International Journal of Computer Science and Information Technology
ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 4, Number 1, Year 2024
DOI: https://doi.org/10.62051/ijcsit.v4n1.30
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Analysis of Key Technologies for Data Security Protection Based on Cloud Computing

Ruizhe Li

Uber Technologies, Inc., San Francisco, USA

ABSTRACT

With the rapid development of cloud computing, more and more enterprises and institutions are migrating their data and business operations to the cloud. However, this shift has brought data security issues in cloud environments to the forefront. This paper primarily investigates the key technologies for data security protection in cloud computing environments. It begins by introducing the basic concepts and architecture of cloud computing and analyzing the fundamental concepts and threats related to data security protection. The paper then delves into critical technologies such as data encryption, access control, data storage and backup security, and data privacy protection, detailing the algorithmic principles and implementation methods of these technologies. Additionally, the paper explores the challenges faced by current cloud computing data security protection technologies and their corresponding solutions, analyzing the effectiveness of these technologies through practical case studies. Finally, the paper forecasts the future development trends of cloud computing data security technologies, particularly the potential impact of emerging technologies on data security protection. This research aims to provide theoretical and technical support for data security protection in cloud computing environments.

KEYWORDS

Cloud Computing; Data Security; Encryption Algorithms; Access Control

1. INTRODUCTION

Cloud computing, as an emerging information technology, has rapidly become a core driving force in digital transformation across various industries worldwide, thanks to its powerful computing capabilities, flexible resource allocation, and efficient cost management. As more enterprises and institutions migrate their critical business operations and vast amounts of data to the cloud, the scope and depth of cloud computing applications continue to expand. However, the openness of cloud computing environments, multi-tenant architecture, and dynamic resource allocation also bring unprecedented data security challenges. Frequent issues such as data breaches, privacy violations, and unauthorized access have made data security protection a critical problem in the field of cloud computing.To address these challenges, it is crucial to research and develop efficient data security protection technologies tailored to cloud computing environments. This paper aims to systematically analyze the data security threats in cloud computing and explore and summarize the existing key data security protection technologies, including data encryption, access control, data storage and backup security, and data privacy protection. By deeply analyzing the algorithmic principles, implementation methods, and practical applications of these technologies, this paper seeks to provide theoretical and practical guidance for data security protection in cloud computing environments [1].

2. OVERVIEW OF CLOUD COMPUTING

Cloud computing is a technology that provides computing resources and services over the internet, fundamentally changing the traditional computing model. It consolidates processing power, storage space, and network services into remote data centers managed by professional service providers, allowing users to access these resources on-demand from anywhere. The core of cloud computing is resource virtualization and on-demand provisioning, enabling businesses and individuals to flexibly adjust resources and reduce the high costs associated with traditional IT infrastructure. As shown in Figure 1, the architecture of cloud computing is divided into two parts: the front end and the back end. The front end consists of client-side infrastructure and the internet, through which users connect to the cloud using various devices such as PCs, smartphones, or tablets. The back end includes applications, services, storage, and infrastructure that support user requests and provide efficient computing resources. In this architecture, security and manageability are crucial, ensuring data confidentiality, integrity, and availability while providing flexible management mechanisms [2].

