to reduce the computational overhead of data owners; The attribute authentication stage verifies the legitimacy of user attributes through attribute matrix and vector operations; The data decryption stage is pre decrypted by the proxy server, and the user only needs constant level calculations to restore the plaintext; The user/attribute revocation phase supports dynamic adjustment of the pre decryption key mapping table to ensure that requests return an invalid state directly after revocation. Security is based on the discriminative bilinear Diffie Hellman assumption, and indistinguishability is proven through IND-CPA attack games. Experimental verification shows that in the key generation, encryption, decryption, and key update stages, this scheme significantly optimizes the computational cost compared to the comparative scheme: the time for attribute set size 100 in the key generation stage is reduced by 85.96%, the time for policy tree depth 5 and leaf nodes 20 in the encryption stage is reduced by 70.81%, the computational complexity of the user end is optimized to a constant level in the decryption stage, and the time for attribute set size 100 in the key update stage is reduced by 89.09%. Blockchain and InterPlanetary File System (IPFS) respectively undertake the functions of access control metadata storage and actual data storage. Through smart contracts, dynamic authorization and auditing are implemented to ensure that data sharing behavior can be verified and responsibility can be traced, while reducing the risk of single point of failure in centralized architecture.

4 Results and discussion

4.1 Simplified Design of Blockchain Data Sharing System

The system adopts a four layer architecture design, and the front-end layer is based on Vue to develop functional interactive pages and API data transmission gateways. It dynamically transmits user data, timestamps, metadata, etc. through the HTTP protocol, and supports real-time display of shard node status, block height, and transaction content; The application layer serves as the system hub, integrating distributed microservice architecture to process data aggregation and shard topology information storage, implementing trust evaluation algorithms based on spatiotemporal feature perception and revocable efficient attribute based encryption algorithms, supporting shard expansion/contraction, transaction set maintenance, and node state management; The blockchain layer includes identity authentication services, smart contracts, functional nodes, and shard models.

Identity authentication is provided by CA nodes for dynamic join auditing and key distribution. Smart contracts implement trust evaluation, transaction queries, status data updates, and other functions. Functional nodes include submission nodes, endorsement nodes, sorting nodes, and ledger nodes to ensure transaction legality, global consistency, and final submission; The data storage layer uses CouchDB to store state data and support rich queries, IPFS to store encrypted data, MySQL to store system public keys, trust state data, and ciphertext location hashes, combined with channel settings to achieve sharded data isolation storage. The core data structure of the system includes three types of contracts: node trust evaluation contracts calculate node trust values through multiple dimensions such as historical interaction records, behavior logs, and direct trust degrees. The data structure includes fields such as node unique identifiers, dynamic trust values, timestamps, status identifiers, and shard identifiers; The node management contract supports real-time query and dynamic update of node status, and the data structure includes fields such as node identifier, latest trust value, shard identifier, transaction set, heartbeat time, etc; The data storage contract implements on chain and off chain data binding, and the data structure includes fields such as file name, username, file address hash, access policy set, transaction hash, block hash, timestamp, file status, and effective time. The core functional module includes four major modules: the shard scheduling module is based on the Apache Flink stream processing engine to achieve real-time data cleaning and preprocessing, and efficient communication between the application layer, blockchain layer, and data storage layer is built through the producer consumer mode. It supports uploading IPFS and location hash after data encryption, and uploading keys to the blockchain; The data encryption and decryption module is based on the revocable CP-ABE algorithm to implement fine-grained access control. It executes encryption or decryption processes according to state parameters. During encryption, the access control structure is transformed into an LSSS matrix and plaintext is encrypted. During decryption, ciphertext is requested through location hashing and decrypted using the user's private key; The trust evaluation module calls smart contracts to execute trust evaluation algorithms based on spatiotemporal feature perception through scheduled tasks, summarizes node evaluation results, and updates trust values; The distributed data storage module stores blockchain state data through MySQL, CouchDB supports complex queries, IPFS stores encrypted data, including ciphertext location hash tables, system public key tables, state query routing tables, and trust state data tables, achieving real-time data access, classified storage, and visual display.