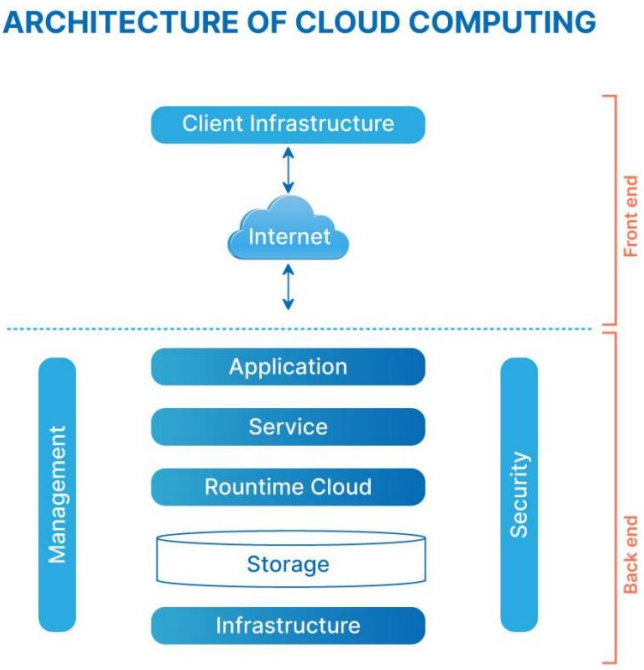

Figure 1. Cloud Computing Architecture

The widespread adoption of cloud computing is mainly due to its unique features. It offers on-demand self-service, allowing users to flexibly acquire resources, reducing time and cost. Cloud computing also features broad network access, enabling users to connect to cloud services using various devices through standardized interfaces, improving efficiency. Additionally, cloud computing implements resource pooling technology to share resources, enhancing system scalability. Its rapid elasticity allows resources to be quickly adjusted based on demand, meeting the fast-changing business needs of enterprises. The pay-per-use model reduces upfront costs, making cloud computing services attractive to businesses of all sizes.Cloud computing has become an integral part of the global IT infrastructure, widely applied in industries such as finance, healthcare, education, and manufacturing. As technology advances, cloud computing is evolving towards greater intelligence, security, and customization. The integration of AI, big data analytics, and the Internet of Things has broadened the application scenarios of cloud computing, enhancing automation and intelligence. Meanwhile, the increase in data volume and the growing demand for privacy protection have made security issues in cloud computing more prominent. Frequent incidents of data breaches, privacy violations, and unauthorized access have driven cloud service providers to continuously optimize security technologies and standards, such as improving encryption algorithms and optimizing access control mechanisms. In the future, cloud computing will continue to develop towards efficiency, security,

and intelligence, integrating emerging technologies like quantum computing and edge computing to enhance processing capabilities and application scope. The continuous improvement of data protection regulations worldwide will also drive further advancements in the security and compliance of cloud computing services.

3. BASIC CONCEPTS OF DATA SECURITY PROTECTION

Data security protection involves using technical and management measures to ensure the confidentiality, integrity, and availability of data throughout its lifecycle. With the widespread adoption of cloud computing, most data storage, processing, and transmission occur in the cloud, posing greater data security challenges. The multi-tenant architecture, virtualization technology, and remote access features of cloud computing, while enhancing resource utilization and flexibility, also increase the risks of data breaches, unauthorized access, and data tampering. Therefore, data security protection has become crucial in cloud computing environments [3].

Figure 2. Overall Framework of Data Security Protection Strategies

Figure 2 illustrates the overall framework of data security protection strategies, which include both technical and management measures to address threats such as data breaches, malicious attacks, and data loss. These strategies encompass encryption, access control, firewalls, policy formulation, risk assessment, and compliance checks. By applying these measures, enterprises can effectively reduce data security incidents and ensure the safety of their data. In cloud computing, the main threats to data include breaches, tampering, and loss. Data breaches can occur through cyber-attacks or malicious actions by insiders, tampering may happen via man-in-the-middle attacks, and data loss can be caused by equipment failure or malware. These threats present multiple challenges for both technology and management in data security protection.To counter these threats, the industry has developed various data protection methods and technologies. Data encryption is a key method for ensuring data confidentiality, with common techniques including symmetric encryption, asymmetric encryption, and homomorphic encryption. Access control, through fine-grained permissions management, ensures that only authorized users can access data, with models such as RBAC and ABAC being widely used. Regular backups, offsite storage, and disaster recovery mechanisms are

effective in preventing data loss. In multi-tenant environments, technologies such as anonymization, pseudonymization, and differential privacy effectively protect data privacy. These methods and technologies, when combined, enhance the level of data security in cloud computing environments. Looking ahead, data security protection technologies will evolve towards greater intelligence and automation to better address complex security threats.

4. DATA SECURITY THREATS IN CLOUD COMPUTING ENVIRONMENTS

As cloud computing becomes more widespread, data security issues have become a major concern for many enterprises and organizations. Data security threats in cloud computing environments are multifaceted, involving data storage, transmission, access, and management. Figure 3 outlines the main data security challenges in cloud computing, which not only include potential data breaches and privacy violations but also impact business continuity and regulatory compliance [4].

Figure 3. Major Data Security Challenges in Cloud Computing Environments

Data breaches are the most common and severe threat. Due to the openness of cloud services and the multi-tenant architecture, unauthorized users may gain access to sensitive data through insecure APIs, unencrypted transmissions, or malicious insider activities. Access control is crucial in cloud computing, but traditional methods face challenges in environments with resource sharing and diverse access methods, which may lead to unauthorized access or data tampering. Insecure APIs serve as the main entry point for attackers to infiltrate systems, potentially bypassing authentication mechanisms, leading to data breaches or tampering. Data loss is another significant issue, which can be caused by equipment failure, malicious attacks, or human error. Although cloud service providers typically offer backup and recovery services, ensuring rapid and complete data recovery in large-scale data processing scenarios remains challenging.In a multi-tenant environment, multiple users share the same physical infrastructure, which, while improving resource utilization, also increases security risks. If the infrastructure has vulnerabilities, attackers may compromise one tenant's data to access others'. As global data protection regulations become more stringent, companies using cloud computing must ensure that their data processing complies with local laws. This is particularly challenging in cross-border operations, where meeting the regulations of different jurisdictions is a significant task. Data privacy and confidentiality also face new challenges, as unclear responsibilities can result in insufficient protection during data transmission and storage, while government and legal requirements for data access further increase the risk of breaches.In summary, data security threats in cloud computing environments are diverse, requiring enterprises to implement comprehensive security measures, including both technical and management strategies, to ensure the safety, integrity, and availability of their data.

5. KEY TECHNOLOGIES FOR DATA SECURITY PROTECTION IN CLOUD COMPUTING

5.1. Data Encryption Technologies

Data encryption is one of the core methods to ensure data security in cloud computing environments. By using encryption, data can be protected during storage and transmission, preventing unauthorized access and information leaks [5]. Common data encryption techniques in cloud computing include symmetric encryption, asymmetric encryption, homomorphic encryption, and attribute-based encryption. Below, we will introduce the basic principles, algorithmic formulas, and applications of these encryption techniques in cloud computing.Symmetric encryption uses the same key for both encryption and decryption. Its main advantages are speed and high computational efficiency, making it suitable for encrypting large volumes of data. In symmetric encryption, the encryption and decryption algorithms are inverses of each other, as shown in formulas 1 and 2:

$$\text{Encryption: } C = E(K, P) \tag{1}$$

$$\text{Decryption: } P = D(K, C) \tag{2}$$

Where P represents plaintext, C represents ciphertext, K represents the key, and E and D are the encryption and decryption functions, respectively. Common symmetric encryption algorithms include AES and DES. For example, AES involves multiple rounds of byte substitution, row shifting, column mixing, and key addition, ensuring high non-linear mapping between ciphertext and plaintext, thereby enhancing data security.

Asymmetric encryption, also known as public-key encryption, uses two keys: a public key (public) and a private key (secret). The public key is used for encryption, and the private key is used for decryption. Asymmetric encryption is advantageous for key management and is well-suited for secure data transmission in cloud environments. Its mathematical representation is as follows:

$$\text{Encryption: } C = E(K_{pub}, P) \tag{3}$$

$$\text{Decryption: } P = D(K_{pri}, C) \tag{4}$$

Where K_{pub} is the public key, K_{pri} is the private key, P is plaintext, and C is ciphertext. Common asymmetric encryption algorithms include RSA and ECC.

Homomorphic encryption allows for operations like addition and multiplication on ciphertexts, which is highly useful for processing encrypted data in the cloud.Attribute-based encryption (ABE) is a technique that encrypts data based on user attributes, allowing the data encryptor to define access policies. Only users who meet specific attribute conditions can decrypt the data. ABE can be key-policy (KP-ABE) or ciphertext-policy (CP-ABE). The mathematical representation of CP-ABE is as detailed in formulas 5, 6:

$$\text{Encryption: } C = E(PK, P, A) \tag{5}$$

$$\text{Decryption: } P = D(SK, C) \tag{6}$$

Where PK is the public key, SK is the user's private key, P is plaintext, C is ciphertext, and A is the access policy. In CP-ABE, ciphertexts are associated with an access structure, and only users with the corresponding attribute-based private key can decrypt the data. CP-ABE is commonly used in cloud storage services for access control, ensuring that only authorized users can access sensitive

information.In summary, data encryption is the cornerstone of data security in cloud computing. Various encryption algorithms effectively prevent data breaches, tampering, and unauthorized access, providing robust security in cloud environments [6].