4.2 Model experiment

The system technology architecture is divided into four layers: the front-end interface layer integrates JWT identity authentication and Vue responsive framework, providing user interaction entry points; The service layer is based on the Spring Boot framework [9], integrating Maven project construction, MyBatis database interaction, and Fabric Gateway client SDK to achieve business logic processing; The blockchain layer adopts Hyperledger Fabric 2.2.1 and deploys shard clusters (such as shard clusters 1/2/3), including submission nodes, endorsement nodes, sorting nodes, and smart contracts, supporting on chain data authentication and channel configuration; The data storage layer combines MySQL 8.0.41, CouchDB 3.3.0, and IPFS 0.10.0 to achieve structured data, chain state data, and encrypted file storage. The system environment component versions include JDK 1.8.0, Fabric SDK 2.2.0, JPBC 2.0.0, Docker 20.0.4, etc. The setup steps are as follows: Start three Ubuntu 20.04 virtual machines, each running a Docker environment containing Order, Peer0, and Peer1 nodes. The node configurations are shown in Table 1

The Java backend communicates with the blockchain network through key files (such as ordererorgadmins, peer0orgadmin) to configure network connection files and channel/chain code parameters; The access control structure (such as (A and B) or (C and D)) is transformed into a binary tree through breadth first search, with attributes stored as leaf nodes and logical operators stored as non leaf nodes. It is then transformed into an LSSS two-dimensional matrix [10],with each row corresponding to a leaf node vector. The transformation process is shown in Figure 1.

Table Name 1 Node Configuration Table

Node Type	IP Address	Service Components
Order	192.168.X.100	Order Service
Peer0	192.168.X.200	Peer Service, CouchDB Component
Peer1	192.168.X.201	Peer Service, CouchDB Component
Peer2	192.168.X.202	Peer Service, CouchDB Component

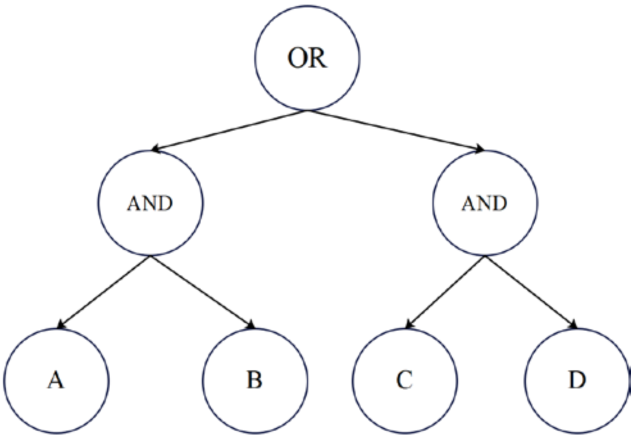

Figure 1 Logic structure diagram of access control

The system embeds encryption and decryption interfaces through chain codes to achieve on chain and off chain collaborative security management.

4.3 Effect analysis

The automated operation approach of the scalable cloud data platform achieves full process collaboration through a modular system architecture: the program automatically initializes the front-end, back-end, blockchain, and database environment after startup, establishes scheduled task synchronization node status, trust evaluation, and data cleaning. The blockchain network is deployed on a shard cluster based on Hyperledger Fabric. Order, Peer0, and Peer1 nodes are launched and smart contracts are loaded through a Docker environment. The Java backend integrates the JPBC encryption library to generate system master keys and public parameters, and combines host mapping to achieve blockchain network connectivity. The data is encrypted and stored using the CP-ABE algorithm. During encryption, the access control policy is converted into an LSSS matrix and linked to metadata on the chain. During decryption, policy matching is verified through attribute private key authentication, supporting fine-grained access control and decentralized permission management. The front-end builds a visual interface based on Vue.js, integrating dashboard, file upload, blockchain management, and data query modules. It interacts with the back-end through RESTful API, supports dynamic definition of access policy tree, real-time synchronization of attribute changes to the blockchain network, and embeds encryption and decryption interfaces to achieve audit tracking and on chain and off chain collaborative management, ensuring the security, controllability, and traceability of data sharing process.