5.2. Access Control Technologies

Access control is another crucial technology for ensuring data security in cloud computing environments. It restricts and manages user access to data resources, ensuring that only authorized users can perform legitimate operations, thus preventing data breaches and tampering. In cloud computing, where multi-tenant architecture and resource sharing are common, access control is particularly important. This section focuses on two main access control technologies: Role-Based Access Control (RBAC) and Attribute-Based Access Control (ABAC).Role-Based Access Control (RBAC) manages and assigns permissions based on user roles. In RBAC, users are assigned one or more roles, each corresponding to a set of permissions. Users inherit the permissions associated with their roles, allowing them to perform specific operations. The core idea of RBAC is to shift permission management from the individual user level to the role level, simplifying the complexity of managing permissions.

The mathematical representation of RBAC is as detailed in formulas 7 through 9:

$$\text{-Role-to-permission mapping:} R = \{r_1, r_2, \ldots, r_n\} \text{ where R represents a role.} \quad (7)$$

$$\text{-User-to-role mapping: } U = \{u_1, u_2, \ldots, u_m\}, \text{ where U represents a user.} \quad (8)$$

$$\text{-Permission set: } P = \{p_1, p_2, \ldots, p_k\}, \text{ where P represents a permission.} \quad (9)$$

When a user u is assigned a role r, they inherit all the permissions p associated with that role, allowing them to perform the corresponding operations. RBAC is efficient in managing and controlling permissions, particularly in complex enterprise environments, but its limitation lies in its lack of flexibility, especially in scenarios requiring fine-grained control.Attribute-Based Access Control (ABAC) is a more flexible and fine-grained access control model. Unlike RBAC, ABAC considers not only user roles but also attributes of users, resources, and the environment. ABAC defines complex policies based on these attributes to determine whether a user should be allowed access to a particular resource, offering greater flexibility and scalability, especially in dynamic cloud environments. The mathematical representation of ABAC is as detailed in formulas 10, 11:

$$\text{Attribute set: } A = \{a_1, a_2, \ldots, a_n\} \text{ where } a_i \text{ represents an attribute (e.g., user, resource, or environmental attribute).} \quad (10)$$

$$\text{Policy set: } S = \{s_1, s_2, \ldots, s_n\}, \text{ where } s_i \text{ represents an access control policy.} \quad (11)$$

In ABAC, access control policies are typically defined as Boolean expressions. For example as formulas12 shown :

$$s1: \text{if } (\text{role} = \text{"admin"}) \wedge (\text{resource\_level} = \text{"low"}) \text{ then allow access} \quad (12)$$

This policy can be flexibly defined based on different attribute combinations, enabling fine-grained access control. ABAC's strength lies in its high flexibility and precision, supporting complex access control needs, particularly in multi-tenant cloud environments where different users can be granted specific access permissions based on their attributes, ensuring resource security and compliance.In cloud environments, RBAC and ABAC are often combined to leverage their respective strengths. For instance, enterprises can use RBAC to manage regular roles and permissions for daily access control,

while ABAC can be used to further refine access control in scenarios involving highly confidential data, ensuring that only users meeting specific attribute conditions can gain access.In practice, RBAC is typically used to establish the basic permission framework, while ABAC is employed to define complex access policies that cater to dynamic environmental needs. This combination enhances system security and increases the flexibility and manageability of permissions.In conclusion, access control is a crucial component of ensuring data security in cloud computing environments. By combining RBAC and ABAC, it is possible to effectively manage and control user access to data, preventing unauthorized access and data breaches, thereby providing a solid foundation for data security in cloud computing [7].

6. CHALLENGES AND SOLUTIONS FOR DATA SECURITY PROTECTION IN CLOUD COMPUTING

In cloud computing environments, data security faces challenges from dynamic data handling, large-scale data processing, and the integration of emerging technologies. Addressing these challenges requires optimizing existing technologies and exploring innovative solutions. Firstly, optimizing dynamic data encryption and decryption algorithms is crucial. Stream encryption allows data to be encrypted byte by byte without waiting for the entire data block, enhancing real-time data processing. Parallel computing and hardware acceleration (such as GPU) can also improve encryption and decryption speeds. Additionally, dynamic key management strategies generate and update keys in real time based on data dynamics, preventing security risks. Secondly, privacy protection in large-scale data processing presents challenges [8]. Differential privacy ensures the privacy of individual data by introducing noise into analysis results. Homomorphic encryption enables computations on encrypted data without decryption, making it suitable for handling sensitive data in cloud computing. Secure Multi-Party Computation (SMPC) ensures privacy protection during cross-organizational data collaboration. Lastly, the combination of machine learning and blockchain technologies brings new opportunities for data security in cloud computing. Machine learning enhances the intelligence of threat detection and encryption algorithms, while blockchain, with its decentralized and tamper-proof nature, strengthens data storage, access control, and audit tracking. Smart contracts in blockchain can effectively manage data sharing, ensuring data confidentiality and privacy. In summary, cloud computing data security faces multiple challenges, but by combining and optimizing these technologies, a more secure and reliable cloud environment can be built, enhancing data security while increasing system flexibility and scalability.

7. CASE STUDIES OF DATA SECURITY PROTECTION TECHNOLOGIES IN CLOUD COMPUTING

7.1. Analysis of Data Encryption Technology in Amazon Web Services (AWS)

Amazon Web Services (AWS), one of the world's leading cloud service providers, offers data encryption technologies that cover multiple stages from data storage to transmission. AWS provides several encryption services, including static data encryption for Amazon S3 and database encryption for Amazon RDS [9]. The encryption options offered by AWS include: Static Data Encryption: Using Customer Master Key (CMK) to encrypt data stored in S3 or RDS. Data Encryption in Transit: Using TLS/SSL protocols to protect data transmission between clients and servers. Customer-Managed Keys: Customers can manage their own encryption keys through AWS Key Management Service (KMS). <Table 1> shows the performance data of AWS when processing static data encryption, displaying encryption and decryption speeds of different algorithms (AES-128, AES-256) and their impact on processing time and resource utilization.

Table 1. Performance Data for Encryption

Encryption Algorithm	Encryption Speed (MB/s)	Decryption Speed (MB/s)	CPU Usage	Memory Usage
AES-128	150	160	30%	10%
AES-256	120	130	35%	12%

The table indicates that AES-128 has advantages in speed and resource usage, while AES-256 provides higher security, suitable for scenarios with stringent security requirements. To optimize data encryption efficiency, AWS introduced parallel encryption technology. Assuming the original data block is D, the mathematical representation of the parallel encryption algorithm is shown as formula 13:

$$C_i = E(K_i, D_i) \text{ for } i = 1,2, \ldots, n \tag{13}$$

Where C_i is the i-th encrypted data block, D_i)) is the i-th data block, and K_i is the i-th key. Through parallel processing, AWS can significantly improve encryption speed, especially in large-scale data encryption tasks [10].

7.2. Data Protection Strategies and Algorithm Implementation in the Financial Industry

The financial industry has extremely high requirements for data security, especially when handling sensitive customer information and transaction data, where the confidentiality, integrity, and availability of data must be ensured. Many financial institutions have migrated part or all of their operations to cloud computing platforms and have established stringent data protection strategies. Common data protection strategies in the financial industry include: Full Lifecycle Data Encryption: Encrypting data at every stage—storage, processing, and transmission. Multi-layer Access Control: Combining RBAC and ABAC to ensure that only authorized users can access specific data.Data Auditing and Monitoring: Real-time monitoring and periodic audits to ensure data usage complies with company policies and regulatory requirements. <Table 2> shows the performance data of a differential privacy algorithm used by a financial institution in a cloud computing environment, employed to protect the privacy of transaction data.

Table 2. Performance Data of Differential Privacy Algorithm

Algorithm Type	Data Processed (Million Records)	Average Latency (ms)	Privacy Protection Strength (ε)	Data Loss Rate (%)
Differential Privacy	1	50	0.1	1.2%
No Differential Privacy	1	30	N/A	0.0%

The table 2 shows that the differential privacy algorithm increases processing latency and data loss rate but significantly enhances privacy protection, meeting the high standards of data security required in the financial industry. To reduce the computational overhead of the differential privacy algorithm, the financial industry has introduced a privacy budget optimization algorithm based on gradient descent, mathematically represented as shown in formula 14:

$$\epsilon_{t+1} = \epsilon_t - \eta \nabla L(\epsilon_t) \tag{14}$$

Where ϵ_t is the privacy budget at iteration t, η is the learning rate, and $L(\epsilon_t)$ is the loss function. Iterative optimization allows meeting privacy protection requirements while minimizing data loss and computational delay. In recent years, frequent data breaches have caused severe economic losses and reputational damage to companies. The following is an analysis of a significant data breach and the