5 Conclusion

The automated operation approach of the scalable cloud data platform achieves full chain collaborative optimization through multi-dimensional technological innovation: adopting a multi factor based blockchain dynamic sharding model, combined with spatiotemporal decay factors to quantify trust timeliness, to achieve real-time sharding reassembly and horizontal expansion capability improvement; Design a revocable attribute based encryption scheme that solves the key custody problem through a private key/pre decryption key separation architecture, and adapts to resource constrained scenarios in conjunction with computing outsourcing protocols to improve policy update efficiency. The system architecture adopts a layered design, integrating the front-end interaction layer, application logic layer, blockchain consensus layer, and distributed storage layer. It modularly implements sharding scheduling, data encryption and decryption, trust evaluation, and storage management functions, and supports full lifecycle data sharing and management. The automated operation path is further extended to the federated learning driven trust dynamic evaluation system. Through distributed machine learning, real-time prediction of node trust values and cross chip load balancing are achieved. Combined with the attribute authentication mechanism of zero knowledge proof fusion, fine-grained policy hiding is realized while ensuring revocation efficiency. A lightweight outsourcing verification protocol oriented to edge computing is developed, and a key escrow defense system for multiple authorization centers is built. Finally, the robustness and scalability of the system in malicious node injection, network topology change and other scenarios are ensured through the piecemeal performance simulation platform and complex scenario test benchmark.

References

[1] *Chithaluru P , Al-Turjman F , Dhatterwal K J S .An enhanced consortium blockchain diversity mining technique for IoT metadata aggregation[J].Future generations computer systems: FGCS, 2024, 152(Mar.):239-253.DOI:10.1016/j.future.2023.10.020.*

[2] *Vali S J .ATTRIBUTE BASED ENCRYPTION IN IOT DEVICES[J].Futuristic Trends in IOT Volume 3 Book 4, 2024:171-179.DOI:10.58532/v3bdio4p2ch1.*

[3] *Inayatulloh, Pelawi D ,Witarsjah,et al.Blockchain Sharding Model to Increase The Efficiency of Waste Management Processes Based on Blockchain Technology[J].2024 4th International Conference on Innovations in Computer Science (ICONICS), 2024:1-5.DOI:10.1109/iconics64289.2024.10824439.*

[4] *Barber T M , Kabisch S , Pfeiffer A F H ,et al.Optimised Skeletal Muscle Mass as a Key Strategy for Obesity Management[J].Metabolites (2218-1989), 2025, 15(2).DOI:10.3390/metabo15020085.*

[5] *Zhen Zhong. Big Data Engineering and Intelligent Analysis Framework for Compliance Investigation. Academic Journal of Computing & Information Science (2025), Vol. 8, Issue 11: 107-115*

[6] *Wu, W. (2025, June). Construction and optimization of intelligent gateway software management platform based on jenkins cluster management under cloud edge integration architecture in industrial internet of things. In International Conference on 6G Communications Networking and Signal Processing (pp. 633-645). Singapore: Springer Nature Singapore.*

[7] Wu, L. (2025, December). Design and Application of Automatic Data Set Generation Tool Based on KLEE in Embedded Memory Management Performance Test Framework. In 2025 IEEE 17th International Conference on Computational Intelligence and Communication Networks (CICN) (pp. 1111-1117). IEEE.

[8] Ye, J. (2025). Multimodal medical data intelligent classification method and system implementation based on improved SVM and similarity learning algorithm. International Journal of World Medicine, 2025, 6 (1), 19, 27.

[9] Truong, T. H. (2026). Research on Risks and Countermeasures of Enterprise E-Commerce Transformation in the Cross-Border E-Commerce Environment.

[10] Qi, Y. (2025, October). Research on Privacy Protection of AI Models in Big Data Using Differential Privacy Technology. In 2025 2nd International Conference on Software, Systems and Information Technology (SSITCON) (pp. 1-5). IEEE.

[11] Hong, Y. (2025). Architecture Design and Performance Optimization of a Large-scale Online Simulation Platform for Business Decision-making. Advances in Computer and Communication, 6(4).

[12] Ye, J. (2025). Design of a Non-Invasive Brain Computer Interface System for Handwritten Text Based on L2 Regularization and Attention Supervision Paradigm, and Optimization of EEG Signal Decoding. International Journal of Big Data Intelligent Technology, 2025, 6 (1), 126, 134.

[13] Hong, Y. (2026). Research on Warehouse Capacity Optimization Methods Based on Predictive Modeling. Engineering Advances, 6(1).

[14] Truong, T. H. (2025). Research on the Application of Digital Healthcare Platforms in Chronic Disease Management. Advances in Computer and Communication, 6(5).

[15] Ye, J. (2025). Optimization of Neural Motor Control Model Based on EMG Signals. International Journal of Engineering Advances, 2(4), 1-8.