corresponding improved algorithm solution. A company experienced a large-scale data breach when migrating customer data to a cloud computing platform, resulting in the unauthorized access of approximately 500,000 customer records. The main reasons for the breach were the company's failure to update access control strategies in a timely manner and the use of weak encryption algorithms during data transmission, allowing attackers to easily crack the ciphertext. The incident exposed the following issues:Inadequate Access Control Strategy: Static access control was used without dynamic adjustment based on user behavior.making them easy to crack. Lack of Multi-layered Data Protection: Data transmission relied only on a single layer of encryption protection, without adopting multi-layered protection measures. To prevent similar incidents, the company decided to implement the following improvements:Dynamic Access Control Algorithm: Combining ABAC and machine learning algorithms to dynamically adjust user access permissions. The improved access control algorithm is represented as shown in formula 15:

$$P = f(A_{user}, A_{resource}, C) \tag{15}$$

Where P is the access permission, A_{user} is the user attribute, $A_{resource}$ is the resource attribute, and C is the context. By analyzing user behavior and context in real-time, access permissions are dynamically adjusted.Enhanced Encryption Algorithm: Adopting AES-256 combined with a distributed Key Management System (KMS) for data encryption. The improved encryption algorithm is represented as shown in formula 16:

$$C_i = E(KMS(K_{user}, K_{session}), D_i) \tag{16}$$

Where KMS is the key management system, K_{user} is the user key, $K_{session}$ is the session key, D_i is the data block, and C_i is the encrypted data block. Distributed key management enhances data encryption security. Multi-layered Data Protection: In addition to using TLS/SSL protocols during data transmission, an additional layer of protection based on differential privacy is introduced at the application layer, ensuring that even if transmission encryption is compromised, the data remains protected. These improvements have significantly enhanced the company's data protection capabilities, effectively mitigating potential data breach risks. In conclusion, by analyzing data encryption technologies and protection strategies in real-world applications, it is evident that the rational application of improved algorithms and multi-layered protection measures can significantly enhance data security in cloud computing environments. These case studies provide valuable reference solutions for other companies and guide the future development of data security protection technologies.

8. CONCLUSION

The development of cloud computing has brought unprecedented flexibility and efficiency to data processing, but it has also introduced new data security challenges. This paper analyzed core data security protection technologies in cloud computing, such as data encryption, access control, and privacy protection algorithms, and explored solutions to these challenges. By incorporating practical case studies, multi-layered protection and dynamic algorithm optimization strategies were proposed, which enhance data security while ensuring the efficiency and availability of cloud computing. As technology continues to advance and emerging technologies integrate, cloud computing data security protection will further evolve, providing stronger data security assurances.

REFERENCES

[1] Yang, Pan, Naixue Xiong, and Jingli Ren. "Data security and privacy protection for cloud storage: A survey." Ieee Access 8 (2020): 131723-131740.

[2] Pawar, Ankush Balaram, Shashikant U. Ghumbre, and Rashmi M. Jogdand. "Privacy preserving model-based authentication and data security in cloud computing." International Journal of Pervasive Computing and Communications 19.2 (2023): 173-190.

[3] Mohammed, Shameer, et al. "A new lightweight data security system for data security in the cloud computing." Measurement: Sensors 29 (2023): 100856.

[4] Sauber, Amr M., et al. "A new secure model for data protection over cloud computing." Computational Intelligence and Neuroscience 2021.1 (2021): 8113253.

[5] Wang, Zhiying, et al. "An empirical study on business analytics affordances enhancing the management of cloud computing data security." International Journal of Information Management 50 (2020): 387-394.

[6] Tai, Wei-Liang, Ya-Fen Chang, and Wen-Hsin Huang. "Security Analyses of a Data Collaboration Scheme with Hierarchical Attribute-based Encryption in Cloud Computing." Int. J. Netw. Secur. 22.2 (2020): 212-217.

[7] Bo S, Zhang Y, Huang J, et al. Attention Mechanism and Context Modeling System for Text Mining Machine Translation [J]. arXiv preprint arXiv:2408.04216, 2024.

[8] Sood S K. A combined approach to ensure data security in cloud computing [J]. Journal of Network and Computer Applications, 2012, 35(6): 1831-1838.

[9] Wang Z, Yan H, Wei C, et al. Research on Autonomous Driving Decision-making Strategies based Deep Reinforcement Learning [J]. arXiv preprint arXiv:2408.03084, 2024.

[10] Baek J, Vu Q H, Liu J K, et al. A secure cloud computing based framework for big data information management of smart grid [J]. IEEE transactions on cloud computing, 2014, 3(2): 233-244.